DISCOVERY OF CD8 T CELL EPITOPES IN VIRAL DISEASES EMPLOYING CONDITIONAL LIGANDS FOR MURINE MHC AND ASIAN HLA VARIANTS

CHANG XIN LEI CYNTHIA

NATIONAL UNIVERSITY OF SINGAPORE

2012

DISCOVERY OF CD8 T CELL EPITOPES IN VIRAL DISEASES EMPLOYING CONDITIONAL LIGANDS FOR MURINE MHC AND ASIAN HLA VARIANTS

CHANG XIN LEI CYNTHIA

(B.Sc.(Pharm)(Hons.), NUS)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

NUS GRADUATE SCHOOL FOR INTEGRATIVE

SCIENCES AND ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2012

Declaration

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Chang Xin Lei Cynthia

11/12/2012

Acknowledgements

This thesis would not have been possible without many kind individuals who in one way or another contributed and extended their valuable assistance, and to only some of whom it is possible to give particular mention here. I would like to first express my utmost gratitude to my main supervisor, Dr. Gijsbert Grotenbreg, for his guidance, expertise, patience and support for the past few years. He has been pivotal to the realization of the project to develop novel conditional ligands for Asian HLA. Dr. Grotenbreg was generous to share his time, ideas and experiences and he had constantly encouraged me throughout the study. I am very grateful that he had taken me in as a graduate student, and I cannot thank him enough. I would also like to thank Prof. Antonio Bertoletti for his supervision, and generous support. He kindly provided human clinical samples, preliminary data for the study and allowed me to work in his laboratory. I had gained much experience in T cell culturing, T cell related assays and analysis as well as novel insights related to the project under his guidance and while working with his lab members. In addition, I thank Dr. Grotenbreg, Prof. Bertoletti and Assoc. Prof. Paul MacAry for their supervision as my thesis advisory committee.

I would also like to thank Dr. Sylvie Alonso for her provision of the dengue virus-infected AG129 mice and her supervision in the murine epitope identification project. I sincerely thank her and her lab members, Ms. Grace Tan and Mr. Jowin Ng for the preliminary data, valuable help in sample collection and guidance. The identification of novel human epitopes would not have been possible without the help of many including Assoc. Prof. Tan Yee Joo and Dr. Janice Oh for the preliminary data related to T cell responses in SARS patients and the provision of clinical samples; Dr.Adam

Gehring for his support in the HLA-C study, guidance in HLA-C restricted epitope identification and provision of clinical samples; Dr. Hoe Nam Leong for the provision clinical samples; Dr. Anthony Tan and Ms. Adeline Chia for providing preliminary ELISPOT data for HBV identification, bioinformatics analysis and guidance in many aspect of the study; Ms. Or Ming Yan and Ms. Lim Pei Yiing for their valuable support in HLA complexes refolding and HLA-stability ELISA; Ms. Toh Kai Yee for her contribution in HLA genotyping of donors; Ms. Karen Nadua for her preliminary data and support in human dengue epitope identification, and Dr. Bahar Shafaeddin Schreve and Ms. Melissa Chng for technical assistance.

I give my special thanks to all the members of the Grotenbreg lab, Bertoletti lab, and friends in the Immunology programme and NUS graduate school of integrative sciences and engineering (NGS). Their help and encouragement had made a huge difference to my life and I greatly treasure all the moments we had shared. Lastly, I sincerely thank my beloved parents, sister, and fiancé for their undying support, love and understanding.

The research was supported by the Singapore National Research Foundation Research Fellowship (NRF2007NRF-RF001-226) and an A*STAR BMRC grant (10/1/21/19/652). I also thank NGS for their financial support in my graduate studies. The funders had no role in study design, data collection and analysis, and preparation of this thesis.

Table of Contents

Summary

The synthesis of soluble recombinant Major Histocompatibility Complex (MHC) molecules is a time-consuming and labour intensive task. However, the specificities of Major Histocompatibility Complex (MHC) molecules incorporated with ultraviolet (UV)-light sensitive conditional ligands may be easily manipulated by a peptide exchange strategy, where the emptied binding-grooves of the MHC after photocleavage of the ligands are refilled with peptides of choice. Development of such conditional ligands has paved the way for high-throughput production of MHC libraries that present defined peptides at will. While MHC-based assays have been successfully adopted for CD8+ T cells epitope identification in murine models, their application in humans are limited as established reagents concentrated on human leukocyte antigen (HLA) restriction elements associated with European ancestry. For Asian variants of HLA, such tools are scarcely available. With a lack of information on epitope specificity for HLAs commonly associated with Asian population, the development of conditional ligands to target this population group is highly valuable.

In the first part of this study, conditional ligands for H-2Db and H-2Kb MHC allelic variants were employed to generate MHC tetramers libraries for dengue virus epitope mapping in a mouse model. Antigen-specific CD8+ T cell responses following a dengue virus infection with D2Y98P in AG129 mice that were deficient in both alpha/beta and gamma interferon receptors (IFN-α/βR-/- and IFN-γR-/-) were studied.

In the second part of this study, we report the design and synthesis of thirty novel irradiation-sensitive conditional ligands for HLA-A*02:03, -A*02:06, -A*02:07, -

A*02:11, -A*11:01, -A*24:02, -A*33:03, -B*15:02, -B*40:01, -B*46:01, -B*55:02, -B*58:01, -C*03:04, -C*04:01, -C*07:02, and -C*08:01 alleles, thereby targeting the technology towards South East Asian (SEA) populations to provide 93%, 63% and 79% for HLA-A, -B and –C respectively. UV light-sensitive ligands were incorporated into soluble HLA products, and a unique ligand that gave the desired complex in sufficient yield was identified for each of the sixteen HLA types. UV irradiation-induced peptide exchange was accomplished for all variants as demonstrated by the rescue of complex disintegration in an ELISA-based MHC stability assay. HLA tetramers with redirected specificity were able to detect antigen-specific CD8+ T cell responses against a variety of infections such as human cytomegalovirus (HCMV), hepatitis B (HBV), dengue virus (DENV) and epstein–barr virus (EBV).

In the last part of the study, the potential of this Asian population-centric HLA library was demonstrated with the characterization of seven novel T cell epitopes from severe acute respiratory syndrome coronavirus (SARS-CoV), HBV and DENV. Post-hoc analysis revealed that the majority of these T cell responses could only have been identified with our unbiased epitope discovery approach. This flow cytometry-based immunotechnology therefore holds considerable promise for monitoring clinically relevant antigen-specific T cell responses in human populations of distinct ethnicity.

List of Tables

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Appendices

List of Symbols

λ	Wavelength
ABTS	2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid)
ANN	Artificial neural network
Anp, J	3-amino-3-(2-nitro)phenyl-propionic acid group
APC	Allophycocyanin
ATP	Adenosine-5'-triphosphate
β2m	Beta-2 microglobulin
B-LCL	B lymphoblastoid cell line
BrdU	Bromodeoxyuridine
BSA	Bovine serum albumin
C	Core protein
CD	Cluster of differentiation
CDR	Complementarity-determined region
CEPP	Consensus prediction program
CFSE	Carboxyfluorescein succinimidyl ester
CTL	Cytotoxic T lymphocytes
CSC	Cytokine secretion and cell surface capture
CXCR	C-X-C chemokine receptor
DCM	Dichloromethane
DENV	Dengue virus
DHF/DSS	Dengue haemorrhagic fever/ dengue shock syndrome
DIC	Differential interference constrast
DIPEA	N,N-Diisopropylethylamine

DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
E	Envelope
EBV	Epstein-barr virus
E.Coli	Escherichia coli
ELISA	Enzyme-linked immunosorbent assay
ELISPOT	Enzyme-linked immunosorbent spot
ER	Endoplasmic reticulum
ERAP	Endoplasmic reticulum-resident aminopeptidase
FCS	Foetal calf serum
FITC	Fluorescein-5-isothiocyanate
Fmoc	Fluorenylmethyloxycarbonyl
FP	Fluorescence polarization
FPLC	Fast protein liquid chromatography
HBV	Hepatitis B virus
HCMV	Human cytomegalovirus
HCV	Hepatitis C virus
HIV	Human immunodeficiency virus
HLA	Human leukocyte antigen
HMM	Hidden markov model
HMPB	4-(4-hydroxymethyl-3-methoxyphenoxy)-butyric acid
HOBt	Hydroxybenzotriazole
HPHPLC	High performance liquid chromatography
HPV	Human papillomavirus
HRP	Horseradish peroxidase

ICS	Intracellular cytokine staining
IEDB	Immune epitope database
IFN-α/βR	Interferon-alpha and beta receptor
IFN-γ	Interferon-gamma
IFN-γR	Interferon-gamma receptor
IPTG	isopropyl β-D-thiogalactopyranoside
IL	Interleukin
i.p.	intraperitoneal
LAMP-1	Lysosomal associated membrane protein-1
LB	Lysogeny broth
LC/MS	Liquid chromatography/mass spectrometry
LDH	Lactate dehydrogenase release
MACS	Magnetic cell sorting
MBHA	4-Methylbenzhydrylamine
MES	2-(N-morpholino)ethanesulfonic acid
MHC	Major histocompatibility complex
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
MWCO	Molecular weight cut-off
NaEDTA	Sodium ethylenediaminetetraacetic acid
NMP	N-Methyl-2-pyrrolidone
NS	Non-structural protein
NP	Nucleoprotein
OD	Optical density
PB	Pacific Blue

PBMC	Peripheral blood mononuclear cell
PBS	Phosphate buffered saline
PCR	Polymerase Chain reaction
PD-1	Programmed death-1
PE	Phycoerythrin
PES	Polyethersulfone
PFA	Paraformaldehye
PFU	plaque forming unit
pH	potential hydrogen
pMHC	Peptide-Major histocompatibility complex
pHLA	Peptide-Human leukocyte antigen
p.i.	Postinfection
prM	Membrane glycoprotein precursor
PTM	Post translational modifications
PyBOP	(Benzotriazol-1-yl-oxy)tripyrrolidinophosphonium hexafluorophosphate
qRT-PCR	Quantitative real-time polymerase chain reaction
RBC	Red blood cells
RPMI	Roswell park memorial institute
SARS-CoV	Severe acute respiratory syndrome-coronavirus
SEA	South East Asia
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SPSS	Solid phase peptide synthesis
SVM	Support vector machine
TAP	Transporter associated with antigen processing

TCEP	Tris(2-carboxyethy)phosphine
TCR	T cell receptor
TFA	Trifluoroacetic acid
TIPS	Triisopropylsilane
TIS	Triisopropyl silane
TNF-α	Tumor necrosis factor-alpha
UV	Ultraviolet
Vβ	Variable-beta

Publications

Chang CX, Tan A, Or MY, Toh KY, Lim PY, Chia A, Froesig T, Nadua K, Oh HL, Leong HN, Hadrup SR, Gehring A, Tan YJ, Bertoletti A, Grotenbreg GM. Conditional ligands for Asian HLA variants facilitate the definition of CD8 T cell responses in acute and chronic viral diseases. (Submitted)

Chang CX, Dai L, Tan ZW, Choo JA, Bertoletti A, Grotenbreg GM (2011) Sources of diversity in T cell epitope discovery. Front Biosci 17: 3014-3035

Chew SL, Or MY, Chang CX, Gehring AJ, Bertoletti A, Grotenbreg GM (2011) Stability screening of arrays of major histocompatibility complexes on combinatorially encoded flow cytometry beads. J Biol Chem 286: 28466-28475

Oh HL, Chia A, Chang CX, Leong HN, Ling KL, Grotenbreg GM, Gehring AJ, Tan YJ, Bertoletti A (2011) Engineering T cells specific for a dominant severe acute respiratory syndrome coronavirus CD8 T cell epitope. J Virol 85: 10464-10471

1

Background: T cell epitope discovergy

1.1 The CD8 T cell receptor and Class I Major Histocompatibility Complex

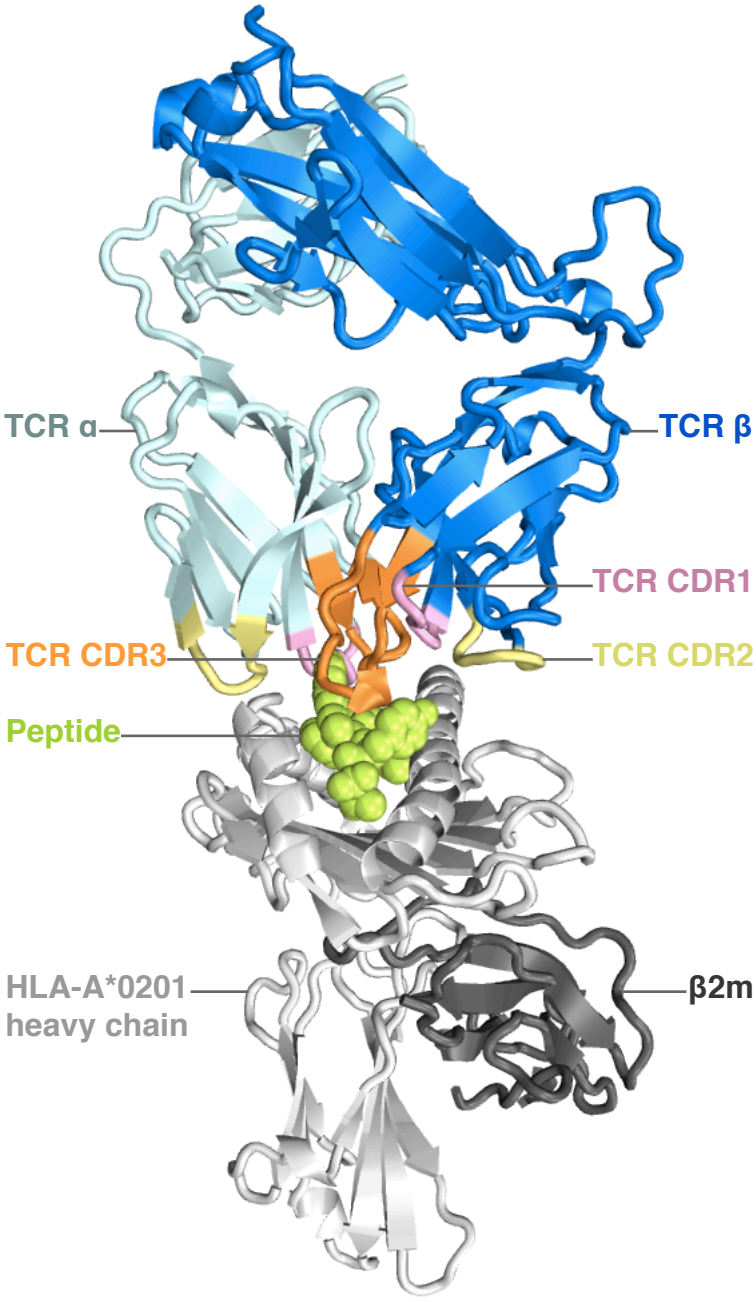
The immune system consists of a complex interplay of various humoral and cellular immune effectors that works to combat infections and malignancies, as well as to develop immunological memory for future protection. Following exposure to foreign antigens, both T and B lymphocytes, as part of the adaptive immune system respond to eliminate infected cells while ensuring proper regulation of immune homeostasis (1). Terminally differentiated B-lymphocytes, also known as plasma cells, release antibodies that directly recognize intact foreign antigens. In contrast, T lymphocytes use their T cell receptors (TCRs) to survey antigenic fragments presented by specialized molecules on the surface of infected cells (1). An infectious insult activates antigen-specific CD8-positive T cells possessing an appropriately matching TCR to transiently undergo clonal expansion and effectuate the clearance of the infected cells through cell-mediated cytolysis and release of cytokines such as interferon-gamma (IFN-γ) (2, 3). Numerous lymphocyte populations with structurally diverse and clonally distributed antigen-specific receptors are produced to deal with the heterogeneity of presented antigens not encountered by the immune system previously (4). A large repertoire of TCRs is therefore essential to ensure that at least one circulating T cell has the ability to recognize the foreign antigen. In human and in

mice, the quantity of distinct TCRs of naïve T cell precursors are estimated to be ~10^6 and ~10^7 respectively (5, 6). The generation of such a diverse TCR repertoire is accomplished through somatic recombination in the early stages of heterodimeric TCR production. The variable (V), joining (J) and diversity (D) gene segments are randomly recombined to the constant region (C) to form the β-chain of the complex, whilst the α-chain undergoes V-J-C genetic recombination. Inaccuracy in the process of joining the gene segments, particularly the variable addition or subtraction of nucleotides, ensures additional diversity (7, 8). In the resulting type I membrane protein complex, the extracellular variable and constant domains for both the disulfide-linked α- and β- or γ- and δ-chains fold in immunoglobulin-like structures, that are followed by a transmembrane and a cytoplasmic domain (9). Even though the hypervariability in the junctions of the α- and β-variable domains is at the basis of TCR diversification, these complementarity-determining regions (CDR) have a relatively modest affinity (~1–100 μM) for the presented antigen (**Figure 1.1**) (9, 10).

The counter structure recognized by the TCR is a composite surface of the major histocompatibility complex (MHC) presenting a peptide at the surface of the antigen-presenting cell (**Figure 1.1**). For CD8-positive T cells, they respond to endogenously processed peptides of allogeneic (foreign) sources, also known as antigens or epitopes, which are associated with class I MHC, a heterotrimeric glycoprotein complex with characteristic immunoglobulin folds. This class I MHC molecule consists of a polymorphic membrane-anchored heavy chain (~43kDa) that is non-covalently associated with an invariant light chain (β2-microglobulin, β2m, ~12kDa) and a peptide ligand (9). The heavy chain can be subdivided into three domains (i.e.

Figure 1.1

Crystal structure of complex between TCR and class I HLA

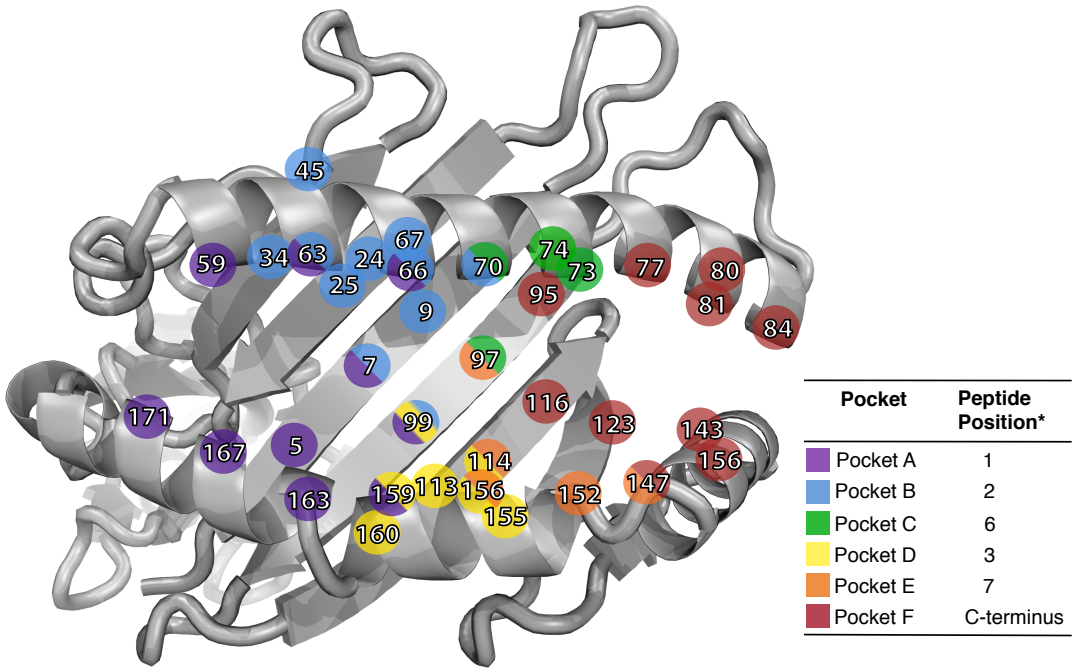


The topology of the human TCR B7 (top) interacting with the HLA-A*02:01 molecule (bottom) containing Tax peptide (green spheres) is shown as a protein cartoon representation of the side view (RCSB protein data bank:1BD2). The CDR3 loops from both TCRα and TCRβ contact mainly the peptide, while the CDR2 and CDR1 loops interact with the HLA residues.

α1, α2, and α3), where the α1 and α2 domains form an antigen-presenting superdomain capable of binding small peptide fragments (~1kDa), and the α3 domain forms an association with the β2-microglobulin. Peptides being presented by the class I MHC, are usually limited to 8 to 11 residues to allow both their N- and C-terminal ends to remain firmly embedded in the closed MHC binding groove (9). This peptide-binding domain of the MHC, which is composed of a beta-sheet platform traversed by two antiparallel alpha-helices that form the sides of the grooves, imposes a restriction for peptide association. Amino acids lining the six distinct binding pockets (Pockets A-F, see **Figure 1.2**) in this groove define the exact geometry and chemical environment for interaction and greatly affects peptide binding complimentarity (11). Peptides displayed to the immune system on class I MHC are typically derived from cellular proteins that have escaped complete destruction during normal intracellular protein turnover, where degradation of short-lived or defective proteins serves to maintain cellular protein homeostasis (12-14). The first step in the generation of antigenic peptides involves the posttranslational modification of a target protein substrate with an ubiquitin chain, which serves as a molecular signal for rapid degradation by the proteasome. The fraction of peptides that survives can be processed further by cytosolic proteases into even smaller peptides or their constituent single amino acids, or be translocated into the ER by the Transporter Associated with antigen Processing (TAP) complex. N-terminally extended precursors can then be trimmed to optimal length by endoplasmic reticulum (ER)-resident aminopeptidases (ERAP) before being loaded on a recipient class I MHC, thus releasing the final complex from its surrounding chaparones including tapasin, calreticulin, and the oxidoreductase ERp57, and initiating its migration through the Golgi to the plasma membrane (12, 15).

Figure 1.2

Peptide binding groove of HLA-A2 molecule



Supertype	B pocket specificity		F pocket specificity	
	Description	**a.a. residues****	**Description**	**a.a. residues****
A01	Small and aliphatic	ATSVLIMQ	Aromatic and large hydrophobic	FWYLIM
A02	Small and aliphatic	ATSVLIMQ	Aliphatic and small hydrophobic	ATSVLIMQ
A03	Small and aliphatic	ATSVLIMQ	Basic	RHK
A24	Aromatic and aliphatic	FWYLIVMQ	Aromatic, aliphatic and hydrophobic	FWYLIVM
A01 A03	Small and aliphatic	ATSVLIMQ	Aromatic and basic	YRK
A01 A24	Small, aliphatic and aromatic	ASTVLIMQFWY	Aromatic and large hydrophobic	FWYLIM
B07	Proline	P	Aromatic, aliphatic and hydrophobic	FWYLIVM
B08	Undefined	-	Aromatic, aliphatic and hydrophobic	FWYLIVM
B27	Basic	RHK	Aromatic, aliphatic, basic and hydrophobic	RHKFWYLIVM
B44	Acidic	DE	Aromatic, aliphatic and hydrophobic	FWYLIVM
B58	Small	AST	Aromatic, aliphatic and hydrophobic	FWYLIVM
B62	Aliphatic	LIVMQ	Aromatic, aliphatic and hydrophobic	FWYLIVM

Numbers indicate the positions of amino acids of the HLA (RCSB protein data bank:1BD2) that form the various peptide binding pockets. Each pocket accommodates a certain position of the binding peptide (*). The commonly associated amino acid residues (a.a. residue) of the binding peptide at the B and F pocket for each HLA supertype is given (**). (11)

1.2 Challenges in T cell epitopes discovery

The development of novel diagnostic, prognostic and eventually immunotherapeutic applications exploiting class I MHC antigen presentation, crucially depends on the exact definition of the antigen, the MHC restriction element, and subsequently on the monitoring of the ensuing antigen-specific T cell response. Some of the major obstacles that need to be overcome for a proper description of the cellular immune response involve dealing with the multiplicity of variables that are at its basis. Specifically, the diversity of MHC molecules present in the individual host and in the population (Chapter 1.2.A), variation of epitopes arising from various cellular processing (Chapter 1.2.B), and the variety of pathogens encountered by the human immune system (Chapter 1.2.C) are physiological variables that affect cytotoxic T lymphocyte (CTL) responses and consequently, the characterization of the T cell epitopes. Furthermore, prior knowledge of T cell epitopes and bioinformatic prediction algorithms (Chapter 1.2.D), together with plethora of technologies that immunologists currently have at their disposal for T cell epitope mapping (Chapter 1.2.E), contributes to the complexity in the choice of tools for an accurate definition of novel antigens.

1.2.A. Human Leukocyte Antigens (HLA) diversity

In humans, MHC molecules are designated as Human Leukocyte Antigens (HLAs). The HLA gene locus, located in the 6p21 region of human chromosome 6, contains ~32 thousand known protein coding sequences, including the genes encoding the class I and class II HLA complexes (16). Typically, exogenously acquired antigens are presented by class II HLA complexes, while endogenously processed antigens are

presented by class I HLA complexes. These HLA molecules are encoded by polygenic and exceptionally polymorphic gene families. The polygenicity is derived from clusters comprising classical genes (class I HLA-A, -B and –C; class II HLA-DP, -DQ, -DR), non-classical genes (class I HLA-E, -F, -G; class II HLA-DM and –DO) and various pseudogenes, and allows an individual to simultaneously present multiple antigens (17). Maternally and paternally inherited alleles add the final layer of HLA diversity, with a high probability for heterozygosity within offspring given the polymorphic nature of the HLA loci, thereby possibly providing a survival advantage against pathogens (18-21). The allelic polymorphisms for each of those genes is primarily restricted to those residues that line the peptide-binding groove of these molecules, thereby largely defining the peptide-binding and presentation repertoire of the individual gene product. As per May 2012, the IMGT database lists the sequences of over 6800 class I and 1647 class II HLA molecules, with the highest polymorphism found in the classical HLA genes with 1884 HLA-A, 2490 HLA-B, 1384 HLA-C and 1194 HLA-DRB alleles, and these numbers are still steadily increasing (22). When clustered either by ethnicity or geographical location, prominent distributions of HLA allelic variants can be discerned (23). HLA-A*02:01 is commonly found in Europeans (~27%) and north Americans (~20%), but is contrastingly low in southeast Asians (~7%) and southwest Asians (~12%). On the other hand, alleles like HLA-A*11:01, HLA-A*24:02 and HLA-B*40:01 are widespread in southeast Asian populations (~21% of HLA-A*11:01; ~30% of HLA-A*24:02, ~17% of HLA-B*40:01) whereas most of these HLA variants are rare in Europeans (7% of HLA-A*11:01; 8% of HLA-A*24:02, 5% of HLA-B*40:01) and north Americans (1% of HLA-A*11:01, 2% of HLA-B*40:01) (24). These are just a few examples for HLA frequency differences across populations, and the consolidated

data is publicly available in online repositories listed in **Table 1.1**. On an individual level, disease susceptibility can be directly influenced by these genetic variations as minor differences in the constellation of expressed HLA alleles can have crucial consequences during challenges to the immune system by infectious disease, and has been shown in dengue virus (DENV) (25), human immunodeficiency virus (HIV) (26, 27), hepatitis B and C virus (HBV and HCV) (28, 29), and human papillomavirus (HPV) (29). As HLA diversity is a critical mediator in the generation of specific T cell responses, an accurate characterization of epitopes should entail the identity of the associated HLAs. Furthermore, such descriptions will likely provide a basis of understanding of the requirements for developing protective immunity against these pathogens in various human populations.

Through the vast repertoire of allelic variants for HLA molecules, each of them capable of binding distinct sets of peptides, a mechanism is created to deal with the large diversity of antigens for which the prior experiences within host's immune system may not have exist. At the same time, significant cross-reactivity in peptide binding to different HLA variants has been observed. For example, it was found that the HLA molecules B*35:01, B*35:02, B*35:03, and B*54:01 share a consensus motif (30). Therefore, a methodology was proposed to cluster HLAs that bind largely overlapping collections of peptides into supertypes, such as the B7 supertype for the above case. A major driving factor behind this reductionists approach was the desire to design and develop subunit vaccines, which are non-infectious formulations consisting only of the immunogenic determinants without any irrelevant component. Vaccination with the antigen alone holds some inherent advantages, as it circumvents the safety concerns associated with whole pathogen vaccination schemes, boost the

Table 1.1

Online available MHC databases.

Database	Comment	URL
Allelefrequencies	Allele frequencies in human populations	http://www.allelefrequencies.net/
dbMHC	Platform for DNA and clinical data related to the human MHC	www.ncbi.nlm.nih.gov/projects/gv/mhc/
dbMHC anthropology	Allele frequencies in human populations	http://www.ncbi.nlm.nih.gov/gv/mhc/ihwg.cgi
IMGT	Genome, sequence and structure immunogenetics data.	http://www.imgt.org/
IMGT/HLA Database	Database for sequences of the HLA including the official sequences for the WHO Nomenclature Committee For Factors of the HLA System	http://www.ebi.ac.uk/imgt/hla/
HLA sequence data	Static database for HLA sequences	http://hla.alleles.org/data/index.html
ImmPort	DNA and protein sequences, sequence features, and population frequencies of MHC alleles	www.immport.org/
PyPop	Software for large-scale population genetic analyses	www.pypop.org/

These databases mainly provide information on HLA allele frequencies in populations and HLA sequences for each HLA variant. The name of the databases, major contents and URL are listed as shown.

efficiency with a concentrated delivery of antigens, and should provide control over the direction and breadth of the immune response. The conceptual reduction in complexity introduced through HLA supertypes can be a valuable tool in the search for cross-reactive antigens capable of providing appropriate protection over a heterogeneous group of HLAs. The initial classification, which so far is limited to allelic variants of HLA-A and –B, identified four HLA supertypes based on their ability to bind similar peptide sequences: their corresponding supermotifs (31). Detailed inspection of the pocket architecture, enabled largely through the availability of crystallographic data on a variety of MHC molecules (32-36), has revealed that peptide binding specificity is governed primarily by the physiochemical properties of the B and F binding pockets in coupled fashion (**Figure 1.2**). This scheme was thereafter expanded to nine HLA supertypes to accommodate the majority of HLA molecules identified at that time. The range of supertypes – A1, A2, A3, A24, B7, B27, B44, B58, B62 – was broadly inclusive of the HLA variants found in most ethnic groups worldwide (37). The refinement and development of novel immunotechnologies has continued to produce an abundance of novel T cell epitopes and spurred a third update and revision of the classification, yet in its most recent incarnation, no new HLA supertypes were identified (11). Moreover, based on overlapping peptide repertoires, some HLAs appeared not to belong to a single supertype but rather required to be categorized in several supertype families. In light of the polymorphic nature of HLA, this promiscuity in peptide binding is not difficult to envisage. Although the notion of HLA supertypes has been applied in a variety of disease settings, and has been evaluated as a diagnostic and prognostic parameter (11, 31, 37), it has its basis in peptide-MHC binding. While the HLA supertype classification provides a system of extrapolating the binding capacity of peptides to

HLAs of the same supertype family, it does not provide evidence for TCR interaction with the entire peptide-HLA complex and therefore limits its application in epitope identification. Future studies shall provide insight on whether the concept of clustering polymorphic HLA according to degeneracy in T cell recognition will continue to provide added value.

1.2.B. Epitope diversity

It is becoming increasingly clear that the diversity of peptides that are subjected to immune surveillance greatly exceeds those that can be predicted from the coding capacity of the genome alone. Mechanisms involved in protein translation and post-translational modifications (PTMs) extend the sequence diversity beyond peptides derived by normal cellular protein turnover, potentially expose the immune system to a byzantine repertoire of antigens and thereby profoundly impact epitope discovery efforts.

The term cryptic peptides, which refers to their atypical origin, is usually reserved for those peptides that arise from translation of sequences other than the primary DNA encoded open reading frame (13, 14, 38-41). Such peptides may emerge from aberrant posttranscriptional regulation events such as retention of intronic sequences in the final transcript (42), incomplete splicing such that novel epitopes are located across an exon-intron junction (43), or exon extension with the concomitant introduction of novel transcription initiation sites (44). The translation process itself provides yet another mechanism for generating unconventional gene products by using alternative sites of initiation. For example, the ribosome is able to bypass conventional initiation codon for a preferred site downstream of an alternative frame, in a process known as

initiation codon scan-through (45, 46). Moreover, the ribosome can commence translation from non-canonical codons (47-49), thus decoding a leucine residue instead of methionine as the first amino acid of the translated product, or include "untranslated" regions on either the 5'- or 3'- end of the transcript (48, 50). Epitopes derived from alternative reading frames have been observed repeatedly in malignancies (51) and in autoimmunity (52), as well as in T cell responses against pathogens with a high rate of nucleotide insertions and deletions such as influenza (45, 46) and HIV (53, 54). Remarkably, evolutionary pressure imposed by the immune response has been shown to drive viral adaptation to generate escape mutants of cryptic peptides, whilst leaving the coding sequence of the original reading frame unaffected, illustrating their relevance to immunity (53).

A variety of post-translational modifications (PTMs) contributes to protein stability and regulates their function, for example by determining the activation state of an enzyme (55). Changes in PTMs may occur during fundamental cell signaling events, inflammation, cellular transformation, and apoptosis, which in turn result in immune recognition of the modified proteins or fragments thereof (56-58). Antigenic peptides presented both by class I and class II MHC have been decorated with additional chemical moieties at their termini as well as on their side chains, with examples ranging from acetyl- (59), cysteinyl- (60), glycosyl- (61), methyl- (62), nitro- (63) to phosphate (64) groups. Alternatively, the chemical nature of the epitope side chains can be enzymatically converted in processes such as deamidation (65, 66) through the action of N-glycanases or transglutaminases, citrullination (40, 67), asparagine-bond isomerisation (68), or cysteine oxidation to form intramolecular disulphide bonds (69). Not only does chemical alteration of potential epitopes direct the T cell

repertoire structurally by influencing peptide transport, peptide binding to chaperones and the MHC, as well as pMHC-TCR affinity, they also directly influence their processing (70, 71). For instance, deamidation of asparagine residues, occurring either spontaneously or during the removal of N-linked glycosyl groups, renders the sequence less vulnerable to proteolytic degradation by asparagine endopeptidases (72). Protein (or peptide) splicing is arguably the most extreme form of PTM, and entails a primary amino acid sequence going through a series of proteolytic and peptide ligation events (73). This posttranslational "cut-and-paste" reaction yields noncontiguous peptide sequences effectively disrupting the genome coding sequence. In unicellular organisms, numerous examples of self-catalyzed protein splicing have been described where a polypeptide segment is excised from the original sequence (an intein) and subsequently intramolecularly conjugated to a flanking sequence (the extein) without the help of proteases or other auxiliary factors (74). A report by Hanada *et al* revealed that T cell epitopes could indeed be generated through protease-catalyzed protein splicing, thus highlighting that non-genetically encoded peptides could add significant complexity to epitope identification (75). Further studies cemented this notion by pinpointing the proteasome, a central player in the generation of antigenic peptide fragments, as the catalytic center where sequence scrambling of the peptide could occur in both linear and reverse order (76, 77). In this context, it would be intriguing to speculate about the existence of chimeric epitopes derived from multiple parent proteins being generated in similar fashion, although the likelihood of creating those in large abundance in a cellular context might be considered minimal.

Essentially, the increased variety of epitopes that cannot be inferred directly from the genome sequence because of alternative modes of translation, through natural PTM of the primary peptide sequence, or through peptide splicing events, will both qualitatively and quantitatively affect the spectrum of T cell responses and therefore impact upon both immunity and disease.

1.2.C. Pathogen diversity

Antigen-independent diversification, through the stochastic process of V(D)J gene rearrangements, provides T cells with a near limitless repertoire of randomly generated antigen receptors in anticipation of pathogens that the adaptive immune system has never encountered thus far. Knowledge of the offending microbial genome is a crucial step in describing the complexity of host's immune response to pathogen, as it essentially provides a catalogue of all the potential virulence factors and immunogens, and allows for T cell epitope discovery that are focused on translated products represented by the genome. In 1995, a landmark publication by The Institute for Genomic Research (TIGR) reported the sequencing of the first genome of a free-living organism, the bacterium Haemophilus influenzae (78). Whole genome sequencing, aided by novel sequencing technologies such as cyclic-array strategies that have been rapidly developed into commercial platforms and have proven successful in replacing the traditional Sanger biochemistry-based methods, has since dramatically accelerated (79, 80). The accumulation of genomic data on microorganisms of medical significance has generated a wealth of publicly available information, where a number of web resources providing this sequence data as well as analysis software are listed (**Table 1.2**). While these data maybe valuable for epitope

Table 1.2

Online available pathogen sequence databases.

Database	Contents	URL
DNA Database of Japan	Primary sequence database	www.ddbj.nig.ac.jp
EMBL-ENA	Primary sequence database	www.ebi.ac.uk/ena
NCBI GenBank	Primary sequence database	www.ncbi.nlm.nih.gov/genbank
EuPathdb	Metadatabase for eukaryotic pathogens	http://eupathdb.org
Gemina	Metadatabase	http://gemina.igs.umaryland.edu
Genomes Online	Metadatabase	www.genomesonline.org
Influenza Research db	Metadatabase for Influenza	www.fludb.org
Los Alamos National Laboratory pathogen research databases	Metadatabase for HIV, HCV, Influenza, oral and STD pathogens	www.lanl.gov/science/pathogens
NIAID Bioinformatics Resource Centers	Metadatabase	www.pathogenportal.org
NMPDR	Metadatabase for food-borne pathogens and STD pathogens	www.nmpdr.org/FIG/wiki/view.cgi
PATRIC	Metadatabase for all bacterial species in the selected NIAID category A-C priority pathogens list	www.patricbrc.org
VectorBase	Metadatabase for invertebrate vectors of human pathogens	www.vectorbase.org
ViPR	Metadatabase for viral pathogens	www.viprbrc.org
varDB	Antigenic variation database	www.vardb.org/vardb

The online available pathogen sequence databases are summarized above, and provides a platform for obtaining suitable genomic and proteomic data related to the pathogen for epitope screening purposes.

mapping, the enormous influx of information will continually require solid bioinformatics infrastructures for the comprehensive analysis and extraction of useful *in silico* information of pathogen diversity (81-83). Nevertheless, properly archived data can provide a useful scaffold for the annotation of multiple related genomes, and permits the comparison of whole genome sequences on the level of phylogenetically related strains and species, as well as between higher order taxonomic ranks. The degree of sequence conservation and variability between genotypic variants of pathogens also provides direct insight into their proteomes and the inherent antigenic variation in immune responses they could provoke. To cite an example of immune variation arising from pathogenic variants, it has been observed that epitope variants of the same protein location from different viral strains, such as dengue virus (84, 85), hepatitis B virus (28), Influenza A (86) or Epstein-barr virus (87), are able to elicit differential $CD8^+$ T cell responses. In view of this, the vast amount of genetic information for different pathogenic strains will present an additional challenge in the quest for comprehensive epitope mapping.

1.2.D. Epitope databases and prediction tools

Apart from databases that catalogue the complete and intact source of the antigen, MHC-binding ligands have also been extensively documented and deposited in publicly available archives (88). This was pioneered in 1995 with a relatively modest collection (as judged by current standards) of a few hundred peptide entries, mostly consisting of human ligands and T cell epitopes, that later were consolidated into the SYFPEITHI database (89, 90). Increasing demand from the immunological research community has driven the development of several databases (**Table 1.3A**) with each

Table 1.3

Online available MHC epitope resources

A.

Database	Contents	URL
AntiJen	Peptide binding information	http://www.ddg-pharmfac.net/antijen/AntiJen/antijencontacts.htm
EPIMHC	Peptide binding information	http://imed.med.ucm.es/epimhc
IEDB	Peptide binding data and T cell responses	www.immuneepitope.org
ISED	Influenza epitope information	http://influenza.korea.ac.kr/ISED2
Los Alamos HIV databases	Published HIV epitopes	www.hiv.lanl.gov
MHCBN	Published epitopes	www.imtech.res.in/raghava/mhcbn
MPID-T2	Sequence-structure-function of MHC/TCR	http://biolinfo.org/mpid-t2
SYFPEITHI	Peptide motifs and ligand binding information	www.syfpeithi.de

B.

Name	Prediction methods	Coverage	URL
BIMAS	QM	Class I MHC binding	www-bimas.cit.nih.gov/molbio/hla_bind
CBS Prediction Servers	Suite of prediction methods	Proteosome cleavage, class I and II MHC (pan-) specific binding	www.cbs.dtu.dk/services/
CEPP	Consensus prediction using published QM	Class I MHC binding	http://jura.wi.mit.edu/bioc/grotenbreg
EpiJen	QM	Proteosome cleavage, TAP transport and class I MHC binding	http://www.ddg-pharmfac.net/epijen/EpiJen/EpiJen.htm
EpiToolKit	Suite of prediction methods	Class I and II MHC binding, and Minor Histocompatibility Antigens	www.epitoolkit.org/
HLABIND	Structure-based modeling	Class I MHC binding	http://atom.research.microsoft.com/hlabinding
IEDB	Suite of prediction methods	Proteosome cleavage, TAP transport, and class I and II MHC binding	http://tools.immuneepitope.org/main/html/tcell_tools.html
Immunomedicine Group	Suite of prediction methods	Proteosome cleavage, class I and II MHC (pan-) specific binding	http://imed.med.ucm.es/Tools/index.html
KISS	Support vector machines (SVM)	Class I MHC binding prediction for alleles with few known binders	http://cbio.ensmp.fr/kiss

MAPPP	Combinations of cleavage models with QM and MMs	Proteosome cleavage, and class I MHC binding	http://www-alt.mpiib-berlin.mpg.de/MAPPP/
MHCPred	Quantitative Structure Activity Relationship (QSAR) models	TAP transport, and Class I and II MHC binding	http://www.ddg-pharmfac.net/mhcpred/MHCPred/
MOTIF_SCAN	Sequence Motifs	Search Class I and II MHC for protein motifs.	www.hiv.lanl.gov/content/immunology/motif_scan/motif_scan
PAProC	Network-based model	Proteosome cleavage	www.paproc.de
Peptidecheck	Matrices	Class I and II MHC binding	http://www.peptidecheck.org/
POPI	SVM	Class I and II MHC binding	http://iclab.life.nctu.edu.tw/POPI
PREDEP	Structure-based modeling	Class I MHC binding	http://margalit.huji.ac.il/Teppred/mhc-bind/index.html
Raghava Group	Suite of prediction methods	Proteosome cleavage, class I and II MHC (pan-) specific binding	www.imtech.res.in/raghava/
SMM	SMM	Class I MHC binding	http://zlab.bu.edu/SMM
SVMHC	SVM	Class I and II MHC binding	http://www-bs.informatik.uni-tuebingen.de/Services/SVMHC
SVRMHC	PSSM	Class I MHC binding	www.sbc.su.se/~pierre/svmhc/new.cgi
SVRMHC	SVM	Class I and II MHC binding	http://svrmhc.biolead.org/
SYFPEITHI	PSSM	Class I and II MHC binding	www.syfpeithi.de/Scripts/MHCServer.dll/EpitopePrediction.htm

Major online resources for (**A**) MHC epitopes databases and (**B**) MHC epitope prediction softwares and servers are listed above.

of them addressing particular niche areas although considerable overlap exists between them. For example, there are repositories that focus exclusively on a particular pathogen (e.g. the Influenza Sequence and Epitope Database (91), or the Los Alamos HIV databases (92)), whereas others concentrate primarily on those data sets for which complete structural characterization of the pMHC complexes is available (e.g. the MHC-Peptide Interaction Database version T2 (MPID-T2) (93)). A more recently developed online repository, The Immune Epitope Database (IEDB), also aims to collate information on antibody and T cell epitopes for humans, as well as other species such as non-human primates and rodents. Experimental data, both positive and negative, is derived and curated from published data obtained from PubMed, as well as through direct submission of large data sets by individual investigators (94, 95). The data covers several immunological domains spanning from infectious diseases (excluding HIV), allergy, autoimmunity, transplantation, to cancer. Separate categories have been created for HIV, and orphan data that does not fit in any of those domains are classified as "Other". As with any archive, the challenge comes in assessing the integrity and quality of the databases' content. Ideally, the data curation needs to apply transparent criteria for data inclusion and several levels of review to ascertain the antigenic entity, its source and the context of immune recognition, be it a description of a T cell response or qualitative and quantitative binding data for peptides to cognate MHC molecules. Also, awareness of the updating process of epitope data, and software interface upgrading of these repositories to deal with the exponential growth of immunologically relevant data will ensure obtainment of the most current data. Furthermore, when querying and extracting information from epitope databases, it is paramount to be able to make a clear distinction between data that describes binding of a peptide sequence to a particular MHC molecule, and a T

cell response against a defined peptide-MHC complex, which are fundamentally different parameters that cannot be used interchangeably.

Databases of immunologically relevant peptides and experimental details, combined with genomic data of a myriad of pathogens, have proven excellent resources for bioinformatics to develop tools to predict candidate epitopes, hence saving valuable time and resources associated with experimental T cell epitope identification. The precise technological features of methods for epitope prediction have been extensively reviewed by the groups of Reche (96), Sette (88), Ren (97), Ranganathan (98), and Nielsen (99). A limited discussion of the more commonly used bioinformatic tools that are freely available as ready-to-use internet-accessible programs (**Table 1.3B**) will be provided here. The main attraction of web-based epitope prediction programs is their easy and public access. Even though they are not ideally suited for handling large data sets and their output data is not normalized making comparison between different methods difficult, they provide an estimation of candidate epitopes based on pre-established predictive algorithms. Most methods in essence predict peptide binding to the MHC because predictive tools describing T cell reactivity for particular peptide-MHC combinations have, to our knowledge, not yet been reliably developed. The factors that influence the performance of prediction methods in terms of accuracy and sensitivity are primarily the type of data used to train the algorithm, and the nature of the predictive algorithm or model. The latter can be broadly categorized in qualitative and quantitative peptide sequence-based models, and structural modeling methods (96). The earliest attempts on predictions were epitope-data driven and comprised the qualitative description of motifs of amino acids commonly found at specific positions within the polypeptides sequence (89, 90, 100).

This was improved upon by the definition of motif matrices, where complete tables were constructed with residue-specific coefficients for each position within the peptide chain, allowing the scoring and ranking of epitopes by summing, multiplying or averaging the individual coefficients derived from a peptide sequence(90). Ever more sophisticated platforms, including hidden markov models (HMMs), artificial neural networks (ANNs) and support vector machines (SVMs), have since been developed (96, 98, 99). Many of these computational strategies have also been applied to quantitative prediction models with the distinction that biophysical affinity data for peptide binding to the MHC serves as the training information (96). The incorporation of parameters such as the likelihood of proteasomal processing or TAP transport have further added theoretical sophistication, although the benefit of their inclusion has not been conclusively demonstrated (88), and predictive methods have yet to include variables such as rates of protein expression, protein localization, proteolytic processing, or the T cell repertoire to which the peptide is presented. Finally, three-dimensional atomic structures for various MHC molecules obtained by X-ray crystallography, each detailing the exact geometry and chemical environment of the peptide binding cleft of the specific MHC, have also been exploited for epitope prediction using strategies such as protein threading, homology modeling and peptide docking (96, 98).

Side-by-side comparison of prediction methods has revealed that artificial neural networks and consensus prediction emerge as the tools of choice, however, the principle factor influencing performance remains the size of the training data (101-103). Increasingly successful mathematical models and algorithms that will continue to guide and focus epitope identification efforts will require the integration of both

novel *in silico* methods and experimentally validated data on peptide binding or T cell reactivity. It is noteworthy that there is considerable disparity in the quantity and quality of epitope-related data for different allelic variants of MHC molecules, with a strong bias towards the archetypical HLA-A*02:01 (**Figure 1.3**). The comprehensive dissection of cytotoxic T lymphocyte (CTL) responses across genetically heterogeneous human populations necessitates epitope identification efforts to focus on the less-studied yet commonly distributed MHC molecules.

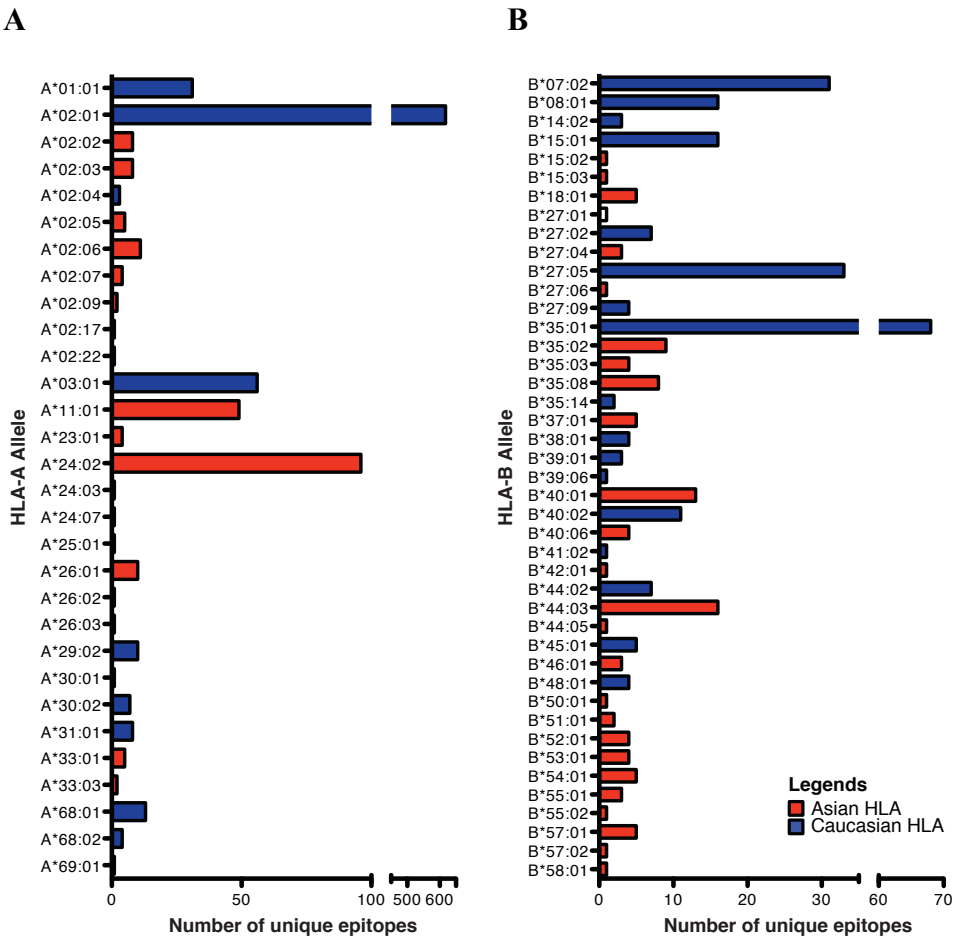
1.2.E. Immunotechnologies for epitope discovery

Many technologies have been added to the immunologist's toolbox to establish the precise identity of $CD8^+$ T cell antigenic determinants, and the majority of those have been reviewed in-depth (104-108). Here, several selected assays are highlighted (**Table 1.4**) and categorized by those that probe peptide-binding to the MHC as measured by the presence of a (labeled) ligand or by stabilization of the resulting protein complex, and assays looking at cellular responses as measured by surface expression of specific TCRs or T cell function such as cytokine secretion.

Biophysical techniques allow assessment of epitope binding and stabilization characteristics for individual MHC products. They involve either systems with cell-lines expressing particular MHC molecules, or cell-free systems that require the use of soluble MHC molecules. Common cell-lines capable of antigen presentation are the murine RMA-S cell-line (109, 110), human TAP-deficient T2 cell line (111) and EBV-transformed B cell lines (112, 113). Peptide binding assays employing these lines depend on the presence of "empty" class I MHC molecules on the cell surface to

Figure 1.3

HLA distribution of epitopes



HLA distribution of epitopes identified in IEDB for (**A**) HLA-A and (**B**) HLA-B. Epitopes data with only HLA serotype reported are omitted. Asian HLAs (red) presented here are HLA alleles found more commonly in Asians who reside in northeast, southeast or southwest asia, although the alleles may be present in the Caucasian population. For HLA-A*02:01, a total of 618 epitopes have been identified.

Table 1.4

Examples of T cell epitope mapping technologies.

Technology	Advantages	Disadvantages
• Cell-based peptide binding	• Established technology • Widely adopted technology • Robust	• Cell lines unavailable for several HLAs • Difficult to isolate sufficient quality and quantities of MHC proteins • Labor intensive
• Radio-labeled peptide binding	• Quantitative measure of peptide binding • Cell-free assay is possible	• Radioactive reagents • Indirect quantification of peptide binding • Labor intensive
• Fluorescence polarization	• Direct measurement of free and bound peptide • Real-time measurement • Fast and simple	• Requires the design and synthesis of fluorescent peptide probe • Peptide under investigation is modified with fluorescent moiety.
• MHC-stability ELISA	• Direct measurement • High throughput screening is possible	• No quantitative measurement • Multiple staining and incubation steps
• LC-MS/MS	• No prior knowledge of the antigen identity needed allowing unconventional epitopes to be identified • High sensitivity	• Only a small percentage of all possible ligands can be analyzed per run • Technically challenging
• MHC-peptide multimers	• Measures frequency of antigen-specific T cells • No requirement for *in vitro* expansion • Non-destructive for sample • Combination with other cell stainings is possible • High throughput screening is possible	• Does not measure effector function of T cells • Requires different multimer reagents for specific MHC-peptide combinations
• Intracellular Cytokine Staining	• Quantification of cytokine product • Multi-cytokine analysis is possible • High throughput screening is possible	• Identification is on based on function only • *In vitro* stimulation is required • Cell fixation is required
• T cell Specific Activation Markers	• Fast staining methods • Combination with other cell stainings is possible • High throughput screening is possible	• Some markers are not exclusive for T cell activation
• Proliferation (BrdU or CFSE)	• Measures function • Measures number of cell cycles (CFSE only) • Combination with other cell stainings is possible • *In vivo* application possible	• Cell fixation required for analysis (BrdU only) • Long *in vitro* culturing is required • Does not measure effector function
• ELISPOT	• Measures T cell function and frequency • High throughput screening is possible • Does not require *in vitro* expansion	• Measures a single cytokine • *In vitro* stimulation is required • Bystander activation not distinguishable
• Cytokine Secretion and Cell Surface Capture (CSC)	• Protein-release is quantified • Can be used for cell sorting	• Multiple cell staining steps required • Require *in vitro* stimulation • Not applicable for high throughput assays
• Cytokine Secretion and Well Surface Capture	• Protein-release is quantified • High throughput assay	• Bystander activation not distinguishable • Measures a single cytokine • Require *in vitro* stimulation • Technically challenging
• MHC-Peptide Microarrays	• Simultaneous detection of T cell binding with cytokine secretion is possible • Simultaneous detection of multiple specificities	• Dependent on T cell diffusion to their target • Technically challenging
• Qrt-PCR	• Simultaneously measurement of different cytokine expression	• No quantification on protein- or cellular level • Technically challenging

Common techniques for T cell epitope mapping, together with their advantages and disadvantages are summarized above.

enable the exogenous addition of peptides. This is achieved in RMA-S cells by maintaining them at 26°C (109, 110), in transfected T2 cell-line because peptides derived from cytosolic antigens are not presented (111, 114), and in other cell types by transfection to induce an overexpression of HLA (115) or by a mild acid treatment to remove the peptide from the complex (112, 113). In the T2 or RMA-S peptide stabilization assays, peptides are added directly to these "empty" MHCs. The surface expression levels of the stabilized complex is then probed with MHC-specific antibodies and quantified by flow cytometry (109, 116, 117). Another avenue pursued with these cell-lines are peptide competition assays, where the inhibition of binding of fluorescein-labeled (112, 113) or radio-labeled peptides (115) by non-labeled peptides-of-interest to the "empty" MHC is measured. This can be accomplished using either flow cytometry or gamma counting of metabolically labeled MHC molecules isolated from lysed cells by gel-filtration or immunoprecipitation. Peptide competition involving radioactive ligands has been demonstrated on pre-purified detergent-solubilized MHC molecules as well (118-122).

Assays involving cell-free systems that directly measure peptide-MHC binding come with a number of distinct advantages. First, the confounding effects of simultaneous surface expression of multiple MHC variants on a single cell are absent. Furthermore, less commonly studied MHCs are directly accessible even when cell-lines for those are not yet available. Finally, the difficult isolation of MHC proteins in sufficient quantity and purity from cultured cells is circumvented, as is the need for continuous maintenance of cells with the inherent potential for culture contamination. As a means of detection, fluorescence polarization has become increasingly popular in cell-free assays. Typically, a non-anchor amino acid residue in the peptide-of-interest is

substituted for a FITC-conjugated lysine (107). The ratio of free and bound peptide-FITC probe to soluble MHC can then be measured directly by fluorescence anisotropy, from which the binding characteristics of the ligand can be determined without the need for any MHC separation steps (107). Peptide competition assays using this detection method are also feasible. Moreover, this method was successfully employed together with UV-mediated peptide exchange, in which class I or II MHC was pre-loaded with a conditional ligand that fragmented upon UV-light exposure and then was exchanged for a fluorescent probe (123, 124). A similar assay has also been demonstrated for class I MHC conjugated to fluorescence beads detectable the flow cytometer (125). Finally, the capability of non-labeled peptides to stabilize the MHC following UV-mediated peptide exchange can be tested with the MHC-stability Enzyme-linked immunosorbent assay (ELISA). Peptide binding is indirectly ascertained with β2m-specific antibodies by probing for its presence or absence, as this subunit will be lost too when a replacement peptide does not adequately stabilize the complex. Subsequent colorimetric development provides an assay read-out similar to traditional ELISA (126, 127). A recent modification to the assay involves performing the MHC-stability ELISA on fluorescence beads and instead provides a readout by flow cytometry (125).

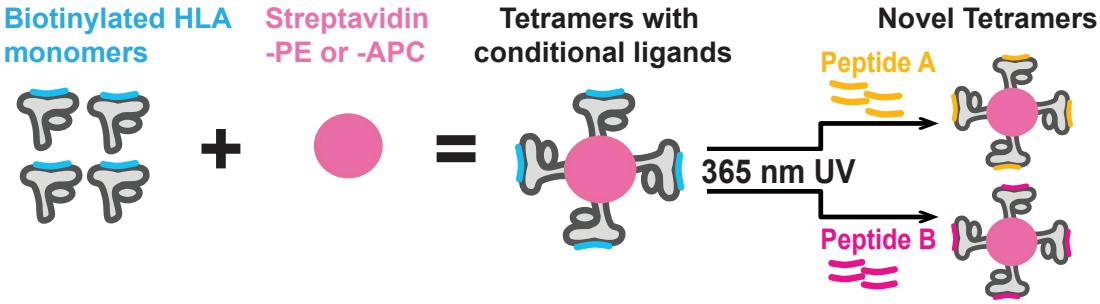
Mass spectrometry (MS) based techniques provide a bottom-up approach to identify peptides that bind to cell surface MHC. Generally, the ligands are extracted by acid treatment and purified by high-performance liquid chromatography (HPLC). The eluted peptide fractions are then analyzed for their ability to induce T cell activation, and the bioactive fractions are probed using liquid chromatography in combination with tandem mass spectrometry (LC-MS/MS) for definitive peptide identification

(128-133). One major advantage of this technology is that it can operate without prior knowledge of the input antigen's identity. Cryptic epitopes or post-translational modifications, as discussed earlier, are thus readily identified. However, the technological sophistication that is required to run such experiments often precludes its wider and more frequent application.

As T cell expansion and function is the most direct measure of the host's immune response to infection, antigenic peptides identified through T cell-based assays are invaluable. Directly analysis of T cells by flow cytometry has been achieved using various labeling strategies that seek to distinguish them by their antigen-specific TCR, their function or combinations thereof. Such techniques are useful for both epitope discovery and for functional characterization of antigen-specific T cells. First described by Altman *et al*, the MHC tetramer is such an enabling technology for the visualization of antigen-specific $CD8^+$ T cells (134). Recombinantly produced MHC molecules of defined specificity are multimerized, traditionally around a streptavidin core that is endowed with a fluorophore label of choice, to increase the avidity of the combined complex without loss of specificity (135-138). Recent advances to further expand the possible ways of reading MHC tetramer-stained cells include synthesizing MHC tetramers with streptavidin coated quantum dots of various fluorescence to increase the number of fluorochromes useful for flow cytometry or with metal-ion conjugated streptavidin for Cytometry by Time-of-Flight (CyTOF) analysis (138-140). As the synthesis of tetramers is time- and labor-intensive, the construction of large tetramer libraries for epitope mapping has not been feasible until Toebes *et al* developed a peptide exchange strategy involving tetramers loaded with UV-sensitive conditional ligands (141) (**Figure 1.4**). The initial ligand bound to the MHC contains

Figure 1.4

HLA tetramers and conditional ligands



Tetramerized HLA are occupied by a synthetic photocleavable ligand in the peptide binding groove (blue). Upon UV irradiation, in the presence of putative peptide ligands (pink or yellow), the conditional ligand is fragmented and the replacement peptide produces MHC tetramers of novel specificity capable of staining individual T cell clones.

a photocleavable residue which fragments the peptide backbone after longwave UV irradiation (141, 142). The resulting complex is unstable unless a rescue peptide is provided to occupy the emptied peptide-binding groove, consequently producing a novel peptide-MHC combination (141). Adopting a similar peptide exchange strategy, chemosensitive ligands that cleave upon exposure to a chemical trigger can also be designed. Specifically, $NaIO_4$ has been employed to induce chemocleavage in a diol-containing conditional ligand (143). By using conditional ligands, the time spent with high-throughput production of MHC tetramers of distinct specificity can be vastly reduced and T cell epitope discovery has been achieved with MHC tetramer libraries for the murine MHC restriction elements H-2K^b, -D^b, and -L^d (144-147) and human HLA-A1, -A2, -A3, -A11 and -B7 (124, 138, 141, 148-150).

Epitope-identification based on T cell function can be accomplished by antigen-specific activation of T cells which results in a transient increase in cytokine expression such as interferon-gamma (IFN-γ), IL-2 and TNF-α (151). These soluble factors can be visualized by intracellular cytokine staining (ICS) after administration of protein transport inhibitors, such as brefeldin A or monensin, to prevent their release into the extracellular milieu (152-155). The accumulated intracellular cytokines can then be detected through appropriate antibody staining of detergent-permeabilized cells (156, 157). The expression of surface proteins that mark particular cellular activation states also provides evidence of antigen-activated T cells. For example, CD107, a marker for activation-induced T cell degranulation, is found on the luminal side of cytotoxic granules but becomes part of the plasma membrane upon cytolytic degranulation. Antibodies specific for CD107a and -b allow the visualization, typically by flow cytometry, of this temporary increase on the cell

surface which can occur in concert with IFN-γ release (158). Further characteristic cell surface markers for activation include increased expression of CD137 (159), CD44 (160), CD69 (160-162), the late activation marker CD25 (162, 163) or low levels of CD62L (160). T cell proliferation is routinely demonstrated using 3H-thymidine, bromodeoxyuridine (BrdU) (164, 165) and carboxyfluoresceine-diacetate–succinimidylester (CFSE). Similar to tritium-labeled thymidine, the non-radioactive thymidine analog BrdU is incorporated into the DNA during replication. As cell fixation is required to give BrdU-specific antibodies access to the nucleus, isolation of intact and vital cells for follow-up assays is not possible, and the number of cell cycles the cells have undergone can also not be measured. These limitations are not imposed when CFSE, a protein-reactive fluorescent dye, is used for tracking proliferating cells. Upon each cellular division, CFSE labeled cells lose half of their fluorescence intensity resulting in a series of spectral peaks (166). Alternatively, CFSE can be used to detect the site of antigen presentation, the origin of the antigen-presenting cells and the proliferation rate of T cells in vivo after CFSE-labeled T cells has been adoptive transferred to a recipient host (167, 168). A combination of MHC tetramer, cell activation and proliferation marker, and intracellular cytokine staining can be analyzed concurrently by multiparametric and polychromatic flow cytometry techniques, thus allowing a complete assessment of CD8+ T cell antigen-specificity, activation state and function (165, 169-171).

T cell responses, specifically the release of cytokines like IFN-γ, can also be visualized by assays other than flow cytometry such as the enzyme-linked immunosorbent spot (ELISPOT) assay (172, 173). In brief, wells pre-coated with cytokine-specific antibodies are used to capture cytokines secreted in the local

environment of sedimented T cells. Secondary antibodies are then used to probe for the presence of the captured cytokine, followed by color development steps, resulting in the occurrence of distinct spots in the well. Each of these spots corresponds to a single cytokine-producing cell when employing an appropriate dilution of cells (174). ELISPOT based on IFN-γ secretion is a mainstay method for the identification of novel epitopes (175-177) and ensuing T cell functional studies (28, 177). The Cytokine Secretion and Cell Surface Capture (CSC) is a novel assay that measures the cytokine release of T cells (178, 179). In CSC, the lymphocytes are labeled with a cytokine-specific "catch reagent", a bi-specific antibody that consists of an anti-CD45 antibody conjugated with an anti-cytokine antibody. This "catch reagent" captures the secreted cytokine onto the surface of the labeled cell. Cytokine-coated cells can then be detected with secondary anti-cytokine antibodies conjugated either to fluorochromes for flow cytometry analysis or to super-paramagnetic particles for magnetic cell sorting (MACS) enrichment (179-181). With this, CSC allows for the isolation of cytokine secreting antigen-specific T cells that can be cultured or used in subsequent experiments (179, 182, 183), and affords the possibility to isolate T cells for adoptive transfers (184-186), yet its application as high throughput assay is uncommon (106). Cytokine Secretion and Well Surface Capture (Cell-ELISA), on the other hand, allows for high throughput screening for antigenic peptides based on T cell cytokine release (187-189). Cytokines can furthermore be visualized by quantitative real-time polymerase chain reaction (qRT-PCR) to determine the cytokine mRNA expression, and has been employed as a strategy for epitope mapping (190, 191). Detection of T cell populations can also be accomplished with cellular array-based screening strategies (192-195). Various peptide-MHC dimers (193) or tetramers (194-196) and antibodies are immobilized in different "spots" to form a

microarray. When a cell suspension is layered onto the array, these "spots" capture either antigen-specific T cells (194) or the cytokines produced by T cells activated by the peptide-MHC "on-the-spot" (192, 195). Occupied "spots" are visualized by direct inspection, differential interference contrast (DIC), or fluorescence microscopy (192-196). Common issues associated with methods that circumvent flow cytometry, as described above, are that the antigen-specific T cell frequency can be difficult to ascertain and the phenotype of the responding cells cannot be determined because of a lack of simultaneous multiparametric analysis (106).

1.2.F. Perspective: current knowledge of T cell epitopes

Vaccines have provided a major contribution to the prevention of human disease. Regardless of the many success stories involving immunization strategies against small pox, polio or hepatitis B, significant challenges from infectious diseases remain. Case in point is the ongoing search for a prophylactic HIV vaccine with the recent highly publicized failed efficacy trials for a vaccine based solely on evoking a T cell response (197-199). The question as to why sufficient levels of immune protection could not be reached is still debated, but the reasons are likely to be manifold as immunity is often best obtained when founded on several layers of protection, for example, through raising a combination of neutralizing antibodies and effective T cell responses. It would, however, be presumptuous to immediately disqualify the employed strategy as a whole, as it merely illustrates our limited understanding of what shapes an effective antiviral T cell repertoire.

Novel opportunities for the development of successful therapeutics and vaccines will need to build on the accurate description of the pertinent T cell responses, and those

efforts will be confronted with the various layers of diversity as described earlier. Particular areas that in our opinion warrant systematic attention are the following: (A) There is no reason to assume that HLAs that are currently less studied (**Figure 1.3**), yet widely present in the general population, do not present a similar abundance of antigenic peptides as those select HLA molecules around which epitope identification has traditionally been centered. Focus will need to be turned towards underrepresented HLA variants in order to provide a full understanding of cellular immunity in geographically and ethnically distinct populations. (B) It has been demonstrated that host ethnicity influences the fine-specificity of the T cell response to HBV (28), HCV (200), HIV (201) and EBV (202), presumably as a result of host-pathogen co-adaptation which has driven both HLA and pathogen diversification (16, 29, 203, 204). Consequently, the HLA supertype classification, which primarily describes promiscuity in peptide binding, should be closely scrutinized for its ability to accurately describe the spectrum of protective T cell responses. (C) The shortage of identified T cell epitopes for a majority of HLA variants impacts on the confidence in corresponding predictive algorithms, as the outcome of epitope prediction for those alleles might become biased towards similar peptide motifs. Improvement of these bioinformatic methods needs to be driven by the generation of experimental data. The above issues can all be addressed directly through epitope discovery, followed by definition of T cell hierarchy and functionality, employing large and clinically well-defined disease cohorts, and will be aided by the development of novel technologies and the harmonization of data reporting. Ultimately, this will allow the proper evaluation of the importance of T cell responses for immunity against a variety of pathological challenges.

1.3 Thesis overview

This thesis aims to address the gap in knowledge of epitopes identified for HLA variants common in Asian population. While many tools are available for the accurate description of CD8 T cell epitopes, the promise provided by MHC tetramers for direct visualization and tracking of antigen-specific T cell, and for in-depth characterization of the epitopes' fine specificity and HLA restriction make it an attractive reagent to pursue. Furthermore, MHC tetramer libraries generated by MHC complexes laden with UV-light sensitive conditional ligands may be adopted in screening for antigenic peptides. We decided to utilize the MHC tetramer and conditional ligand technology in T cell research. This technology and its use in epitope identification is a central theme in the following chapters:

Chapter 2: Identification of novel CD8 T cell epitopes in the AG129 dengue mouse model - How MHC tetramer libraries generated through the use of conditional ligands may be employed in epitope discovery will be first illustrated in a dengue mouse model, which involves the combined IFN-α/βR and IFN-γR knockout mice (AG129) infected with dengue virus-2 (strain D2Y98P). The expertise used in this study will follow closely that of published literature (145-147).

Chapter 3: Construction of conditional ligands and HLA tetramer libraries for Asian HLA variants - Epitope idenfication efforts will be directed towards human infections, particularly in the South East Asian population. As the HLA tetramers were initially developed to cater to the Western population, many HLA variants common in the Asian population have been neglected which limits the use of this technology. Conditional ligands for the appropriate HLA variants common in Asian populations

have to be designed, synthesized, refolded with their corresponding HLA molecules, and properly validated for function.

Chapter 4: T cell epitope discovery in the Asian population - The conditional ligands and HLA reagents developed as described in *Chapter 3* will then be precursors to the generation of HLA libraries for measuring qualitatively the strength of peptide binding to the HLA, or for the synthesis of HLA tetramers to determine the specificity of the TCRs in T cell lines derived from patients previously infected with either hepatitis B virus, severe acute respiratory syndrome (SARS) coronavirus or dengue virus. For the first time since the development of HLA tetramers, novel epitopes restricted by HLA-B*40:01, B*55:02 and B*58:01 are identified through the use of HLA tetramer screening methodologies.

Chapter 5: Conclusion - Summarizes the key findings and points to future directions.

2

Identification of novel CD8 T cell epitopes in the AG129 dengue mouse model

2.1 Introduction: AG129 Dengue mouse model

Dengue virus (DENV) is a member of the *Flaviviridae* family and is transmitted by mosquitoes *Aedes aegypti* and *Aedes albopictus* (205). Humans infected with any one of the four DENV serotypes will present with varying clinical outcomes, from a lack of signs and symptoms to lethal complications such as dengue haemorrhagic fever and dengue shock syndrome (DHF/DSS) (206). First recognized in 1950s during a dengue epidemic in the Philippines and Thailand (207), DHF is a leading cause of hospitalization and death among children in Asia (208). Unfortunately, the absence of therapeutic drugs or vaccines in the market, coupled with the lack of full understanding in dengue-related disease pathogenesis (209), will only spell continual health and economic burdens in high-risk tropics and sub-tropical region (210, 211). For the advancement of antiviral agents against DENV and to better understand the mechanisms by which DHF/DSS is caused, various animal models have been developed for potential pre-clinical studies (212). As DENV only replicate well in humans or mosquitoes, infection in non-human primate models generally did not result in DHF/DSS (212, 213). While the attention has been shifted to the development of dengue mouse models, the search for a suitable model to reflect the clinical pathology after DENV infection was met with difficulties (213). Just like in

non-human primates, replication of DENV in wild-type mice was unfavorable (214). Interestingly, 129/Sv mice deficient in both alpha/beta and gamma interferon receptors (IFN-α/βR and IFN-γR), also known as the AG129, were found to be highly susceptible to DENV infection from all four serotypes (215-217). Nevertheless, the physiological relevance of the exhibited symptoms and the unnatural route of infection had made the use of these models questionable (214). Improving on established AG129 dengue mice models, Grace *et al* (218) isolated a non-mouse-adapted DENV-2 strain named D2Y98P that originated from a 1998 dengue-2 (DEN-2) Singapore human isolate, and found the strain to be highly infectious in AG129 mice. Infection with 10^4 plaque forming units (PFU) of D2Y98P by intraperitoneal (i.p.) administration resulted in 100% mortality at 20 days post infection (p.i.). Severe impairment of splenic architecture and vascular leakage were observed at moribund state. In addition, localized areas of damage were observed in the liver accompanied by a moderate but significant increase in systemic ALT and AST levels. Infection with an even higher viral dose 10^7 PFU resulted in 100% mortality by day 5 p.i. with severe damage in the spleen, liver and intestines (218). The use of DY298P was attractive as it was adapted in *Aedes aldopictus* cells, virulent in AG129 and induced physiologically relevant clinical symptoms. In addition, subcutaneous infection of a plaque-purified strain (D2Y98P-PP1) derived from D2Y98P was met with similar success (219)

We aimed to further understand the cellular immune response in AG129 following DENV-2 (D2Y98P) infection by examining the relevant cell populations, specifically dengue-specific cytotoxic T lymphocytes (CTL). For the study of DENV-specific $CD8^+$ T cell responses in this disease model, we chose to utilize MHC tetramers as a

tool to track these cell populations. Adopting pre-established techniques in H-2K^b and H-2D^b epitope identification and the conditional ligand technology for high-throughput production of H-2K^b and H-2D^b tetramers (146, 147), we proceeded to discover H-2K^b and H-2D^b restricted DENV CD8$^+$ T cell epitopes in AG129 mice following D2Y98P infection, and characterized the dominant DENV-specific T cell population based on cell surface marker expression. As the purified virus was unavailable at the beginning of our studies, our epitope identification efforts were directed towards both the dominant strain present in the unpurified D2Y98P and a plaque purified clone D2Y98P-PP1.

2.2 Materials and methods

2.2.A. Media and Buffers

The phosphate buffered saline (PBS) used commonly is prepared by a 10 times dilution of pre-made 10 X PBS (1st Base) and filtered using a 0.2μM Supor machV PES membrane (Nalgene). Media used for bacteria culturing are the lysogeny broth (LB) that contains 2.5% w/v Bacto™ Tryptone powder (BD) in water, and the lysogeny broth agar that contains 2.5% w/v Bacto™ Tryptone powder and 1.5% w/v Bacto™ Agar powder in water. The Red blood cell (RBC) lysis buffer (pH= 7.35) contains 0.828 % w/v ammonium chloride (Sigma-Aldrich), 0.1% w/v potassium bicarbonate (Sigma-Aldrich) and 0.003% w/v EDTA. The solution was filter sterilized through a 0.2μm filter.

Buffers used in protein inclusion bodies purification include the (1) solution buffer (pH = 8) that contains 50 mM Tris hydrochloride (Tris HCl, 1st Base), 25% sucrose (Kanto chemicals), 1mM ethylenediaminetetraacetic acid (EDTA, 1st Base), and 0.1% sodium azide (Sigma-Aldrich), (2) lysis buffer (pH = 8) that contains 50mM TrisHCl, 100mM sodium chloride (NaCl, Kanto chemicals), 1% Triton X-100 (USB Corporation), 1% sodium deoxycholate (optional), and 1mM EDTA, (3) Washing buffer with Triton (pH = 8) that contains 50 mM Tris HCl, 100mM NaCl, 1mM EDTA , 0.1% sodium azide, and 0.5% Triton X-100, (4) Washing buffer without Triton (pH = 8) that contains 50 mM Tris HCl, 100mM NaCl, 1mM EDTA , and 0.1% sodium azide, and the (5) 8M Urea buffer (pH = 6) that contains 8M Urea (Sigma-Aldrich), 25mM 2-(N-Morpholino)ethanesulfonic acid (MES, Sigma-Aldrich), and 10mM EDTA, 0.1mM Dithiothreitol (DTT, Fermentas).

Buffers used in protein refolding are the (1) refolding buffer (pH = 8) that contains 100mM Tris HCl, 400mM of L-arginine monohydrochloride (Sigma-Aldrich), and 2mM EDTA, and the (2) guanidine solution (pH = 4.2) that contains 3M guanidine hydrochloride (Sigma-Aldrich), 10mM sodium acetate (Sigma-Aldrich), and 10mM EDTA.

For the biotinylation reaction, the buffer used is prepared from two separate components and used immediately after dilution and mixing. They two components are (1) Biomix-A (10X concentration,

pH 8.3) that contains 0.5M of N,N-Bis(2-hydroxyethyl) glycine (Bicine, Sigma-Aldrich) buffer, and (2) Biomix-B (10X concentration) that contains 100mM Adenosine-5'-triphosphate (ATP, Sigma-Aldrich), 100mM Magnesium acetate (Sigma-Aldrich) and 500μM of d-biotin (Sigma-Aldrich) in water.

2.2.B. Anti-mouse antibodies

Target	Host	Clone	Conjugation	Source
Mouse CD8a	Rat	53-6.7	PB	BD Biosciences
Mouse CD8a	Rat	53-6.7	FITC	BD Biosciences
Mouse CD8a	Rat	53-6.7	PE	BD Biosciences
Mouse CD8	Rat	53-6.7	APC	BD Biosciences
Mouse CD107a	Rat	1D4B	FITC	BD Biosciences
Mouse IFN-γ	Rat	XMG1.2	PE	BD Biosciences
Mouse CD4	Rat	RM4-5	PE	BD BIosciences
Mouse CD44	Rat	IM7	FITC	BD Biosciences
Mouse CD62L	Rat	MEL-14	PE	BD Biosciences
Mouse CXCR3	Armenian Hamster	CXCR3-173	PE	eBioscience
Mouse PD-1	Rat	RMP1-30	FITC	eBioscience
Mouse CD122	Rat	TM-b1	PE	eBioscience
Mouse CD69	Armenian Hamster	H1.2F3	FITC	eBioscience
Mouse CD25	Rat	PC61.5	PE	eBioscience
Mouse LAG3	Rat	eBioC9B7W	PE	eBioscience
Mouse Vβ 2 TCR	Rat	B20-6	FITC	BD Biosciences
Mouse Vβ 3 TCR	Armenian Hamster	KJ25	FITC	BD Biosciences
Mouse Vβ 4 TCR	Rat	KT4	FITC	BD Biosciences
Mouse Vβ 5.1/5.2 TCR	Mouse	MR-94	FITC	BD Biosciences
Mouse Vβ 6 TCR	Rat	RR4-7	FITC	BD Biosciences
Mouse Vβ 7 TCR	Rat	TR310	FITC	BD Biosciences
Mouse Vβ 8 TCR	Mouse	F23.1	FITC	BD Biosciences
Mouse Vβ 8.1/8.2 TCR	Mouse	MR5-2	FITC	BD Biosciences
Mouse Vβ 9 TCR	Mouse	MR10-2	FITC	BD Biosciences
Mouse Vβ 10b TCR	Rat	B21.5	FITC	BD Biosciences
Mouse Vβ 11 TCR	Rat	RR3-15	FITC	BD Biosciences
Mouse Vβ 12 TCR	Mouse	MR11-1	FITC	BD Biosciences
Mouse Vβ 13 TCR	Mouse	MR12-3	FITC	BD Biosciences
Mouse Vβ 14 TCR	Rat	14-2	FITC	BD Biosciences
Mouse Vβ 17a TCR	Mouse	KJ23	FITC	BD Biosciences

2.2.C. T cell epitope prediction

The epitope prediction is based on scoring matrices that was previously published (220). This program is available online at http://jura.wi.mit.edu/bioc/grotenbreg.

2.2.D. Peptides

SV9-P7*, the conditional ligand for H-2D^b and H-2K^b is a gift from Dr. Grotenbreg (147). Peptides used for screening were synthesized by Mimotopes (Australia) and supplied in crude lyophilized form. These were dissolved in 80% DMSO (Signma-Aldrich) in MilliQ water containing 10mM tris(2-carboxyethyl)phosphine (TCEP, Sigma Aldrich) and stored at -20°C.

2.2.E. Generation of MHC tetramers and peptide exchange

The production of MHC tetramers was performed as described previously (127, 142, 221). The major steps included the transformation of plasmids into *Escherichia coli*, expression and purification of protein inclusion bodies, protein refolding, biotinylation and purification of MHC. The plasmids encoding H-2D^b or H-2K^b heavy chain genes engineered with a C-terminal sequence for BirA-catalyzed site-specific biotinylation, as well as plasmids encoding the murine β2-microglobulin light chain genes were gifts from Dr. Grotenbreg (147). For the transformation into plasmids into *Escherichia coli*, BL21 *E. coli* competent cells made using the $CaCl_2$ method and contained a T7 polymerase under control of the lac operon promoter were mixed with 100ng of plasmids and subjected to a heat shock of 42°C incubation for 2 mins. The selection of successfully transformed *E. coli* for protein expression was based on antibody (i.e. Kanamycin for H-2D^b, H-2K^b and murine β2m plasmids) resistant conferred by the plasmid as shown by growth of cells in LB agar containing antibody.

The MHC heavy or light chains were then overexpressed in *E. coli* BL21 following isopropyl β-D-thiogalactopyranoside (IPTG, Sigma Aldrich) induction. Typically, IPTG was added into *E. coli* grown in LB culture media when the optical density (OD) of the media reaches 0.5 to 0.6, and left to incubate in 37°C for 3 hours. For the isolation of proteins from inclusion bodies in the cells, the bacteria after IPTG induction was pelleted by centrifugation and resuspended in 13mL of solution buffer *(See 2.2A-Media and buffer)*, 2 tablets of *complete* mini protease inhibitor cocktail tablets (Roche), 200 µL of 1M DTT, 50 µL of 10 mg/mL DNAse I, 100µL of 50mg/mL Lysozyme, and 13 µL of 1M $MgCl_2$. The suspension was incubated on ice for 30 mins. 13mL of lysis buffer *(See 2.2A - Media and buffer)* was then added followed by another 15 mins incubation on ice. The suspension was sonicated for a few

cycles of 30 seconds per cycle until a homogenous mixture was obtained. 560 µL of 0.5M EDTA was added, followed by centrifugation at 120 000 rpm at 4°C. The supernatent was discarded and the pellet was washed in 10 mL of washing buffer with Triton *(See 2.2A- Media and buffer)* with 10µL of 1M DTT for a few times. This was done by suspending the pellet in the washing buffer with sonication, followed by centrifugation at 120 000 rpm at 4°C and removal of the supernatant. When a white coloured pellet was obtained, the pellet was resuspended with sonication in 10 mL of washing buffer without Triton *(See 2.2A- Media and buffer)* and 10µL of 1M DTT, followed by centrifugation at 120 000 rpm at 4°C and removal of the supernatant. The pellet was then dissolved in 10mL of 8M Urea buffer *(See 2.2A- Media and buffer)* with sonication and left overnight at 4°C. The suspension was then centrifuged at 15000 rpm for 30 min at 4°C and the supernatant containing solubilized denatured proteins was collected and stored in -80°C for further protein refolding.

The solubilized heavy and light chains were refolded with a 10-30 fold molar excess of the conditional ligand for 36-72 h. For each refolding reaction, 1 tablet of '*cOmplete'* mini protease inhibitor cocktail tablets (Roche), 62 mg of L-glutathione oxidased (Sigma-Aldrich), 308 mg of L-glutathione reduced (Sigma-Aldrich) and 12mg of 10mg/mL peptides to be used for refolding were dissolved in 200 mL of refolding buffer *(See 2.2A- Media and buffer)* using a magnetic stirrer and left to cool in 4°C. 13.4 mg of solubilized MHC light chains and 18.6 mg of solubilized MHC heavy chains were each separated into 3 portions, and each portion was mixed with 500 µL of guanidine solution *(See 2.2A- Media and buffer)* and stored in -20°C. 1 portion of both heavy and light chains was thawed and added into the refolding buffer containing peptides and stirred for at least 12 hours at 4°C before the next addition.

Soluble proteins were dialyzed into 10 mM Tris-Cl buffer (pH=8), concentrated and biotinylated with recombinant BirA. In details, the refolding mix was centrifuged at 3000 rpm for 15 mins at 4°C, and the supernatant containing the soluble proteins was concentrated to 10mL using Vivaflow-50 with a 100,000 Dalton MWCO Polyethersulfone (PES) membrane (Sartorius stedim), and then dialyzed into 10 mM Tris-Cl buffer (pH=8) using Spectral/Por® dialysis membranes with 3500 Dalton MWCO (Spectrum®labs). The dialyzed proteins were concentrated to less than 10 mL using Vivaspin™ with a 10,000 Dalton MWCO membrane (GE healthcare). The proteins containing MHC monomers were then

biotinylated using recombinant biotin ligase (BirA enzyme) in accordance to pre-established protocols by the manufacturer (Avidity). 200μL of 1X Biomix-A, 200μL of 1X Biomix-B *(See 2.2A- Media and buffer)* and 30μL of 0.5 mg/mL of BirA enzyme were added into 1.4mL of protein solution and left to incubate overnight at 4°C.

Finally, the MHC monomers were purified by fast protein liquid chromatography (FPLC) on an S200 size exclusion column into a buffer containing 20mM Tris, 50mM NaCl (pH=8) and concentrated using Vivaspin™ to approximately 1mg/mL protein concentration. For the generation of MHC tetramers, the MHC monomers were conjugated with Streptavidin-PE or -APC (Invitrogen) at a 4:1 molar ratio. The amount of Streptavidin-PE or APC to be added was split into 5 additions, with at a minimum of one-hour interval in between each addition. The MHC tetramers were stored at 4°C and valid for use within 3 months.

For the UV-light mediated peptide exchange, MHC tetramers were diluted to 40 μg/mL in PBS with 0.1% sodium azide containing 200 μg/mL of replacement peptide, followed by 15 min longwave (365 nm) UV exposure by an Ultraviolet Crosslinker (CL-1000L, UVP). After an 1h incubation, the tetramers were centrifuged at 13,000 rpm and the supernatant was used for cell staining.

2.2.F. Dengue-2 virus infected AG129 mice

D2Y98P and D2Y98P-PP1 virus propagation, viral quantitation and AG129 mice infection were performed as previously described (218, 222). Briefly, eight to nine week-old mice were administered with 10^4 PFU of heat-inactivated Dengue-2 virus strain D2Y98P or D2Y98P-PP1 in 0.4 mL of sterile PBS via the intraperitoneal (i.p.) route. Mouse splenocytes were obtained by mechanical disruption of spleen in cold RPMI 1640 medium (Invitrogen, Carlsbad, CA) through a 70μM nylon cell strainer to form single-cell suspension. The suspension was centrifuged at 2000 rpm, at 4°C for 5 minutes and the supernatant was removed. The cell pellet was resuspended in 5mL of RBC lysis buffer *(See 2.2A- Media and buffer)* and incubated on ice for 5 mins with occasional stirring. 40mL of PBS was then added to stop the lysis and the suspension was centrifuged at 2000 rpm, at 4°C for 5 minutes. The

supernatant was removed and the cells were resuspended in an appropriate volume of PBS for cell staining.

2.2.G. MHC tetramers and cell surface markers staining

Cells were washed with PBS and their viability was ascertained with LIVE/DEAD® fixable near-IR stain (Molecular Probes) prior to tetramer staining. The concentration of the LIVE/DEAD® fixable near-IR stain and the cell staining procedures used were in accordance to the manufacturer's specification. The cells were then washed with PBS containing 0.1% sodium azide and stained with 40 μg/mL of PE- or APC-conjugated MHC tetramers on ice for 20 min, with 50μL of reagents used per cell staining. Other cell surface marker antibodies, including anti-CD8 antibodies, (*See 2.2.B-List of antibodies used*) were concurrently added to cells for 15 minutes. The cells were washed and fixed with 1% paraformaldehyde (Sigma-Aldrich) in PBS. Data was acquired with a BD LSRII flow cytometer (BD biosciences) and analyzed using FlowJo (Treestar Inc).

2.2.H. *In vitro* stimulation and intracellular cytokine staining

For *in vitro* stimulation, 1 million cells were grown in 1 ml of RPMI 1640 (Invitrogen) supplemented with 10% v/v heat-inactivated fetal calf serum (Thermo Fisher Scientific). CD107a-FITC antibody (Clone 1D4B; BD biosciences) was added, followed by stimulation with 10 μg/ml peptide or 5μL of Leukocyte Activation Cocktail with BD GolgiPlug™ (contains PMA, Ionomycin & Brefeldin A; BD biosciences). After 2 hours incubation at 37°C, 5% CO_2, 10μL of BD GolgiPlug™ (contains Brefeldin A; BD biosciences) was added to cells followed by another 2 hours incubation. Cells were washed, and surface marker staining was performed as described before (*See Chapter 2.2.G-MHC tetramers and cell surface markers staining*). Intracellular cytokine staining was done using anti-mouse IFNγ-PE antibodies (Clone XMG1.2; BD biosciences) and BD Pharmingen™ Mouse Intracellular Cytokine Staining Starter Kit (BD biosciences, San Jose, CA) according to manufacturer's specification.

2.3 Results

2.3.A. Epitope prediction for Dengue 2 Virus Strains

The use of bioinformatics to aid $CD8^+$ T cell epitope identification in various infection mouse models had previously garnered positive results (145-147, 176). In the consensus prediction program (CEPP) used, four algorithms were employed to predict and rank the MHC binding capabilities of all 8-mer or 9-mer peptides within the input protein sequence. The median rank of all peptides based on these four scoring-matrix predictions were evaluated, and the top 0.5% of total ranked peptides should cover the majority of the truly antigenic epitopes (220). We therefore adopted similar strategies for our model, DENV-2 (D2Y98P) infected AG129 mice, to retrieve a set of candidate H-$2K^b$ and H-$2D^b$ epitopes for D2Y98P.

The virulent determinent of D2Y98P, named D2Y98P-PP1, was later found through further plaque purification. Injection of AG129 mice with this strain produced similar kinetic and clinical manifestations as the unpurified D2Y98P (data from Tan G.K and Alonso S., personal communification) (222). Comparison of the dominant strain in D2Y98P and D2Y98P-PP1 revealed 251 amino acid mutations throughout the proteome. As the purified virus (D2Y98P-PP1) was unavailable at the time of study, the AG129 mice used in this study were infected with the unpurified D2Y98P. As such, we performed CEPP for the proteome sequence of the dominant strain in D2Y98P and D2Y98P-PP1 to predict possible antigenic peptides based on their binding to H-2^b MHC molecules (**Appendix A.1**). A total of 3384 8-mer peptides were ranked for predicted H-$2K^b$ binding and 3383 9-mer peptides were ranked for predicted H-$2D^b$ binding for each viral strain. The top 47 8-mer peptides ranked for

binding H-2K^b and top 47 9-mer peptides ranked for binding H-2D^b were selected and synthesized for subsequent MHC tetramers screening. This covers 1.38% of top 8-mer H-2K^b or 9-mer H-2D^b binding-peptides, which in principle is sufficient to include majority of the epitopes.

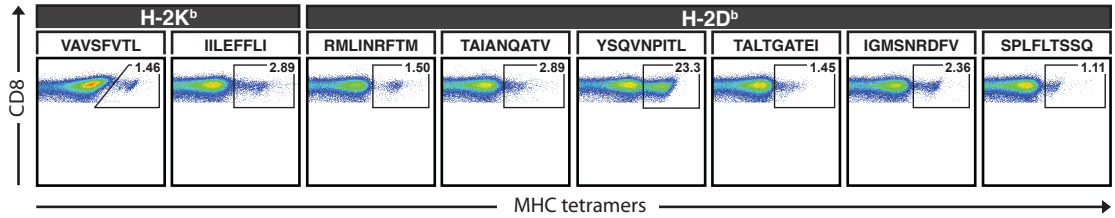
2.3.B. Discovery of CD8 T cell epitopes in the AG129 mouse challenged with Dengue virus

H-2D^b and H-2K^b tetramers containing UV-sensitive conditional ligand SV9-P7*, were used as precursor material for synthesizing staining reagents to identify H-2D^b or H-2K^b antigen-specific CD8$^+$ T cells in various mouse models of infectious disease (145-147). We opted for the same class I MHC tetramer design and peptide exchange strategy as published literature and produced 47 H-2K^b tetramers and 47 H-2D^b tetramers from the highest ranking predicted epitopes for D2Y98P (**Appendix A.1**). We also made H-2K^b tetramers for 8-mer peptides and H-2D^b tetramers for 9-mer peptides from the D2Y98P genome that contained the serine (S) residue at 178 in NS2A and phenylalanine (F) at 47 in NS4B, which were amino acid mutations that distinguished D2Y98P from the P13 strain, a variant from an earlier passage (**Appendix A.2**). Mice injected with D2Y98P rather than P13 presented with more severe clinical symptoms (Tan G.K and Alonso S., personal communication). The synthesized tetramers were used to stain splenocytes from AG129 mice at day 10 post-infection (p.i.) with 10^4 PFU of unpurified D2Y98P via the i.p. route. A total of two H-2K^b and six H-2D^b epitopes were identified by positive tetramer stainings and confirmed in independent experiments **(Figure 2.1**). We compared our results with published DENV2 epitopes based on infection of a mouse-passaged strain S221 in IFN-α/βR$^{-/-}$ mice, and found an epitope variant for NS4B$_{99\text{-}107}$ with a C99Y mutation

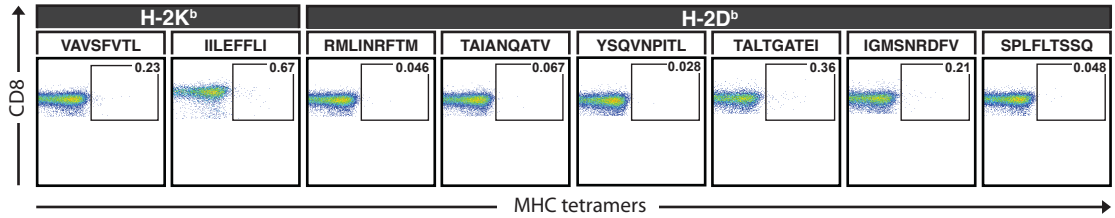
Figure 2.1

Screening for DENV2-specific CD8+ T cell

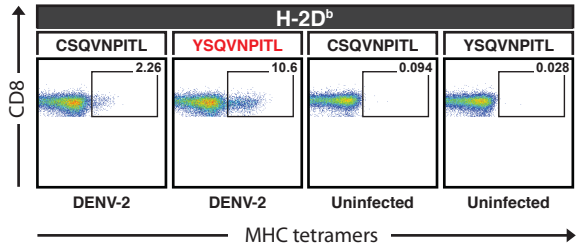
A



B



C

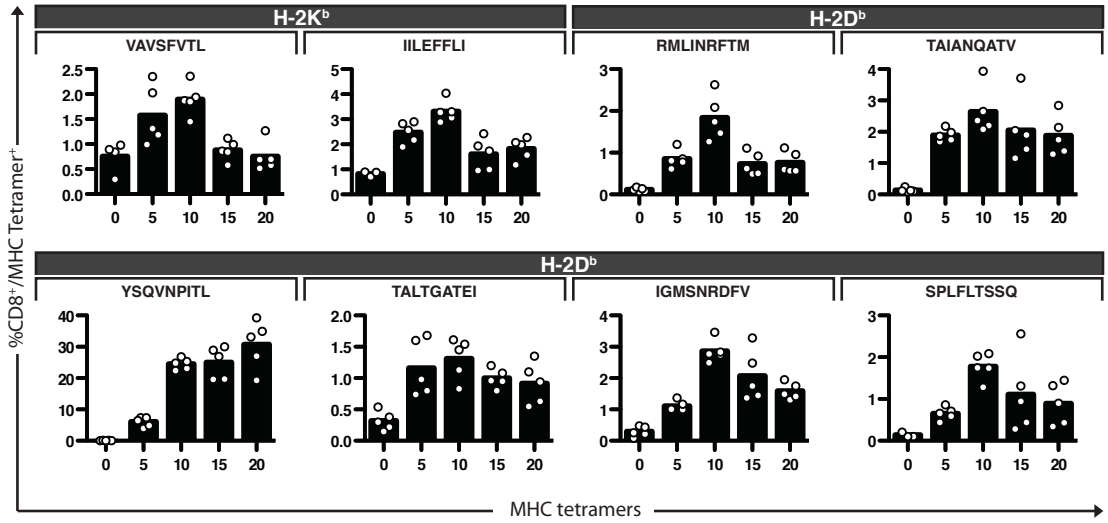


Splenocytes from (**A**) AG129 mice at day 10 post infection with 10^4 PFU of unpurified D2Y98P, (**B**) uninfected mice and (**C**) AG129 mice at day 10 post infection with 10^2 PFU of unpurified D2Y98P (DENV-2) or uninfected mice (Uninfected) were stained with H-2K^b and H-2D^b tetramers to identify antigen-specific CD8$^+$ T cell populations. Cells shown were pre-gated for CD8$^+$ and viability, and then gated for MHC tetramer (black box). The percentages of tetramer-positive cells out of the total CD8$^+$ T cell population are shown in the top right hand corner of each plot. MHC tetramer staining of uninfected mice (**B**), and MHC tetramer staining of infected AG129 mice for other tested peptides that did not yield any MHC tetramer positive population (data not shown) were used as references to determine the gates used to distinguish cells stained positively with MHC tetramers (**A**). Data shown represent stains from one of five mice.

(i.e. **C**SQVNPITL or **Y**SQVNPITL). The staining efficacies of tetramers made with **Y**SQVNPITL were compared to that of **C**SQVNPITL, and they were found to reveal a 4.6-fold higher staining of tetramers positive cells (**Figure 2.1C**). By conducting MHC tetramer stainings of these identified antigen at various intervals post DENV-2 infection, the frequencies and kinetics of the antigen-specific T cell population per epitope were analyzed. The T cells for most epitopes expand maximally at day 10 p.i. and the responses wane afterward until time of death at day 20 p.i. (**Figure 2.2**). This correlates with the kinetics of systemic IgM titre of D2Y98P-infected AG129 mice (218). However, the level of YSQVNPITL-specific $CD8^+$ T cells increased steadily even after 10 days p.i. and peaks at time of death with the highest response recorded to be 39.2% of total $CD8^+$ T cells (**Figure 2.2**). A functional analysis of $CD8^+$ T cells from DENV-2 infected AG129 mice showed that they were capable of producing IFN-γ and upregulating CD107a, a degranulation marker after 4 hours *in vitro* stimulation with the identified antigenic peptides (**Figure 2.3**). While for most antigens the percentage of $CD8^+$ T cells that upregulated CD107a upon antigen stimulation correlated well with the amount of antigen-specific $CD8^+$ T cells (**Figure 2.1**), there were a few exception (i.e. IILEFFLI, RMLINRFTM, SPLFLTSSQ) where a much larger population of CD8+ T cells upregulated CD107a (i.e. 23.3%, 17.6%, 31.0% respectively) or produced IFN-γ (i.e. 4.4%, 2.75%, 8.81% respectively, **Figure 2.3**) compared to the antigen-specific CD8+ T cell population revealed by tetramer staining (i.e. 2.89%, 1.50%, 1.11% respectively, **Figure 2.1**). These non-antigen specific $CD8^+$ T cells may be activated after stimulation by 'bystander' or TCR-independent signaling, such as by cytokines or novel activating receptors (223).

Figure 2.2

Kinetics of $CD8^+$ T cell response to DENV2 infection

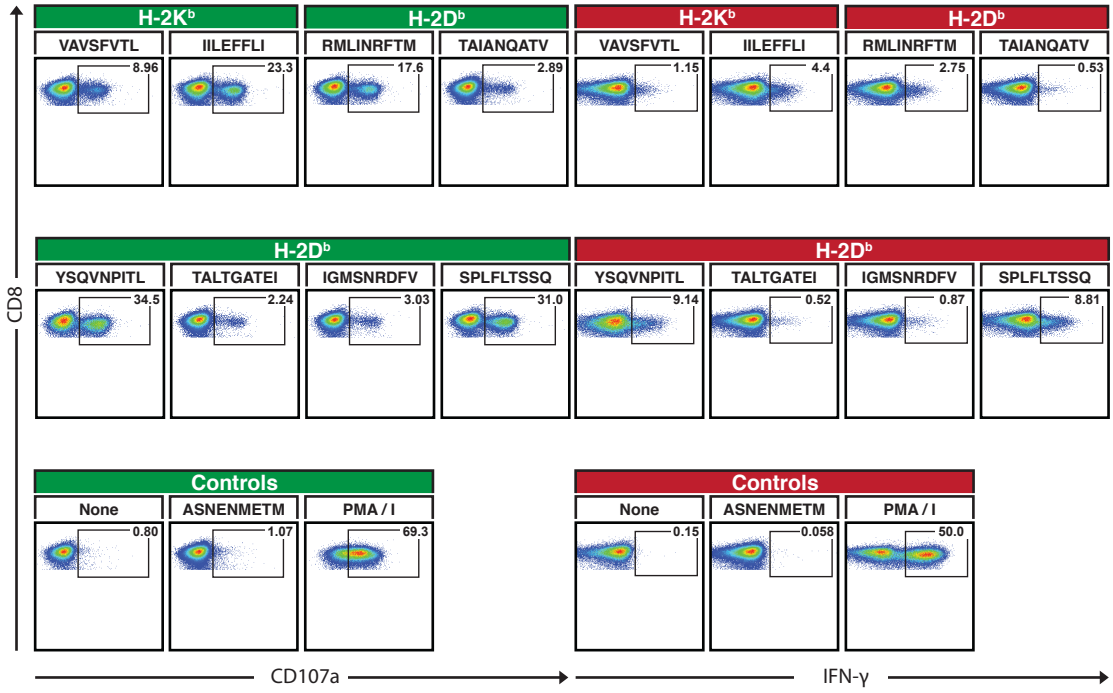


AG129 mice were infected with 10^4 PFU of unpurified D2Y98P. Splenocytes were stained with H-2Kb and H-2Db tetramers at the indicated days post infection. The percentages of CD8+ and MHC tetramer positive cells out of total viable splenocytes of each mouse were plotted (white circle). The mean percentages of CD8+/MHC Tetramer+ cells for each time point are represented by the bar graph.

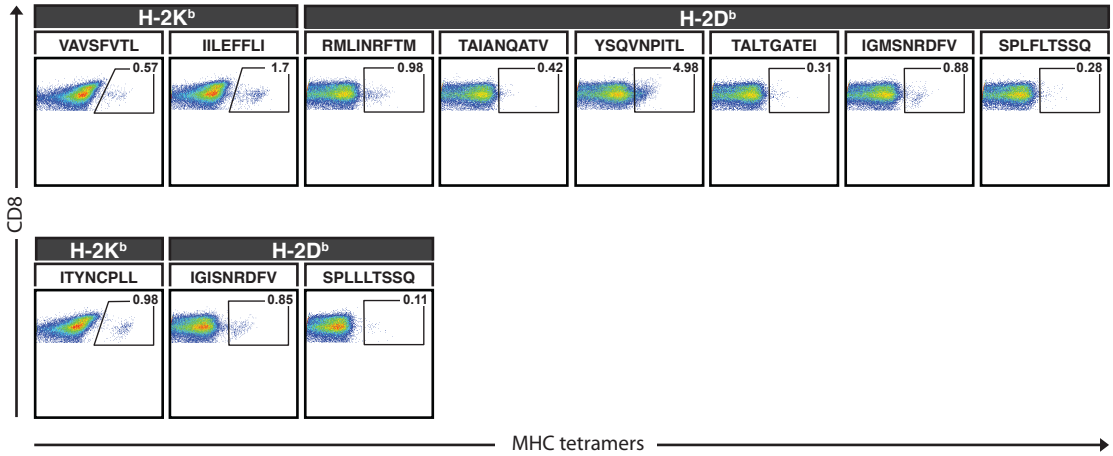
Figure 2.3

CD107a upregulation, IFN-γ production and MHC tetramer staining of T cells

A



B



(**A**) Splenocytes from AG129 mice infected i.p. with 104 PFU of D2Y98P (day 10 p.i.) were stimulated *in vitro* for 4 hours with H-2K^b or H-2D^b binding DENV2 antigens, or phorbol myristate aceate and ionomycin (PMA/I), or Influenza-A derived H-2D^b binding peptide ASNENMETM. Non-stimulated cells were included as controls (None). Cells were stained for CD107a during stimulation or intracellular interferon-gamma (IFN-γ) after stimulation. Data shown are representative of three replicate experiments. (**B**) Splenocytes from AG129 mice infected with 10^4 PFU of D2Y98P-PP1 (day 10 p.i.) were stained with H-2K^b and H-2D^b tetramers to identify antigen-specific CD8$^+$ T cell populations. Cells shown were pre-gated on CD8$^+$ and for viability, then gated for MHC tetramer (black box). The percentages of tetramer-positive cells out of the total CD8$^+$ T cell population are shown in the top right hand corner of each plot. Data shown represent stains from one of two mice.

The subsequent availability of a purified strain of D2Y98P (D2Y98P-PP1) that contained 250 mutations from the earlier strain prompted us to repeat the MHC tetramer screenings for the top predicted H-2K^b and H-2D^b binding peptides (**Appendix A.1**). Variants of the novel DENV-2 epitopes, such as NS2A$_{178-186}$ (SPLLLTSSQ) and E$_{4-12}$ (IGISNRDFV) were also included for screening. AG129 mice infected with purified D2Y98P-PP1 were used for this particular screening to exclude convoluting responses from the other viral strain. Comparable tetramer staining of the CD8$^+$ T cells were observed between the peptide variants (i.e. SPL**F**LTSSQ vs SPL**L**LTSSQ and IG**M**SNRDFV and IG**I**SNRDFV) (**Figure 2.3B**).

Another novel H-2K^b epitope at position 49 to 56 of the membrane glycoprotein precursor (prM), of the sequence ITYNCPLL, was identified while the dominant epitope remained to be NS4B$_{99-107}$ (YSQVNPITL). Altogether, the 9 identified epitopes span across the proteome of DENV-2 and majority of them originate from the non-structural proteins (NS) (**Figure 2.4**).

2.3.C. Characterization of dominant antigen-specific CD8$^+$ T cell populations

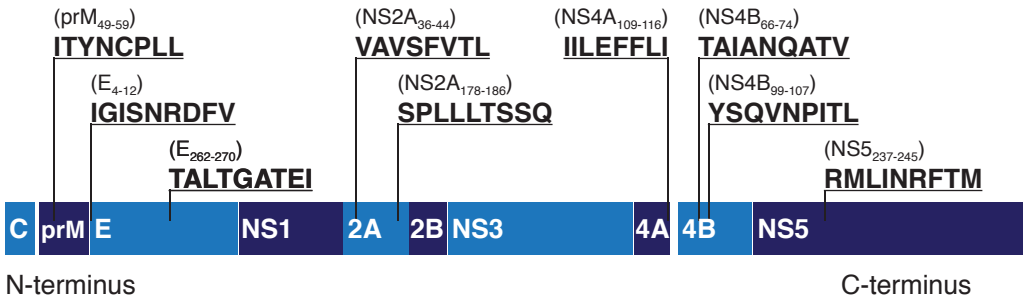
Despite the loss of viremia in i.p. infected AG129 mice at day 10 (218), the population of YSQVNPITL-specific (YSQ) CD8$^+$ T cells continues to expand differing from the trend of the other identified antigen-specific CD8$^+$ T cells (**Figure 2.2**). With the aid of H-2D^b YSQ-tetramers, we characterized the phenotype of this T cell population using various cell marker antibodies. While the total CD8$^+$ T cell population displayed mixed phenotypes in CD44 and CD62L expression, YSQ-specific CD8$^+$ T cells selectively displayed an upregulation of CD44, a marker for

activated T cells (224), as well as a downregulation of CD62L at 15 p.i. (**Figure 2.5**). This $CD44^{high}$ $CD62L^{low}$ phenotype characteristic of YSQ-specific T cells are common in activated T cells (160, 224, 225). The activated status of this T cell population even after the loss of viremia in the infected mice at day 10 was also suggested by the upregulation of chemokine receptor CXCR3 (**Figure 2.5**). CXCR3 is selectively expressed on activated $CD8^+$ T cells but not on resting T cells and is involved in the selective recruitment of $CD8^+$ T cells (226). Collectively, YSQ-specific T cells were $CD44^{high}$ $CXCR3^{high}$ $PD1^{high}$ $CD122^{high}$ $LAG3^{high}$ $CD62L^{low}$ $CD25^{low}$ $CD69^{low}$ (**Figure 2.5**).

As preferential usage of a particular TCR variable gene-segment combination by antigen-specific T cells following antigenic stimulation has been previously observed in various human or murine antigen-specific $CD8^+$ T cells (8, 227, 228), we determined the T cell receptor (TCR) variable-beta (Vβ) chain usage of YSQ-specific $CD8^+$ T cells. AG129 splenocytes 15 days post DENV-2 infection were co-stained with $H\text{-}2D^b$ YSQ-tetramers and antibodies for the various TCR Vβ. The TCR Vβ repertoire for total $CD8^+$ T cells was diverse with Vβ8 as the dominant TCR. However, the TCR Vβ repertoire for YSQ-specific $CD8^+$ T cells was significantly narrower and largely limited to Vβ7, 8 and 9 (**Figure 2.6**) following DENV infection. This confirmed our speculation for preferential usage of TCR Vβ in YSQ-specific $CD8^+$ T cells.

Figure 2.4

Organization of proteins from the DENV-2 genome and location of identified $CD8^+$ T cell epitope.

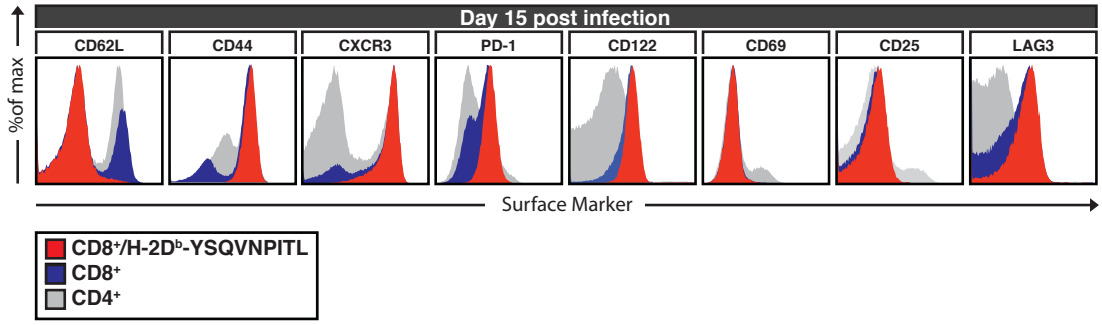


The newly identified DENV-2 epitopes were mapped in the proteome of D2Y98P-PP1. The position of each epitopes in the translated proteins, including core protein (C), membrane glycoprotein precursor (prM), envelope protein (E), and various non-structural protein (NS1, 2A, 2B, 3, 4A, 4B and 5) and a schematic representative of the epitope position within each protein are shown. For the dominant strain in unpurified D2Y98P, variant epitopes are IGMSNRDFV found in $E_{4\text{-}12}$, and SPLFLTSSQ in $NS2A_{178\text{-}186}$.

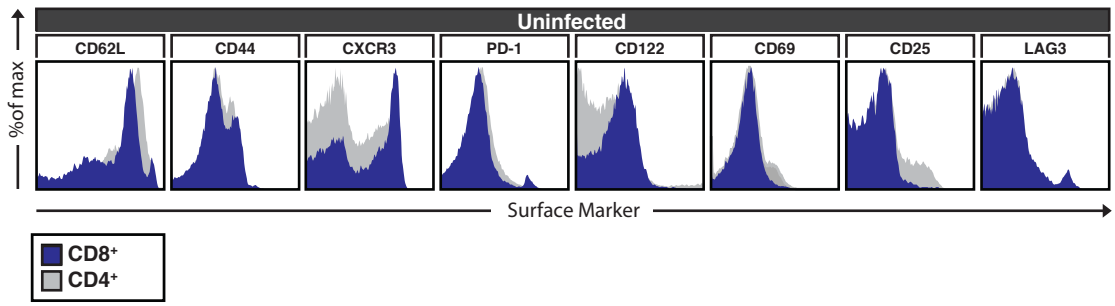
Figure 2.5

Molecular profile of YSQVNPITL-specific T cells

A.



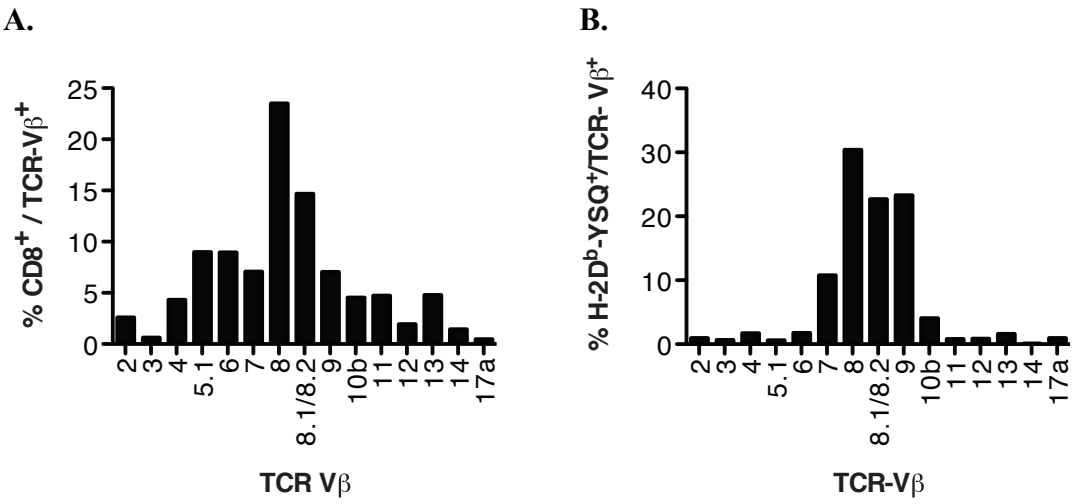
B.



Splenocytes from (**A**) AG129 mice infected with 10^4 PFU of unpurified D2Y98P at day 15, and (**B**) uninfected mice were analyzed. Viable $CD4^+$ (grey) or $CD8^+$ (blue) or $CD8^+$/H-$2D^b$-$YSQVNPITL^+$ T cells (red) were stained with various cell marker antibodies. Data shown represent stains from one of three mice.

Figure 2.6

TCR Vβ repertoire of T cells from DENV-2 infected AG129



The percentages of TCR Vβ positive cells for the particular TCR in (**A**) total $CD8^+$ T cell population and (**B**) YSQVNPITL-specific $CD8^+$ T cells at day 10 post infection with 10^4 PFU of unpurified D2Y98P were measured based on flow cytometry analysis after staining the splenocytes with H-$2D^b$-YSQVNPITL tetramer and various TCR Vβ antibodies. Data shown is representative of two replicates.

2.4 Discussion

A few major hypotheses for the development of clinical symptoms following DENV infection have been circulating as the search for the mechanisms involved in dengue pathology continues (212, 229). One prominent example is the antibody-dependent enhancement of disease (ADE) hypothesis, which speculates that a more severe clinical symptoms after a secondary virus infection is attributed to cross-reactive antibodies generated during a primary infection of a heterologous serotype that increases the uptake of virions into Fc-receptor positive cells (230, 231). A second hypothesis is the idea of 'original antigenic sin', in which cross-reactive memory T cells are causative of DHF/DSS. Supporting this, high frequencies of dengue antigen-specific T cells that were induced by the previous infection have been found in patients with secondary DENV infection (232). However, the protective or pathogenic roles of immune responses to DENV infection are still unclear (209). A thorough examination of the relevant cell population, such as dengue-specific CTL, would be essential to further the understanding of cellular immune mechanisms following DENV infection. We were therefore interested to adopt the MHC tetramer technology to improve our knowledge of such T cell populations. For this purpose, we chose to study a recently developed DENV-2 strain (D2Y98P) for its ability to induce a virulent phenotype in AG129 (IFN-$\alpha/\beta R^{-/-}$ and IFN-$\gamma R^{-/-}$) mice with relevant clinical symptoms (218). While AG129 mice have been shown to be susceptible to other virus strains (212, 214), they exhibited signs of paralysis, an irrelevant clinical manifestation to human dengue diseases (216, 217). As such, the D2Y98P strain was attractive to study, and we pursued to understand antigen-specific $CD8^+$ T cells responses after DENV-2 infection with D2Y98P in AG129.

Employing a predictive algorithm to rank peptides from DENV based on H-2K^b and H-2D^b binding, peptides were selected for the generation of MHC tetramers libraries (**Appendix A.1**). Splenocytes from infected mice were screened for tetramer stainings, and this led to the identification of 8 antigenic peptides in the dominant strain in the unpurified D2Y98P, 2 epitope variants for D2Y98P-PP1, a plaque purifed clone from D2Y98P, and an additional antigen specific only to D2Y98P-PP1 (**Figure 2.3B**). Our identified epitopes reside across various structural and non-structural proteins in the DENV-2 proteome, from membrane glycoprotein precursor, envelope proteins to nonstructural proteins NS2A, NS4A, NS4B and NS5 (**Figure 2.4**). This finding coincides with previous epitope screening efforts by IFN-γ ELISPOT in IFN-α/βR$^{-/-}$ mice infected with DENV-2 (clone S221) (176), while presenting seven previously unknown epitopes (i.e ITYNCPLL, IGISNRDFV and its variant IGMSNRDFV, TALTGATEI, SPLFLTSSQ and its variant SPLLLTSSQ, and IILEFFLI). Further investigation of the initial eight D2Y98P epitopes revealed that CD8$^+$ splenocytes from infected mice produced IFN-γ and upregulated CD107a upon stimulation by these peptides (**Figure 2.3A and B**). With the use of H-2K^b and H-2D^b tetramers, we directly tracked the presence of various antigen-specific CD8$^+$ T cells at different times post infection and mapped the kinetics and hierarchy of the initlial eight epitopes. YSQVNPITL-specific CD8$^+$ T cell population was identified to be the dominant response, and the kinetics of its expansion was unique where it is the only T cell response that increased steadily until time of death at day 20 p.i. instead of decreasing after day 10 p.i. as seen in the other responses (**Figure 2.2**). This observation is perplexing as there were no viruses detected in the liver, spleen, brain, and blood at day 10 p.i. of AG129 mice infected with 10^4 P.F.U of D2Y98P (218). In this study, YSQVNPITL-specific CD8$^+$ T cells were found to be CD44^{high} CD62L^{low} ,

a phenotype typical of activated T cells and exhibited preferential TCR Vβ 7, 8 and 9 usage (**Figure 2.5 and 2.6**). Future work to investigate other cellular markers such as those related to apoptotic genes will possibly aid us in understanding the death resistance of this particular dengue specific $CD8^+$ T cell population.

The antigen YSQVNPITL is part of the NS4B transmembrane protein that was postulated to play a role in anchoring the viral replication complex to the endoplasmic reticulum (ER) membrane (233). Various studies involving amino acid mutations had suggested the importance of NS4B in DENV infectivity and replication. In a recent study, AG129 mice infected with MT-5, a mutated clone from DENV-2 (strain D2Y98P-PP1) with a single amino acid F52L substitution in NS4B, displayed complete absence of disease state (222). In DENV-4 (clone rDen4), viruses having a P101L substitution in NS4B promoted viral replication in mammalian cells but showed decreased replication in C6/36 *Aedes aldopictus* cells (234). Interestingly, based on a published protein sequence of NS4B in DENV-4 clone rDen4 (NCBI Reference Sequence: NP_740324.1) (235), the amino acids before and after position 101 in NS4B (YSQVNPTTL at $NS4B_{96\text{-}104}$ in clone rDEN4) corresponded with our identified epitope (YSQVNPITL at $NS4B_{99\text{-}107}$ in clone D2Y98P-PP1). In a separate study on DENV-1 (clone 02-20, GenBank: BAD16759.1), a mutation at nucleotide position 7152 of the entire genome, which led to a T2353I amino acid substitution in the polyprotein, caused an increase in viral replication in Vero cells (236). The amino acids before position 2353 (YSQVNPLTL at position 2344 to 2351 of the polyprotein in clone 02-20) are also related to YSQVNPITL. Therefore, it might be worth for future studies to investigate single amino acid mutations in D2Y98P-PP1 within this particular epitope, such as a proline to leucine substitution at position 104 in NS4B,

and record the subsequent effect in virulence, clinical signs and symptoms as well as capability to induce T cell responses in AG129 mice after infection with the mutants.

The protective or pathogenic role of DENV-specific CD8+ T cells in dengue pathology remains controversial to this day. Previous studies have suggested a protective function of $CD8^+$ T cells in dengue-infected mice models, as illustrated by an increased viral load after depletion of $CD8^+$ T cells in IFN-$\alpha/\beta R^{-/-}$ before DENV-2 (strain S221) infection (176). Contrary to this, another study had suggested that $CD8^+$ T cells could contribute to dengue pathology in mice (237). HepG2 grafted severe combined immunodeficient (HepG2-grafted SCID) mice that were inoculated i.p. with DENV-specific $CD8^+$ cell clone (2D42 cells) before DENV-2 (strain Tr1751) i.p. infection showed more severe dengue-like clinical manifestations and mortality at an earlier time point than those that did not receive prior 2D42 cells inoculation (237). Resolution of these contradictory findings demands further studies on the specific function of each DENV-specific T cell population. With advances in MHC tetramer technologies, such as antigen-specific T cell isolation for adoptive transfer (238-240), or depletion of antigen-specific T cell populations *in vivo* by saporin-coupled tetramers (241, 242), it may be possible to tease out individual T cell subsets following DENV infection. If the study of the cellular mechanisms of DENV-specific CTL in humans is pursued instead, direct tracking and characterization of these cells could be done through the construction of MHC tetramers loaded with conditional ligands for the appropriate HLA allelic variants commonly found in the tissue donor cohort under investigation (232).

3

Construction of conditional ligands and HLA tetramer libraries for Asian HLA variants

3.1 Introduction: Conditional ligands for the HLA

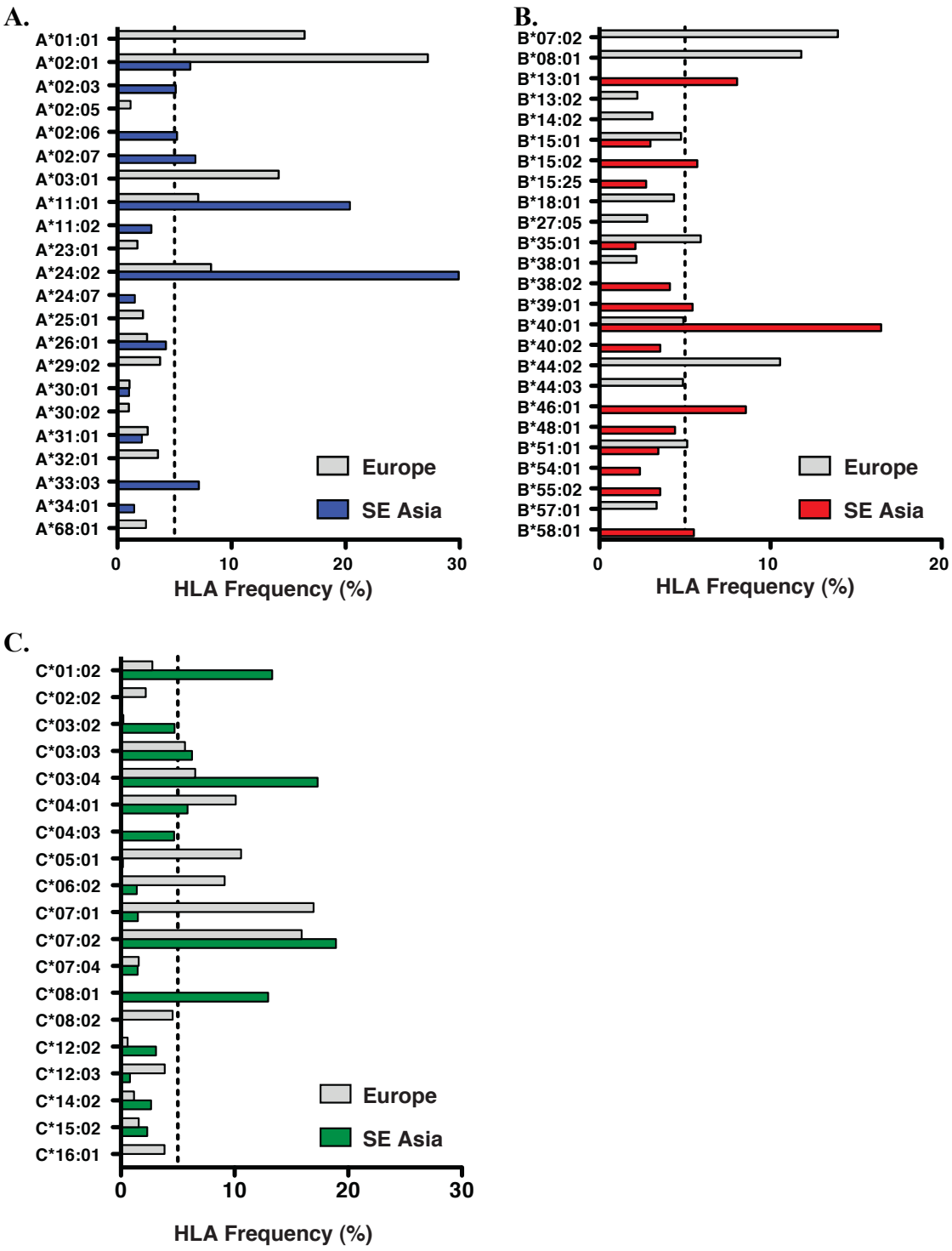
For the monitoring of T cell mediated immunity, successful identification, enumeration, phenotypic characterization and isolation of antigen-specific $CD8^+$ T cells response have become feasible with the advent of class I MHC tetramers (137, 243). The unique specificity of the tetramers for particular TCRs, such that they are able to distinguish between different antigen-specific T cells populations, is the hallmark of this T cell staining reagent (244). This remarkable specificity of MHC tetramers is dictated by two factors, which are the amino acid sequence of the peptide within the antigen-binding groove of the MHC, and the overall conformation of the complex as determined by its protein sequence (137). By keeping the MHC allelic variant constant, the use of different MHC-binding peptides to generate multiple complexes will result in the formation of MHC tetramer libraries that comprise of varying TCR specificities. Nevertheless, the laborious and time-consuming process in the construction of these complexes is a major hurdle to the creation of such libraries. Typically, the MHC heavy and light chain are recombinantly expressed and refolded in the presence of the MHC-binding peptide that is often obtained through solid phase peptide synthesis. The MHC monomers are then biotinylated enzymatically, purified and tetramerized to a streptavidin core decorated with a fluorophore to yield the final

product (142, 245). High-throughput production of peptide-MHC (pMHC) reagents with diverse specificities is now made possible with a recent technological breakthrough - the development of conditional ligands for HLA complexes (124, 141, 142) and murine MHCs (145-147). These ligands are essentially MHC-binding peptides containing a UV-light sensitive 3-amino-3-(2-nitro)phenyl-propionic acid group (Anp, J). Longwave UV-light exposure triggers the photolysis of the Anp residue and causes the ligand to fragment, while other amino acids of the complex remain unaffected (141). When MHC molecules loaded with conditional ligands are exposed to UV-irradiation, the cleaved peptides will dissociate from the MHC, leaving behind unstable empty complexes that can potentially unfold, aggregate and precipitate from solution. However, when the UV irradiation is performed in the presence of peptides that are able to refill the emptied MHC peptide-binding groove, a stable pMHC results. This peptide-exchange strategy therefore allows for the rapid generation of MHC libraries from a common precursor MHC loaded with a conditional ligand, and eliminates the need for pMHC protein refolding with every individual peptide (141).

The unequal distribution of HLA variants across different populations is a great hurdle to the design of conditional ligands. This is due to the variability in peptide-binding motif in different HLAs that prohibits interchangeable use of the conditional ligand for other HLA gene products (11). HLA variants commonly found in the European population are not necessary common in populations from other regions of the world, such as in South East Asia (SEA) (**Figure 3.1**) (24). For example, the archetypical HLA-A*02:01 has the highest frequency of occurrence in Europeans of

Figure 3.1

Frequency of HLA alleles in the European and South East Asian population.

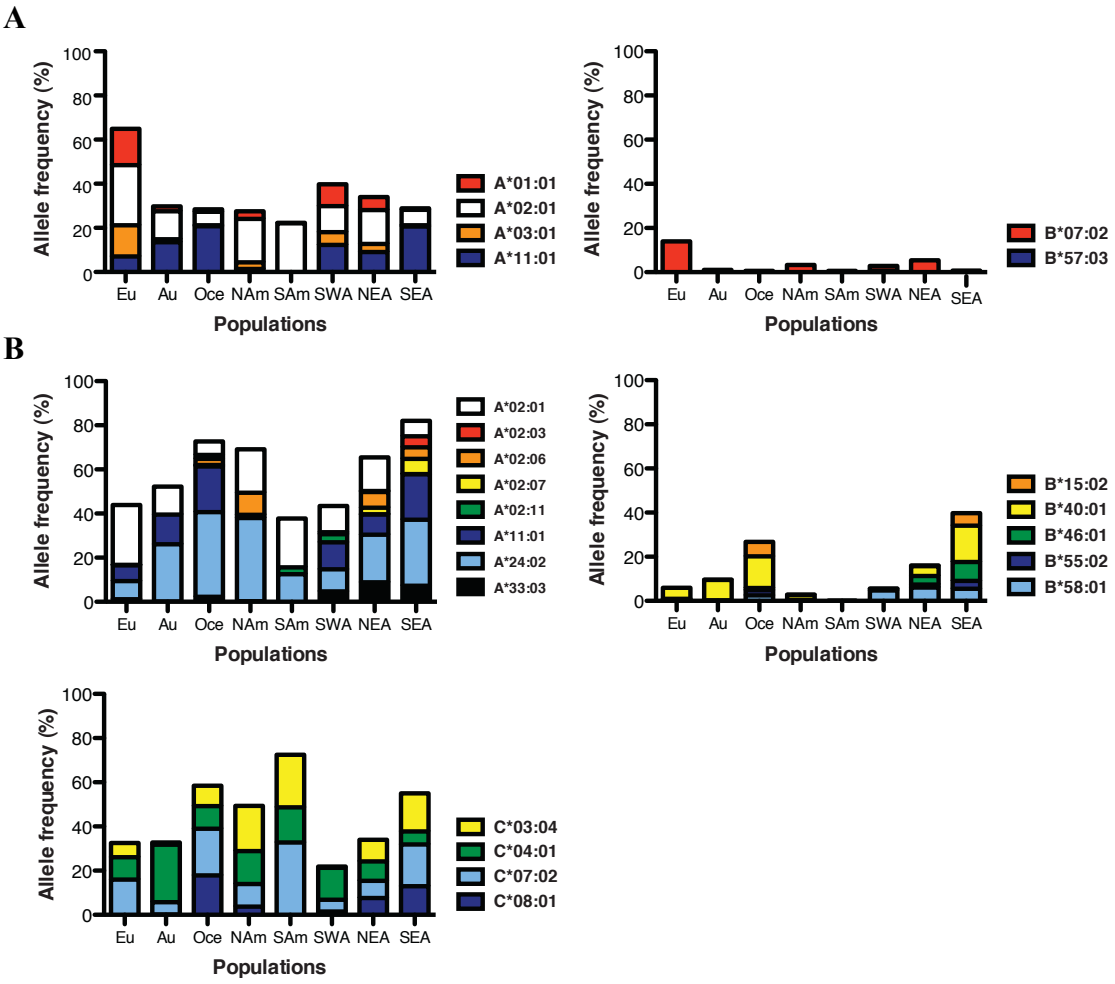


Frequency of HLA-A, HLA-B and HLA-C alleles in the European and South East Asian population. The alleles frequencies, based on data from dbMHC at NCBI, for (**A**) HLA-A , (**B**) HLA-B, and (**C**) HLA-C loci were tabulated for both European and South East Asian populations (24). Alleles with frequencies less than 1% (HLA-A and HLA-C) or 2% (HLA-B), were omitted. The dotted line indicates the 5% cut-off value.

~27%, but only 6% in South East Asians. Instead, HLA-A*24:02 is the most prevalent in the SEA population (i.e. ~29.9%), but not so in the Europeans (~8%). Thus, all of the conditional ligands previously designed for HLA-A*02:01 (141, 142), A*01:01, A*03:01, A*11:01, B*07:02 (124), and B*57:03 (246) to provide a significant coverage of the Western European population (**Figure 3.2**) have limited use in other regions, especially those residing the SEA. Neglecting the conditional ligand for populations originating from or inhabiting regions in Asia equates to limiting the use of this reagent in over 60% of the world's population (247). Hence, novel conditional ligands design for HLA variants common in Asian populations will greatly push the current boundary of the technology. We therefore report the peptide-exchange technology for seven HLA-A, five HLA-B, and four HLA-C restriction elements commonly associated with Asian ancestry to enable comprehensive assessment of the corresponding $CD8^+$ T cell responses. The HLA alleles to be covered include HLA-A*02:03, -A*02:06, -A*02:07, - A*02:11, -A*11:01, -A*24:02, -A*33:03, -B*15:02, -B*40:01, -B*46:01, -B*55:02, -B*58:01, - C*03:04, -C*04:01, -C*07:02, and -C*08:01, to target the technology towards South East Asian (SEA) populations, and provide 82%, 40%, and 55% allele frequency coverage for HLA-A, -B, and -C, respectively with HLA-A*02:01 included (**Figure 3.2**) (24). These will also cover the majority of HLA-A, -B and -C molecules having allele frequencies of >5% in Asia (**Figure 3.1**) and they are often unique to its inhabitants with negligible occurrence in Caucasians. Moreover, A*11:01, A*24:02, B*40:01, C*03:04, and C*08:01 dominate in SEA populations with allele frequencies at least 10% higher than in Europe (24). As we envisage the application of this technology in the accurate definition of T cell epitopes (248), the development of these tools for these HLAs can potentially address the prominent lack of novel epitope discovery and

Figure 3.2

Frequencies of HLA alleles for which conditional ligands have been designed.



The frequencies of HLA-A, -B and -C alleles for which conditional ligands have been designed or targeted for design are shown in percentages (%). Various HLA alleles frequencies for European (Eu), Australian (Au), Ocenian (Oce), North American (NAm), South American (SAm), South West Asian (SWA), North East Asian (NEA) and South East Asian (SEA) populations are indicated. (**A**) Conditional ligands were previously designed for HLA-A*02:01, A*01:01, -A*03:01, -A*11:01, -B*07:02, and -B*57:03. These HLAs are more prevalent in the European population (**B**) We further expanded the conditional ligand technology to include HLA-A*02:03, -A*02:06, -A*02:07, - A*02:11, -A*11:01, -A*24:02, -A*33:03, -B*15:02, -B*40:01, -B*46:01, -B*55:02, -B*58:01, - C*03:04, -C*04:01, -C*07:02, and -C*08:01. This provides more coverage for the Asian populations.

the characterization of the corresponding cellular immune responses related to Asian HLA variants. Hence, we generated tetramers for known EBV, HCMV or influenza antigens using the peptide-exchange strategy, and validated the function of these tetramers. Success in HLA tetramer staining of the appropriate T cells will provide proof-of-concept that functional tetramer libraries for epitope screening may be synthesized at will.

3.2 Materials and methods

3.2.A. Media and Buffers

The phosphate buffered saline (PBS) used commonly is prepared by a 10 times dilution of pre-made 10 X PBS (1st Base) and filtered using a 0.2μM Supor machV PES membrane (Nalgene). For T cell culturing, RPMI with 5% AB Serum is used and it contains RPMI 1640 with L-glutamin (Invitrogen) with 40μM of 2-Mercaptoethanol (Sigma-Aldrich), 100μg/mL of streptomycin and 100IU/mL of Penicillin (Invitrogen) and 5% heat inactivated AB Human Serum (Invitrogen).

3.2.B. Anti-human antibodies

Target	Host	Clone	Conjugation	Source
Human CD8	Mouse	RPA-T8	PB	BD Biosciences/ Biolegend
Human CD8	Mouse	HIT8a	FITC	BD Biosciences
Human CD8	Mouse	HIT8a	PE	BD Biosciences
Human CD8	Mouse	RPA-T8	APC	BD Biosciences
Human CD107a	Mouse	H4A3	FITC	BD Biosciences
Human IFN-γ	Mouse	25723	PE	R&D systems
Human Beta 2 Microglobulin	Mouse	D2E9	HRP	Abcam

3.2.C. Solid phase peptide synthesis

Conditional ligands were manually constructed on a 4-methylbenzhydrylamine (MBHA) resin (Advanced Chemtech) functionalized with a 4-(4-hydroxymethyl-3-methoxyphenoxy)-butyric acid (HMPB) linker (Merck KGaA) by standard Fmoc-based solid-phase peptide synthesis (SPPS). Deprotection of amino acid side-chains and release from the resin was achieved with 95% trifluoroacetic acid (TFA, Sigma-Aldrich) containing 2.5% water and 2.5% triisopropyl silane (TIS, Sigma-Aldrich), followed by precipitation in cold diethyl ether. Synthesized peptides were dissolved in dimethyl sulfoxide (DMSO, Sigma Aldrich) and their identities were confirmed by liquid chromatography/mass spectrometry (LC/MS) analysis on a LC/MS IT-TOF instrument (Shimadzu). Details of SPSS carried out are as follow:

(i) MBHA-HMPB-Resin preparation: MBHA resins (2 mmol) were added into a SPSS flask that has a frit column plate with an outlet connected to a vacuum pump via a collecting flask. The resins were washed 3 times using 4mL of N-Methyl-2-pyrrolidone (NMP, Sigma-Aldrich). For each wash, the flask containing the solvent and resins was shaken for 2 minutes and the solvent was subsequently removed by vacuum filtration. 0.5 mmol of HMPB, benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP, Novabiochem) and 0.5 mmol of Hydroxybenzotriazole (HOBt, Advance Chemtech) dissolved in 4 mL of NMP, together with 15 mmol of N,N-Diisopropylethylamine (DIPEA, Sigma-Aldrich) were added to the resins and the reaction mix was left shaking overnight. The solvent was removed and the resins were washed 3 times using 4mL of NMP.

(ii) Installation of first amino acid: Installation of the first amino acid starts from the C-terminus of the desired peptide. The N-terminus of the amino acid to be installed is protected by a Fluorenylmethyloxycarbonyl (Fmoc) group. The MBHA-HMPB-Resins (2 mmol) were washed 3 times using 4mL of Dichloromethane (DCM, Sigma-Alrdich). 5 mmol of Fmoc-Amino acid-OH , 25 mg of 4-Dimethylaminopyridine (DMAP, Sigma-Aldrich), and 5mmol N,N'-Diisopropylcarbodiimide (DIC, Sigma-Aldrich) dissolved in 4 mL of DCM were added to the resins, and left to shake for at least 2 hours. The solvent was removed and the resins were washed 3 times using 4 mL of DCM. Again, 5mmol of Fmoc-(1^{st}-Amino acid)-OH (Advance Chemtech), 25 mg of 4-DMAP, and 5 mmol DIC were added to the resins, and left to shake for at least 2 hours. The solvent was removed and the resins were washed 1 time using 4 mL of DCM, then 3 times using 4 mL of NMP, and another 1 time with 4 mL of DCM. The resins were dried by vacuum for 30 minutes and then stored in -20°C.

(iii) Fmoc Deprotection/ Amino acid coupling cycle: 0.2mmol of resins were washed 3 times using 4 mL of NMP. 20% piperidine (Merck) in NMP was used for the Fmoc deprotection reaction, and was added to the resin for 30 minutes. The solvent was removed and the resins were washed 3 times with NMP. 1 mmol of Fmoc-Amino acid-OH, 1 mmol of PYBOP and 1 mmol of HOBt dissolved in 4 mL of NMP, together with 3.0 mmol of DIPEA were added to the resins, and left to shake for at least 2 hours. The solvent was removed and the resins were washed 3 times using 4mL of NMP. These steps were repeated using the desired amino acid for coupling to obtain the peptide-of-interest.

(iv) Cleavage of peptides from resins: Any solvent remaining with the resin (0.1 mmol) was removed by vacuum filtration. 4 mL of a solution containing 95% TFA, 2.5% TIPS and 2.5% H_2O was added into the dried peptides pellet in the peptide flask and left to stand for at least one hour. The TFA solution containing peptides was eluted and collected into a 50 mL falcon tube. The above step was repeated again with another 4 mL solution of 95% TFA, 2.5% TIPS (Sigma-aldrich) and 2.5 % H_2O. 4 tubes of 45 mL Diethyl ether were cooled on dry ice for half hour. A magnetic stirrer was added into the tube of Diethyl ether and made to stir the contents in the tube. The TFA solution containing peptides were slowly dripped into 2 tubes of Diethyl ether, to a final ration of 1:10, and left to cool in the dry ice for at least 1 hour to allow for precipitate formation. The tubes were centrifuged at 2500 rpm for 10 minutes, and the supernatant was removed using a vacuum suction. The pellet was resuspended in the same volume of Diethyl ether as earlier and left to cool in the dry ice for at least 1 hour to allow for more precipitate formation. Again, the tubes were centrifuged at 2500 rpm for 10 minutes, and the supernatant was removed using a vacuum suction. The pellet containing peptides was left to dry overnight. The dried peptide pellets in DMSO were dissolved to a concentration of 10mg/mL.

(v) Liquid Chromatography-Mass spectrometry analysis of peptides: The identities of the peptides were confirmed by liquid chromatography/mass spectrometry (LC/MS) analysis on a LC/MS IT-TOF instrument (Shimadzu) with a C18 column for liquid chromatography. 5µL of 0.2 mg/mL peptide were used for each analysis.

3.2.D. Commercial peptides

Peptides used for screening and for the formation of HLA tetramer libraries were synthesized either by Mimotopes (Australia) and supplied in crude lyophilized form or by GenScript (Piscataway, NJ) supplied in lyophilized form at >70% purity. These peptides were dissolved in 80% DMSO in MilliQ water containing 10mM tris(2-carboxyethyl)phosphine (TCEP, Sigma Aldrich) and stored at -20°C.

3.2.E. Synthesis of HLA complexes and tetramers

The production of HLA complexes, tetramers and UV-mediated peptide exchanges were similar to that of murine MHC and caged tetramers *(See Chapter 2.2.E-Generation of MHC tetramers and peptide exchange)* and has been previously described in the literature (127, 142). The plasmids encoding human β2-microglobulin or any of the HLA-A*02:01, -A*02:03, -A*02:06, -A*02:07, -A*02:11, -A*33:03, -B*15:02, -B*40:01, -B*46:01, -B*55:02, -B*58:01, -C*03:04, -C*04:01, -C*07:02, and -C*08:01 heavy chains were synthesized by GenScript (Piscataway, NJ) and the genetic constructs were optimized for *Escherichia coli* production. Plasmids containing the heavy chain gene for HLA-A*11:01 (PTC48) and HLA-A*24:02 (PTCF6) were gifts from the NIH tetramer core facility. The HLA heavy chains were engineered with a C-terminal sequence for BirA-catalyzed site-specific biotinylation. Additional details on the plasmid sequences can be found in **Appendix B.2**. Biotinylated HLA monomers were purified by fast protein liquid chromatography (FPLC) on an S200 size exclusion column into a buffer containing 20mM Tris, 50mM NaCl (pH=8) and concentrated using Vivaspin™ to approximately 1mg/mL protein concentration.

3.2.F. HLA-stability ELISA

The assay was carried out as reported (127, 142). Briefly, wells of a 384 clear flat bottom polystyrene microplate (Corning) were coated with 50μL of 2 μg/mL streptavidin (Invitrogen) in PBS (2 h at 37°C) and washed with 0.05% PBS-Tween 20 (Sigma-Aldrich). 50μL of 2% BSA in PBS was added to each well and incubated 30 min at room temperature. The peptide exchange, initiated by 15 min UV-irradiation, was performed in eppendorfs containing 50 μM peptide, and 0.5 μM recombinant HLA in PBS at 4°C. The solution was then diluted to 1.6 nM HLA, 25μL added to each well, and incubated for 1 h at 37°C. Plates were washed four times with 0.05% Tween 20 in PBS, and 25μL of 1 μg/mL HRP conjugated anti-β2m antibody (Clone D2E9, Abcam) was added to each well, followed by 1 h incubation at 37°C. Plates were washed four times again with 0.05% Tween 20 in PBS, and 25μL of ABTS solution (Invitrogen) was added. After 15 min color development at room temperature, the reaction was stopped with 12.5 μL of 0.1% sodium azide (Sigma-Aldrich) in 0.1 M citric acid. The absorbance (λ= 415 nm) was measured using a Spectramax M2 plate reader (Molecular Devices).

3.2.G. *In vitro* stimulation of PBMCs and T cell lines

Approval for the study protocols was provided by the Institutional Review Board of the National University of Singapore. Healthy volunteers were recruited with informed consent and HLA-genotyped *(See 3.2J-HLA genotyping)*. Blood samples were obtained from volunteers with suitable HLA-type. PBMCs were isolated from fresh heparinized blood by density gradient centrifugation using Ficoll-paque plus (Amersham biosciences) in accordance to manufacturer's protocol. Cells, 1 million/mL, were cultured in RPMI with 5% AB serum *(See 3.2A- Media and buffer)* at 37°C, 5% CO_2. Antigenic peptides were added at 10 μg/mL, and 25 U/mL of interleukin-2 (IL-2, R&D systems Inc.) was added 2 days post-stimulation. Half media change and IL-2 supplementation was done every 2-3 days starting from day 5 post-stimulation.

3.2.H. HLA tetramer staining

PBMCs obtained from direct ex vivo isolation or from *in vitro* stimulation were washed with PBS. Cell viability was ascertained with LIVE/DEAD® fixable near-IR stain (Molecular Probes) prior to tetramer staining. Cells were washed and stained with 1 μg/50 μL of PE- or APC-conjugated HLA tetramers in PBS with 0.1% sodium azide on ice for 20 min. For the generation of A*33:03 tetramers, the p8* instead of p4* conditional ligand were used. Cells were subsequently stained with either anti-CD8-Pacific Blue™ (Clone RPA-T8, BD biosciences) or anti-CD8-FITC (Clone HIT8A, BD biosciences) for 15 min, washed and fixed with 1% paraformaldehyde in PBS. Data was acquired with a BD LSRII flow cytometer (BD biosciences) and analyzed using FlowJo (Treestar Inc).

3.2.I. *In vitro* stimulation and intracellular cytokine staining

The cells were grown at 1 million/mL in RPMI with 5% AB serum *(See 3.2A- Media and buffer)*. CD107a-FITC antibody (Clone H4A3; BD biosciences) and 10μg/mL of Brefeldin A (Sigma-Aldrich) were added, followed by stimulation with 10 μg/mL peptide or a mix of 100 ng/mL PMA and 10 ng/mL Ionomycin (Sigma-Aldrich) for 5 h at 37°C, 5% CO_2. The cells were washed, and HLA tetramer and surface marker staining was performed as above. Intracellular cytokine staining was done

using anti-human IFNγ-PE antibodies (Clone 25723, R&D Systems) and Human Intracellular Cytokine Staining Starter Kit (BD biosciences) according to manufacturer's specification.

3.2.J. HLA genotyping

The principles used in HLA genotyping are similar to published protocols (249, 250). Briefly, genomic DNA extraction was performed on 1mL of blood using PureLink™ Genomic DNA kit (Invitrogen) in accordance to manufacturer's protocol. Locus-specific amplification of HLA Class I genomic materials was achieved by Polymerase Chain Reaction (PCR). The reagents used per PCR reaction are 14.3 μL of MilliQ water, 3μL of Expand High Fidelity buffer (10X) with 15mM Magnesium chloride ($MgCl_2$) (Expand High Fidelity PCR System, Roche), 3μL of 2mM dNTPs (Promega), 1,5μL of DMSO, 3μL of 5μM primer mix (Sigma-aldrich), 1.5μL of DMSO (Sigma-aldrich), 0.2μL of 3.5U/μL of Expand High Fidelity enzyme mix containing Taq polymerase (Expand High Fidelity PCR System, Roche), 5μL of 10ng/μL extracted DNA to give a total volume of 30μL. For the primer mix, 5μL of each of the reverse and forward primers for HLA-A, -B or -C were mixed with 90μL of water to obtain 100μL of 5μM primer mix. The sequences of the 5'→3' primers used (from Tepnel) for HLA-A are *TggCCCCYggTACCCgT* (reverse, target Exon 3: 272) and *gAAACSgCCTCTgYggggAgAAgCAA* (forward, target Intron 1:21), for HLA-B are CCATCCCSggCgAYCTAT (reverse, target Intron 3:54) and gggTCCCAgTTCTAAAgTCCCCACg (forward, target -63 Exon1), and lastly for HLA-C are ggAgATggggAAggCTCCCCACT (reverse, target Intron 3:12) and AgCgAggKgCCCgCCCggCgA (forward, target Intron 1:42). The PCR products should be 910bp, 1106bp and 915 bp for the HLA-A, B and C primers respectively. The cycling conditions are based on the specifications by the manufacturer of Taq polymerase in Expand High Fidelity PCR System (Roche), and are in the sequence of 1 cycle of 94°C (2 mins), 10 cycles of 94°C (15s), 62°C (30s) and 72°C (1min), 20 cycles of 94°C (15s), 60°C (30s) and 72°C (1min), 1 cycle of 72°C at 7min, and hold at 4°C at the end of the PCR. The PCR products are purified using PureLink™ PCR Purification kit (Invitrogen) according to manufacturer's specification. The size of the purified products were checked by agarose gel electrophoresis and extracted for DNA sequencing. For HLA sequencing analysis, the electropherograms were matched with DNA sequences of different HLA-A, -B or -C allelic variants.

3.3 Results

3.3.A. Design and synthesis of conditional ligands

We aimed to develop conditional ligands for the principle South East Asian (SEA) variants with a panel consisting of HLA-A*02:03, -A*02:06, -A*02:07, -A*02:11, -A*11:01, -A*24:02, -A*33:03, -B*15:02, -B*40:01, -B*46:01, -B*55:02, -B*58:01, -C*03:04, -C*04:01, -C*07:02, and -C*08:01. For the design, we selected parent peptides based on documented affinity for the corresponding HLA product, such as the availability of a crystal structure, a validated *ex vivo* T lymphocyte response, or *in vitro* HLA binding data. Based on these specifications sixteen peptides, listed in **Table 3.1**, were selected for the pertinent HLAs. Strategic amino acid substitutions were primarily aimed at maximizing the likelihood of generating a ligand with unperturbed HLA affinity and several designs were tested (**Figure 3.3**). The introduction of Anp in central positions of the conditional ligand endeavored effective liberation from the peptide-binding cleft after UV-induced fragmentation. Whenever possible, the Anp-moiety should replace an aromatic residue (*e.g.* tryosine, tryptophan or phenylalanine) because of the chemical resemblance between them. Residues known to be problematic for the peptide synthesis and purification, such as arginine, tryptophan, glutamine or asparagine where side-chain deprotection can be non-quantitative or result in contaminating side products, and oxidation-prone cysteine and methionine were avoided to promote the quality of the conditional ligand. Following manual solid phase peptide synthesis (SPSS) of these ligands, liquid chromatography–mass spectrometry (LCMS) analysis were used to ascertain the peptide's identity (**Appendix B.1**). Each conditional ligand was subsequently refolded with their corresponding heavy and light chains (**Appendix B.2**), biotinylated, and purified by size exclusion chromatography to produce HLA monomers (127, 142).

Table 3.1

Origin of the parent peptides on which the conditional ligands were based.

Epitope Sequence	HLA	Organism	Protein	Location	ID / Ref
SVRDRLARL	A*02:03	Epstein-Barr virus	EBNA3	596 to 604	62305
LTAGFLIFL	A*02:06	Epstein-Barr virus	LMP2A	453 to 461	39790
LLDSDYERL	A*02:07	(-)	(-)	(-)	(251)
KMDIGVPLL	A*02:11	Dengue virus 1 Singapore/S275/1990	Polyprotein	2329 to 2337	32295
AIFQSSMTK	A*11:01	Human immunodeficiency virus 1	Pol polyprotein	263 to 271	1913
VYGFVRACL	A*24:02	Homo sapiens	Telomerase reverse transcriptase	461 to 469	99851
FYVDGAANR	A*33:03	Human immunodeficiency virus 1	SF2-Pol	594 to 602	(252)
ILGPPGSVY	B*15:02	Homo sapiens	Ubiquitin-conjugating enzyme E2 E1	83 to 91	27079
TEADVQQWL	B*40:01	Human parvovirus B19	Non-capsid protein NS-1	456 to 464	63285
KMKEIAEAY	B*46:01	Neovison vison	HSC70/HSP70	126 to 134	32336
KPWDVIPMV	B*55:02	Dengue Virus Type 1	polyprotein (Non-structural 5 protein)	2822-2830 (329 to 337)	32940
ISARGQELF	B*58:01	Kaposi's sarcoma-associated herpesvirus	ORF 57; immediate early protein homolog	212 to 220	28435
FVYGGSKTSL	C*03:04	Epstein-Barr virus	EBNA1	508-517	18328
QYDPVAALF	C*04:01	Human Cytomegalovirus	65 kDa lower matrix phosphoprotein	341 to 349	52886
VRIGHLYIL	C*07:02	Homo sapiens	MAGE-A12 protein	170 to 178	(253)
KAAVDLSHFL	C*08:01	Human immunodeficiency virus 1	Nef protein	80 to 89	101766

The conditional ligands were based on peptides found to be associated to the various HLA molecules. The epitope sequence, HLA restriction element, organism and protein source of the peptide and its location within the protein, and the literature references are given. For (-), further information is not available, and ID stands for the epitope identification number in the immune epitope database and analysis resource (254).

Figure 3.3

Design of UV-sensitive conditional ligands.

HLA	p1	p2	p3	p4	p5	p6	p7	p8	p9	Yield (mg/L)
A*02:01	G	I	L	G	F	V	F	●	L	4.35
A*02:03	S	V	R	D	●	L	A	R	L	12.00
A*02:06	L	T	A	●	F	L	I	F	L	4.90
A*02:07	L	L	D	S	●	Y	E	R	L	4.90
A*02:07	L	L	D	S	D	●	E	R	L	6.35
A*02:07	L	L	D	S	D	Y	●	R	L	4.70
A*02:11	K	M	D	I	●	V	P	L	L	5.50
A*11:01	A	I	●	Q	S	S	M	T	K	11.10
A*11:01	A	I	F	●	S	S	M	T	K	12.40
A*11:01	A	I	F	Q	S	S	●	T	K	14.60
A*24:02	V	Y	G	●	V	R	A	C	L	45.15
A*33:03	F	●	V	D	G	A	A	N	R	0.15
A*33:03	F	Y	V	●	G	A	A	N	R	12.90
A*33:03	F	Y	V	D	G	A	A	●	R	0.45

HLA	p1	p2	p3	p4	p5	p6	p7	p8	p9	p10	Yield (mg/L)
C*03:04	F	V	Y	G	●	S	K	T	S	L	2.20
C*04:01	Q	Y	D	●	A	V	Y	K	L		1.80
C*07:02	V	R	I	●	H	L	Y	I	L		11.20
C*08:01	K	A	A	●	D	L	S	H	F	L	14.7

HLA	p1	p2	p3	p4	p5	p6	p7	p8	p9	Yield (mg/L)
B*15:02	I	L	G	●	P	G	S	V	Y	1.55
B*15:02	I	L	G	P	●	G	S	V	Y	3.60
B*15:02	I	L	G	P	P	G	●	V	Y	5.70
B*40:01	T	E	A	D	V	●	Q	W	L	2.90
B*40:01	T	E	A	D	V	Q	●	W	L	3.50
B*40:01	T	E	A	D	V	Q	Q	●	L	3.05
B*46:01	K	M	K	●	I	A	E	A	Y	1.05
B*46:01	K	M	K	E	I	●	E	A	Y	0.90
B*46:01	K	M	K	E	I	A	●	A	Y	1.55
B*55:02	K	P	W	D	●	I	P	M	V	5.85
B*58:01	I	S	A	●	G	Q	E	L	F	4.25
B*58:01	I	S	A	R	G	●	E	L	F	3.00
B*58:01	I	S	A	R	G	Q	●	L	F	4.25

A single amino acid within the parent peptide was replaced with 3-amino-3-(2-nitro) phenyl-propanoic acid (Anp) at the position indicated by the filled black circle (l). Amino acid position 1 at the N-terminus to position 10 at the C-terminus are indicated by p1 to p10, respectively. Possible anchor residues based on known HLA-binding motifs are highlighted (green text), except for HLA-A*02:03, -A*02:11, and C*08:01 where the information is unavailable (90, 255). All of these ligands were synthesized by manual solid phase peptide synthesis and used for HLA refolding reactions. The yield represented by the weight of homogenous HLA monomer per litre of refolding reaction, was listed. When multiple conditional ligands were designed, such as in HLA-A*02:07, -A*11:01, -A*33:03 -B*15:02, -B*40:01, -B*46:01, -B*58:01, ligands with the highest yield were selected for further studies (shaded blue).

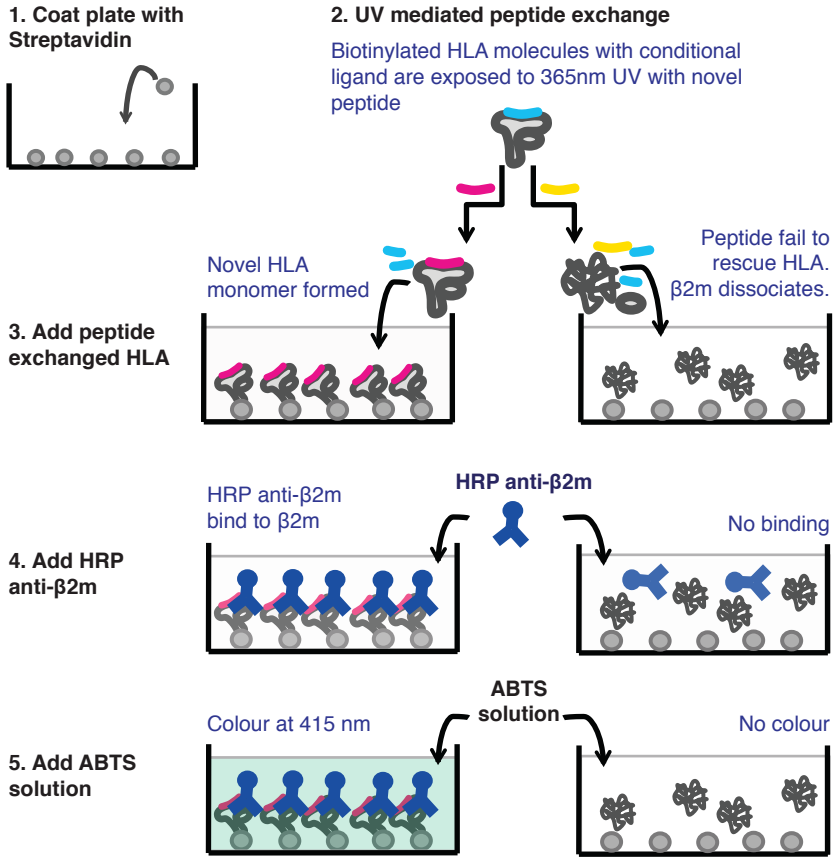
Although comparing yields is contentious, where the refolding reactions concern distinct molecular complexes, several noteworthy trends were discernable (**Figure 3.3**). For example, the HLA-A products with an average yield of ~10 mg/L outperformed HLA-B which furnished only ~3.2 mg/L. Variation could be as high as 300-fold between HLA products, with A*24:02 providing the highest (~45 mg/L) and A*33:03 (~0.15 mg/L) the lowest yield, presumably reflecting differences both in ligand affinity as well as the overall stability of the complex. Even for analogs designed from the same parent peptide, 100-fold differences could be observed (~13 mg/L to ~0.15 mg/mL for A*33:03). Based on these results, the conditional ligands with the most successful refolding were selected (**Figure 3.3**), and all of which had Anp substitutions between p4 and p8, with p7 being the most prevalent.

3.3.B. Stability of HLA complexes generated from UV-light mediated peptide exchange

Conditional ligands permit the manipulation of HLA complexes' specificities through the swapping of photocleavable ligands with peptides-of-interest after UV irradiation. Upon liberation of the cleavage by-products, the stability of the resultant complex varies according to the replacement peptide's HLA-binding competence, yielding either a novel HLAs or an unfolded protein (141, 142). We evaluated the structural integrity of our HLA complexes following peptide exchange by monitoring for the presence of non-covalently associated β2m subunits (**Figure 3.4**). Through the utilization of anti-β2m antibodies in an *in vitro* HLA-stability enzyme-linked immunosorbent assay (ELISA), we could verify the extent of HLA disintegration during various stages of the photocleavage reaction as an indicator for effective peptide exchange (127, 142). Moreover, this technique can be further exploited to

Figure 3.4

HLA-stability ELISA



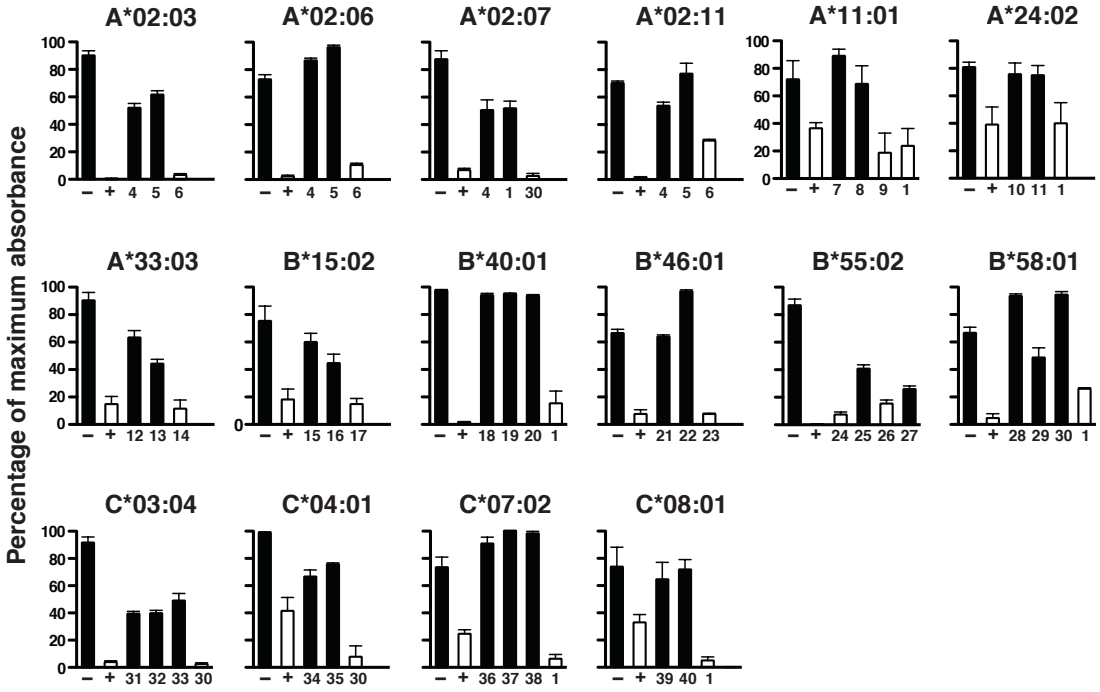
The HLA-stabiilty ELISA was used for the detection of stable HLA complexes following UV-mediated peptide exchange. The immobilized HLA complexes were probed with anti-human β2m antibody conjugated with horseradish peroxidase (HRP) for the presence of the non-covalently associated β2m subunit. Horseradish peroxidase catalyzed 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) (ABTS) colour development was tracked by its absorbance measured at λ=415nm.

define HLA restriction and peptide fine-specificity of novel CD8 T cell epitopes (*See Chapter 4*), without the consumption of valuable peripheral blood mononuclear cell (PBMC) samples (124, 125, 141, 142). We hence conducted HLA-stability ELISA for our constructed conditional ligand-loaded HLAs, in the presence and absence of longwave UV at 365nm, and HLAs loaded with known EBV, human cytomegalovirus (HCMV), HBV, DENV and influenza A virus antigens via the peptide-exchange strategy (**Appendix B.3**). These antigens had been previously identified through various epitopes identification techniques, such as ^{51}Chromium release killing assay, IFN-γ based ELISA, IFN-γ based ELISPOT, intracellular cytokine staining for IFN-γ and HLA tetramer or multimer staining (**Appendix B.4**).

For the sixteen pertinent HLAs, no appreciable degradation of conditional ligand-loaded complexes was observed in the absence of UV-irradiation (**Figure 3.5**). Exposure to UV light resulted in the expected near-quantitative loss of signal for the majority of HLA products examined as their structural integrity were compromised. However, A*11:01, A*24:02 and C*08:01 displayed exceptional stability with up to 39% of β2m still present under cleavage conditions which is reminiscent of the production yield of those complexes (**Figure 3.3 and 3.5**). After UV-fragmentation (141, 142), peptides capable of reoccupying the peptide-binding grooves would confer stability to the complexes by preventing their unfolding, aggregation and precipitation, while delivering novel specificities. For antigens derived from EBV, HCMV, HBV, DENV and influenza A virus that conform to the reported restriction elements (**Appendix B.3**), restoration of each of the 16 complexes with at least two separate ligands proved feasible in individual exchange reactions (**Figure 3.5**). However, peptides lacking the proper complementarity failed to rescue the HLA after

Figure 3.5

Stability of the recombinant HLA with conditional ligand subjected to UV-mediated peptide exchange.



An ELISA assay was used to compare the stability of biotinylated HLA complexes without any treatment (–), after 15 mins UV irradiation at 365nm in the absence of rescue peptide (+), or in the presence of indicated peptides (Peptide 1, 4 to 40, Supplemental Table 1). The peptides used are CLGGLLTMV [1], FLPSDFFPSV [4], FLLTKILTI [5], GLSRYVARL [6], AVFDRKSDAK [7], SIIPSGPLK [8], ATIGTAMYK [9], TYGPVFMSL [10], QYDPVAALF [11], SVNVHNPTGR [12], DTPVLPHETR [13], SVNVHNPTG [14], LEKARGSTY [15], GQGGSPTAM [16], LLDFVRFMGV [17], HERNGFTVL [18], KEVNSQLSL [19], CEDVPSGKL [20], HLAAQGMAY [21], CSAEWPTF [22], VQPPQLTQV [23], KPWDVIPMV [24], KPWDIIPMV [25], KPWDVVPMV [26], KPWDVLPMV [27], RGINDRNFW [28], IALYLQQNW [29], VSFIEFVGW [30], AYAAQGYKVL [31], PYFVRAQGLI [32], FVYGGSKTSL [33], KDVALRHVV [34], PYKRIEELL [35], VYPEYVIQY [36], KRQEILDLWVY [37], VRIGHLYIL [38], KAAVDLSHFL [39], VVCAHELVC [40]. Independent assays were conducted for each HLA, with various peptides efficient at stabilizing the HLA (black bars) and others not capable of maintaining its integrity (white bars), as indicated. The data was normalized, and the averages of four replicate measurements are shown with the standard error of the mean. At least two independent experiments were conducted for each HLA.

UV-irradiation, especially when mismatched in HLA supertype, which is characterized by significant differences in physicochemical properties of peptide binding pockets (11). For example, the removal of C-terminal primary anchor residue Arg in SVNVHNPTGR (peptide 12) to provide SVNVHNPTG (peptide 13) abolished the HLA-A*33:03-stabilizing properties of the peptide (256). Furthermore, the EBV-derived $LMP2_{426-434}$ (CLGGLLTMV) restricted by A*02:01 (*i.e.* HLA-A2 supertype) proved unable to conserve the integrity of HLAs from other subtypes such as HLA-B*40:01 and B*58:01. Also, HLA micropolymorphisms can profoundly affect peptide binding, where both HBV-derived $Env_{183-191}$ (FLLTKILTI) and $Pol_{455-463}$ (GLSRYVARL) associate well with A*02:01 (28), but only $Env_{183-191}$ proved capable of stabilizing the closely related variants A*02:03, A*02:06 and A*02:11 (**Figure 3.5**). Collectively, the results demonstrate successful temporary stabilization and peptide-exchange using conditional ligands tailored to Asian HLA variants.

3.3.C. Functional assessment of HLA tetramers

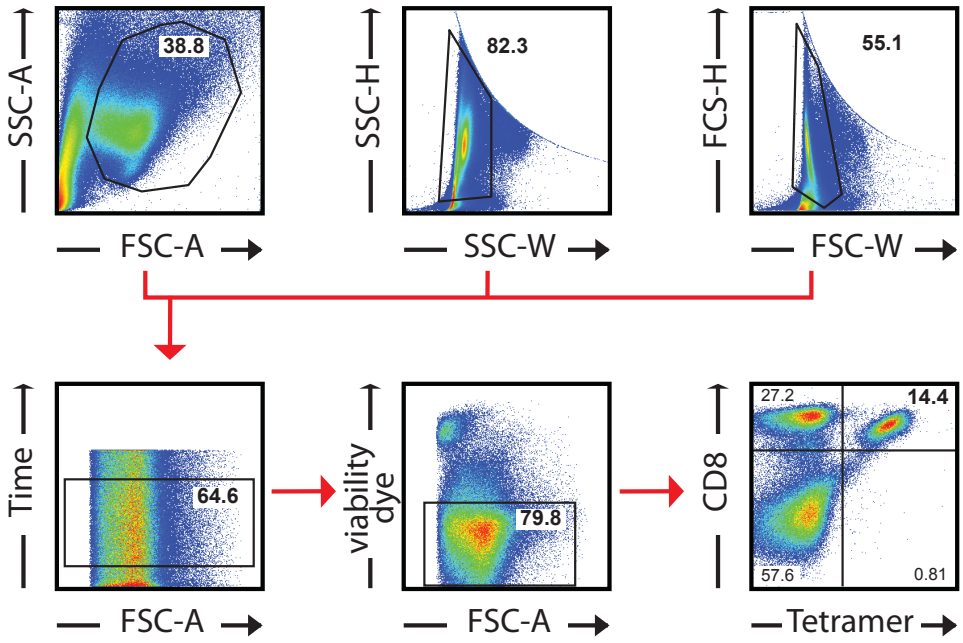
Our panel of HLA tetramers formed using complexes loaded with conditional ligands may be subjected to UV-light mediated peptide exchange (**Figure 1.4**). Exposure of these tetramers to longwave UV induces the dissociation of cleaved ligands from the complexes. Although emptied HLA products could be rescued from decay with other HLA-binding peptides, the question remained whether the resultant complexes present their antigenic peptide indistinguishably from naturally loaded HLAs.

We addressed this by staining antigen-specific T cell populations with the HLA tetramers. This will allow us to discern the capabilities of our tetramers that were catered towards the Asian population. We therefore recruited 45 healthy volunteers

and genotyped their HLA-A, -B, and -C alleles (**Appendix B.5**). Multiple fluorescently labeled HLA tetramers for known EBV, HCMV, HBV, DENV, and influenza A virus epitopes (**Appendix B.3**) were generated by UV-exchange. HLA-matched tetramers, together with stainings to differentiate viable and CD8 positive cells, were then used for direct *ex vivo* staining of isolated fresh PBMCs from healthy volunteers (**Figure 3.6**). Flow analysis of the *ex vivo* isolated T cells revealed in the majority of samples no clearly detectable responses (**Appendix B.6**) and, if any, were less than 1% of the CD8+ T lymphocytes. For example, a distinct HLA-A*1101 AVFDRKSDAK tetramer positive population could be visualized, but this represented only 0.046% of total CD8+ T cells. This data insufficiently validated the above-described HLA tetramers, and we therefore increased the antigen-specific $CD8^+$ T cell population by *in vitro* culturing them in AB serum-rich media containing IL-2 for 14 days following antigen stimulation. Visualization of these amplified CD8+ T cells by HLA tetramers was possible after 10 days in culture and the expansion of these cells continued even up to day 14 post stimulation (**Figure 3.7**). The propagated antigen-specific lymphocytes, could be successfully visualized by A*02:01, A*11:01, A*24:02, A*33:03, B*40:01, B*55:02, or B*58:01 tetramers, and represented a significant proportion –on average 8.65%– of cells that overall were fluorescently well-separated from the principle population (**Figure 3.8**). The percentage of tetramer stained cells ranged from 0.36% to an gratifying 40.3% (**Figure 3.8**), which was in contrast to the direct *ex vivo* tetramer stainings (**Appendix B.6**). Moreover, we subjected HLA-A*02:01 restricted FLPSDFFPSV-specific cells, engineered by TCR gene transfer (27), to staining with HLA-A*02:03, -A*02:06, -A*02:07 and -A*02:11 tetramers loaded with the HBV Core 18-27 peptide to

Figure 3.6

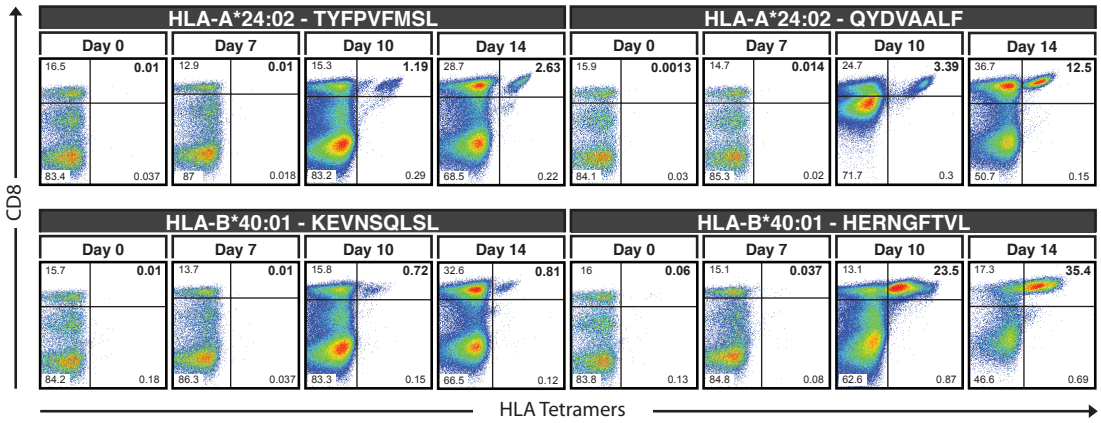
Gating strategy used to analyze HLA tetramer staining of $CD8^+$ T cells



Cells were first analyzed for their size and granularity based on their forward scatter area (FSC-A) and side scatter area (SSC-A) profile. Doublets were excluded using side-scatter height (SSC-H) vs side-scatter width (SSC-W) and forward-scatter height (FSC-H) vs forward-scatter width (FSC-W) parameters. Cells present in all three gates were analyzed for their time vs forward scatter area (FSC-A) to exclude cells that were read at the beginning and end of the flow cytometry run. Viable cells were gated based on their negative staining with the Live/Dead® dye (Invitrogen), and these cells were finally analyzed for anti-human CD8 antibodies and HLA-tetramer stainings. All data were analyzed using FlowJo (Tree Star).

Figure 3.7

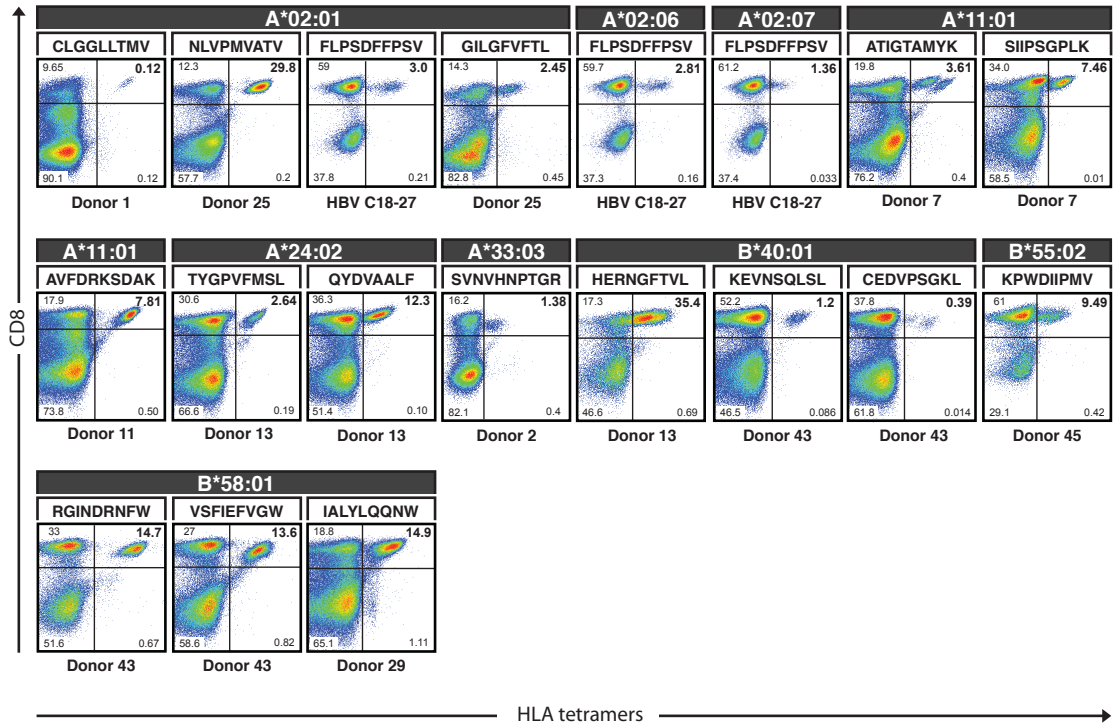
Culturing of PBMCs to develop antigen-specific T cell lines.



PBMCs isolated from healthy volunteer's (Donor No.13) blood were stimulated by antigens from EBV-LMP2$_{419–427}$ (TYGPVFMSL), HCMV-pp65$_{341-349}$ (QYDPVAALF), HCMV-IE-1$_{42-50}$ (KEVNSQLSL) or HCMV-pp65$_{267-275}$ (HERNGFTVL) and cultured in 25 U/mL IL-2, 5% AB serum in RPMI for 14 days. HLA tetramer staining for antigen-specific T cells was performed on day 0, 7, 10 and 14 days post stimulation. Cells were also stained for viability and co-stained with anti-human CD8a antibodies. The cells shown here were gated for viable cells, and for each antigen, the data was obtained independent HLA tetramer staining of the same batch of cultured cells. The apparent increase in CD8 staining of the botton-left quadrant in HLA-A*24:02-QYDVAALF stained cells at day 10 is likely due to inadequate washing of unbound anti-CD8 antibodies.

Figure 3.8

HLA tetramer staining of virus-specific $CD8^+$ T cells in donors' PBMCs.

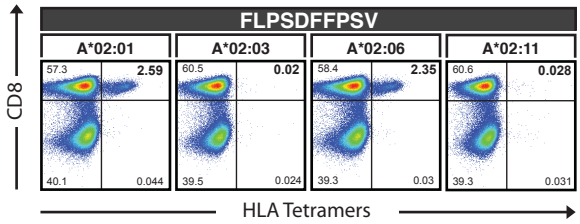


PBMCs were expanded for 14 days with their respective peptide, with the exception of FLPDFFPSV-specific T cells (HBV C18-27) that were engineered by TCR gene transfer of lymphocytes (257). Cells were stained for viability as well as with HLA tetramers presenting the same peptides as used for stimulation, and co-stained with anti-human CD8a antibodies. Numbers shown in the flow cytometry analysis are percentages of total live cells for each gated quadrant. Cells in the top-right quadrant represent HLA tetramer and CD8 double-positive cells. The complete HLA genotyping of PBMC donors is listed in Table 3.5. The cells shown here were gated for viable cells, and the data was obtained from independent experiments. The apparent increase in CD8 staining of the botton-left quadrant in HLA in donor 45 is likely due to inadequate washing of unbound anti-CD8 antibodies.

evaluate the capacity of the TCR to discriminate between micropolymorphisms in HLA-A2 subtypes common in the SEA population. Cross-specific recognition was observed for HLA-A*02:06 and -A*02:07 presenting the HBV peptides (**Figure 3.8**), yet no positive staining was obtained with HLA-A*02:03 or -A*02:11 tetramers (**Figure 3.9**). A closer look at a few antigen-specific T cells for TYGPVFMSL, VSFIEVGW, and IALYLQQNW cultured from the PBMCs of donor 19 revealed that HLA tetramer stained cells upregulated IFN-γ and CD107a with simultaneous slight downregulation of CD8 and TCR when stimulated with the corresponding antigens (**Appendix B.7**). Altogether, we validated most of our conditonal ligands in terms of their capability to generate functional tetramers. The dearth of known epitopes for Asian HLA variants precluded confirmation of tetramer functionality for HLA-A*02:03, -A*02:11, -B*15:02 and -B*46:01. HLA-C restricted responses were similarly not detectable in our donor cohort.

Figure 3.9

HLA Tetramer stainings of HBV C18-27 T cell clone



FLPDFFPSV-specific T cells (HBV C18-27) that were engineered by TCR gene transfer of lymphocytes (257) were stained with HLA tetramers of the HLA-A2 subtypes loaded with peptide FLPDFFPSV. Cells were stained for viability and co-stained with anti-human CD8a antibodies. The cells shown here were gated for viable cells.

3.4 Discussion

While there may be an abundance of CD8 T cells circulating within the host, only a small percentage of T cells are relevant to destroy virally infected cells (258). Recent developments in the MHC tetramer technology enable such antigen-specific T cell populations to be teased out from the masses of T cells through visualization in flow cytometry (243). Cell enrichment techniques exploiting MHC tetramers have also been made available (239, 259). The specificity and avidity of MHC tetramers to the T cells relies heavily on the identity of the presented peptide. In situations where the antigen in question is unknown, MHC tetramer libraries screenings can be employed to search for class I MHC-associated antigens (124, 127, 138, 145-147, 149, 150). The construction of such libraries was unfortunately a painstaking process, until the invention of UV-light sensitive conditional ligands. These photocleavable MHC-binding peptides allow the rapid synthesis of MHC molecules of distinct specificity through a peptide-exchange protocol, and allow the generation of MHC-based multimer arrays for high-throughput flow cytometry analysis of antigen-specific T cells. As conditional ligands had only been developed for a selected few HLA variants that are mostly prevalent in Western population, we sought to extend this technology specifically for Asian populations. Therefore, we designed and synthesized thirty UV-light sensitive conditional ligands for HLA-A*02:03, -A*02:06, -A*02:07, -A*02:11, -A*11:01, -A*24:02, -A*33:03, -B*15:02, -B*40:01, -B*46:01, -B*55:02, -B*58:01, -C*03:04, -C*04:01, -C*07:02, and -C*08:01, and constructed the corresponding HLA molecules. By doing so, we expanded the portfolio of conditional ligand-loaded HLAs nearly three-fold and for the first time included HLA-C allelic variants. Despite the polymorphic nature of HLA, this library includes the majority of alleles in Asia (24); the world's most populous region. The

stability of HLA complexes containing the conditional ligand were validated and all HLA variants proved cooperative with ligand replacement (**Figure 3.3**). We further investigated the integrity of the complexes following UV-mediated peptide exchange and found some previously known antigens to be capable of imparting stability to complexes that were made unstable through conditional ligand cleavage (**Figure 3.5**). The corresponding tetramers could also detect common virus-specific $CD8^+$ T lymphocyte responses in healthy volunteers (**Figure 3.8**).

We now possess conditional ligands with the capability to synthesize HLA libraries for the major Asian HLA allelic variants. Moreover, we have briefly demonstrated the adoption of these HLA libraries in conducting ELISA that determines the HLA complexes' stability. This may be extended to multiplexed beads, which have been successfully used as platforms for HLA-stability ELISA to enable flow cytometry analysis of each HLA's stability simultaneously (125). We also showed the employment of conditional ligands in the synthesis of HLA tetramers that are capable of distinguishing antigen-specific T cells and these reagents have the potential in high-throughput screens for novel T cell epitopes (124, 127, 138, 145-147, 149, 150). In principle, the peptide-exchange strategy may be applied to other recently established peptide-HLA tetramer reagents, such as toxic-coupled MHC tetramers that destroy antigen-specific T cells (241), 'reversible' MHC multimers that may be detached from T cells following staining (239), quantum-dots conjugated MHC multimers for combinatorial encoding of MHC tetramers (138, 140, 260), or precious metal-labelled MHC tetramers for cytometry time-of-flight (CyTOF) analysis (139). The use of conditional ligands, however, is not limited to the above-mentioned examples. Competitive binding fluorescence polarization (FP) assays of peptides and

labeled-tracer peptide for the 'emptied' MHC after UV-cleavage of conditional ligands provides yet another method to quantify a peptide's affinity for the MHC (124, 125). Large collections of crystalline peptide-MHC produced by UV-light induced ligand exchange within the MHC protein crystals provided high-throughput X-ray structural analysis (261).

As opposed to UV-mediated cleavage, there have been conditional ligands designed for chemical cleavage using mild oxidant $NaIO_4$ (143). Even though FP assays were successfully conducted using these ligands, there were a few caveats to this application. $NaIO_4$ causes partial oxidation of sulfur-containing amino acid groups in cysteine or methionine, and affects the quality of the resultant HLA complex. It is noteworthy that with our photolabile ligands, other amino acids or any cellular processes are unaffected during the photocleavage of the Anp residue (141). Not only is the photocleavage reaction highly specific, the photolysis rate of Anp is fast and 15 minutes of UV-irradiation is sufficient to facilitate peptide-exchange (124, 147, 262). Our conditional ligands are therefore, the most ideal for the synthesis of HLA libraries for Asian HLA variants, and there is every reason to believe that further applications using these reagents are plausible.

4

T cell epitope discovery in the Asian population

4.1 Introduction: Definition of CD8 T cell responses restricted by Asian HLA variants

Antigen-specific T cells recognize infected or neoplastic cells and respond by activation leading to target lysis, production of cytokines, and proliferation. This diverse set of effector functions generally contributes to the clearance of infection but can also lead to immunopathology. The unraveling of such T cell responses against pathogens, however difficult to study, are important in the development of immunotherapies that aim to manipulate specific cellular immune responses as a therapeutic strategy (105, 238, 263, 264). While the events leading to T cell activation may be intricate, what are the crucial elements to trigger this response? The T cell receptors (TCRs) of circulating antigen-specific $CD8^+$ T cells probe for class I HLA glycoprotein and interact with the composite surface of the surveyed molecule. When an appropriate intracellularly derived antigenic peptide ligand is being presented by the Human Leukocyte Antigen (HLA) protein heterodimer, such that the TCR has a favorable affinity towards the peptide-HLA complex, a T cell response results. These antigenic peptides, also known as epitopes, originate from the proteins of the invading virus, and are an important determinant for eliciting $CD8^+$ T cell functions. For a specific T cell clone, replacement of the epitope with other dissimilar HLA-binding peptides in the HLA will lead to failure in T cell activation. The HLA family of

restriction elements, which are important determinants of antigen presentation, are the products of the human class Ia gene and are encoded by three separate loci (*i.e.* HLA-A, -B, and -C) that are highly polymorphic. Amino acid variations of the HLA gene product often occur in the HLA peptide-binding pockets, affecting peptide affinity and consequently, the sets of epitopes it presents. As such, an individual's response and susceptibility to viral infections is influenced by their HLA constellation, and this has been evident in studies concerning diseases caused by dengue virus (DENV) (25), human immunodeficiency virus 1 (HIV-1) (26, 27, 201), hepatitis B and C virus (HBV and HCV) (28, 265, 266), Epstein-barr virus (202), human papilloma virus (HPV) (265). Given the significant correlation between HLA variability and geographical distribution and that HLA allelic variants mark distinct ethnic groups (267, 268), it is no surprise that the fine-specificity of T lymphocyte reactivity to viral afflictions is different for various ethnic groups (28, 201, 202, 266). In view of this, the majority of the epitope restricted by HLA-A*02:01 (254), have little relevance to the Asian population, as the HLA variant dominates in Caucasians but less so in Asians (24). The lack of confirmed antigenic peptides for Asian HLA variants has to be addressed so that future immunotherapies may be catered for these populations.

We have developed UV-sensitive conditional ligands for HLA variants common in Asian populations, such as HLA-A*02:03, -A*02:06, -A*02:07, - A*02:11, -A*11:01, -A*24:02, -A*33:03, -B*15:02, -B*40:01, -B*46:01, -B*55:02, -B*58:01, -C*03:04, -C*04:01, -C*07:02, -C*08:01, and then refolded corresponding HLA complexes containing these ligands (*See Chapter 3*). By the manipulation of the HLA's peptide cargo via exchange of the photoclevable ligand, caged HLA tetramers libraries may be generated and utilized for the precise definition of novel antigens as

demonstrated before (124, 127, 138, 145-147, 149, 150). Bearing in mind the lack of T cell epitopes description for asian HLA variants, we demonstrate the application of our conditional ligand in T cell epitope discovery for dengue Virus (DENV), severe acute respiratory syndrome corona virus (SARS-CoV), and hepatitis B virus (HBV) in the relevant South East Asian population. Our discovery pipeline comprises an initial enzyme-linked immunosorbent spot (ELISPOT) screen with pools of overlapping 15-mer peptides. The caged HLA complexes subsequently facilitate the rapid determination of peptide fine-specificity and HLA-restriction by separate assessment of peptide binding to the HLA, and peptide-HLA (pHLA) complexes to the TCR, respectively. This alleviates an epitope discovery bottleneck and conserves PBMC samples in the process. In addition, we investigated T cell cross reactivity to epitopes by generating HLA tetramers of dengue epitope variants. A posteriori assessment of our screening demonstrates that the majority of these novel epitopes would not have been found when only select restriction elements or antigenic proteins would have been considered, or bioinformatic selection of peptides would have been applied; restrictions that are commonly imposed on discovery efforts (138, 149, 150, 269).

4.2 Material and methods

4.2.A. Culture Media

For culturing of T cell line, AIM-V® with 2% AB human serum contains AIM-V® media (Invitrogen) with 2% heat inactivated AB Human Serum (Invitrogen). This media may be enriched with IL-2, IL-2, and IL-15 (R&D Systems) to give a final concentration of 20 units/mL of IL-2, 10ng/mL of IL-7 and 10ng/mL of IL-15 in the media. R10 media for culturing B lymphoblastoid cell lines (B-LCL) lines consists of 0.1% fetal bovine serum (FBS) (Gibco®), 20mM HEPES (Invitrogen), 100μg/mL of Penicillin and Streptomycin (Gibco®), 0.02% MEM Amino acid (Gibco®), 1mM sodium pyruvate (Gibco®), 0.01% non-essential amino acids (NEAA) (Gibco®), 0.01% L-glutamine (Gibco®) and 0.01μg/mL of Plasmocin (Invitrogen) in RPMI 1640 (Gibco®).

4.2.B. *In vitro* stimulation of novel T cell lines

The SARS-CoV-specific T cell lines derived from recovered SARS individuals (6 years post-infection), and HBV Core 18-27 TCR-re-directed T cells were obtained in previous studies (257, 270). The HBV Pol33 and HBV Env72 specific lines and DENV NS3-108 and DENV NS5-66 specific lines were developed as follows: antigen-specific $CD8^+$ T cells short-term lines were re-stimulated and grown from frozen PBMCs that were previously stimulated *in vitro* by 15-mer peptides. For re-stimulation, B lymphoblastoid cell lines (B-LCL) lines were pulsed with 5 μg/mL 15-mer in R10 media (*See 4.2A-Culture media*) for 1 h. 500 000 T cells were then cultured with 1.7 million gamma-irradiated feeder PBMCs (2500 rad) and 300 000 peptide-pulsed EBV-transformed B-cells (4000 rad) in 1.5mL of AIM-V media (Invitrogen) with 2% pooled human AB serum and supplemented with IL-2, IL-7 and IL-15 (R&D systems) (*See 4.2A-Culture media*) per well for 10 days at 37°C in 5% CO_2. Only lines containing $CD8^+$ T cells that upregulated CD107a and produced IFN-γ after 6 h peptide stimulation were used for epitope characterization.

4.2.C. Bioinformatics

The NetMHCpan-2.3 algorithm (http://www.immuneepitope.org/) was used to predict IC_{50} values of all 8-, 9-, 10- , and 11-mer peptides found in the proteome of HBV, DENV, and SARS coronavirus. The input data used were amino acid sequences for the all proteins present in HBV, DENV, and SARS coronavirus. This includes the core, X, envelope, and polymerase proteins for HBV Genotype B (NCBI accession AF121243), or HBV Genotype B containing the sequence "SVIWMMWYW" in the Envelope protein (NCBI accession EU796066) or HBV Genotype C (NCBI accession AB112063), as well as the capsid, membrane glycoprotein, envelope protein, NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5, proteins and the 2k peptide for DENV (NCBI accession NC_001474), and also the nucleocapsid protein and 3a protein for SARS-CoV (NCBI accession AY283798). The predictions for the relevant proteins were made for all six HLA-A, -B, and -C alleles per donor. The output IC_{50} of the peptides were ranked with an affinity threshold fixed at 1000nM.

4.2.D. Repeated protocols

Solid phase peptide synthesis, preparation of commercial peptides, synthesis of HLA complexes and tetramers, HLA-stability ELISA, HLA tetramer stainings, *in vitro* stimulation and intracellular cytokine stainings, and HLA genotyping were performed as previously described *(See Chapter 3.2-Materials and Methods)*.

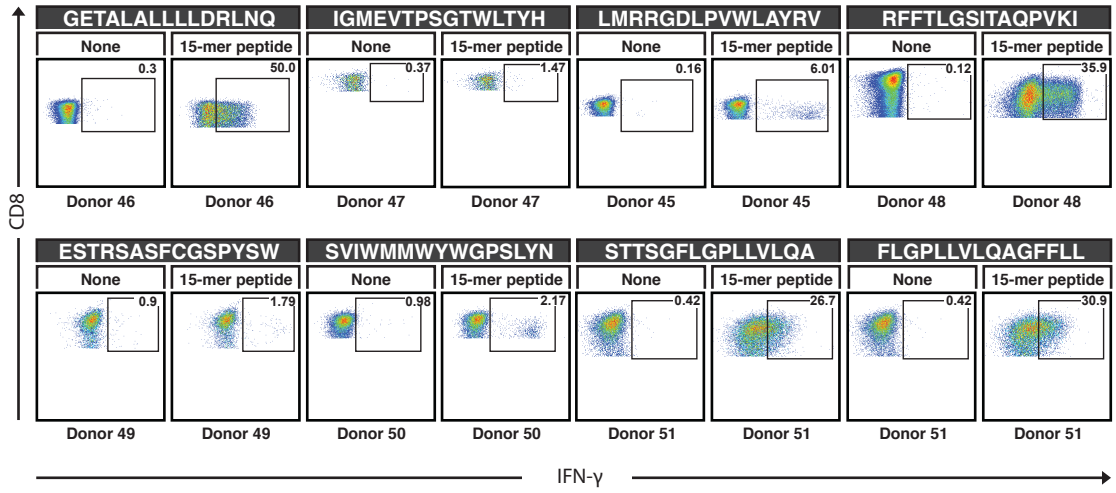
4.3 Results

4.3.A. Determination of peptide binding by HLA-stability ELISA

After the validation of our library of HLA products, we tested their practicality with antigen identification, as epitopes for Asian variants are scarce compared to their Caucasian counterparts and are often not fully characterized with regard to fine-specificity and HLA restriction (248, 267). We screened for $CD8^+$ T cell responses in individuals with a history of acute HBV, exposure to DENV, or SARS-CoV. PBMCs were *in vitro* stimulated with peptide pools of 15-mer peptides designed with 10 amino acid overlaps, thus covering the entire viral proteome of Dengue-2 (660 peptides), HBV genotype B (313 peptides), SARS-CoV nucleocapsid protein (83 peptides) and SARS-CoV 3a protein (53 peptides) (28, 270). IFN-γ secretion from T cells was assessed by ELISPOT. Positive pools were deconvoluted by stimulating the PBMCs with individual peptides followed by measuring IFN-γ production by intracellular cytokine staining (**Figure 4.1**). We identified with this conventional screening method novel $CD8^+$ T cell responses against 15-mers derived from HBV polymerase (HBV Pol33) and envelope (HBV Env72, HBV Env34), DENV non-structural protein 3 (DENV NS3-108) together with antigens embedded in the SARS-CoV nucleocapsid proteins (NP), and 3a proteins (270). Based on the HLA-A, -B, and -C genotyping of the short-term T cell lines (**Table 4.1**) and the known binding motifs of the various HLAs (11, 255, 256) (**Appendix C.1**), we anticipated the restriction elements of the HBV Pol33 and Env72 lines to be B*58:01, while that of DENV NS3-108 line was difficult to deduce (**Appendix C.2**). In addition, SARS NP65, SARS 3a2, and HBV Env34 peptides were known to be presented by B*40:01, B*58:01, C*03:04 respectively, based on inferring the HLA restriction from panels of

Figure 4.1

Functional status of $CD8^+$ T cell lines



IFN-γ production of seven T cell lines to the corresponding 15-mer epitope following 5 hours peptide stimulation (15-mer). As controls, the T cell lines were incubated for 5 hours in the absence of any peptide (none). The sequences of the 15-mer peptide used for stimulation to culture the T cell lines are shown. Cells shown were pre-gated on $CD8^+$, then gated for IFN-γ production (black box). The percentages of IFN-γ^+ cells out of the total $CD8^+$ T cell population are shown in the top right hand corner of each plot.

Table 4.1

HLA typing of donors previously infected with SARS-CoV, HBV or DENV

Donor	Epitope	HLA-A	HLA-A	HLA-B	HLA-B	HLA-C	HLA-C
45	KPWDIIPMV LMRRGDLPVWLAYRV	A*02:06	A*24:07	B*55:02	B*27:06	C*03:03	C*03:04
46	GETALALLLLDRLNQ	A*02:07	A*02:01	B*40:01	B*46:01	C*01:03	C*15:02
47	IGMEVTPSGTWLTYH	A*31:01	A*02:01	B*40:01	B*13:01	C*03:04	C*03:04
48	RFFTLGSITAQPVKI	A*33:03	A*11:01	B*58:01	B*55:02	C*03:02	C*03:03
49	ESTRSASFCGSPYSW	A*02:03	A*33:03	B*58:01	B*38:02	C*03:02	C*07:02
50	SVIWMMWYWGPSLYN	A*11:01	A*33:03	B*58:01	B*38:02	C*03:02	C*07:02
51	STTSGFLGPLLVLQA	A*11:01	A*02:01	B*15:02	B*53:04	C*08:01	C*04:01

The HLA-A, HLA-B and HLA-C genotyping of donors previously infected with SARS-CoV, HBV or DENV and showed T cell response to peptides from the respective viruses (epitope). The donor's PBMCs have been used to developed $CD8^+$ T cell lines used in this study. HLA alleles for which we report conditional ligands for have been indicated in red.

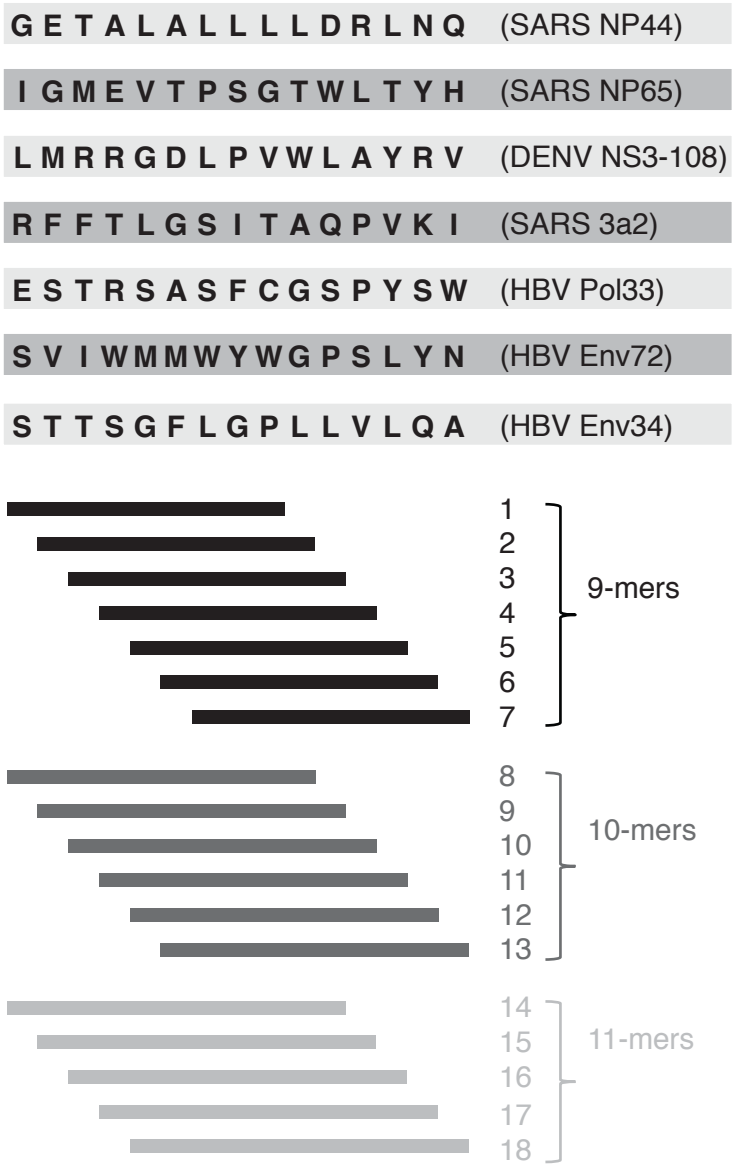
peptide-pulsed B lymphoblastoid cell lines (B-LCL) lines (*Dr. A. Bertoletti, personal communication*). To move away from this methodology of estimating the HLA restriction of peptides, which is consumptive on patient samples, and given the unreliability of using peptide-binding motifs for prediction, we obtained eighteen distinct overlapping 9-, 10-, and 11-mer peptides from each 15-mer epitope and tested all using a HLA-stability ELISA (**Figure 4.2, Appendix C.3**). The assay allowed us to rapidly screen the ability of each peptide in the respective set to rescue various HLA complexes from denaturation (127, 142). The SARS NP44 15-mer (270), for example, had three embedded peptides restricted by either A*02:01 or B*40:01 (**Appendix C.4**). For all SARS-CoV, DENV and HBV 15-mer antigens, we similarly found at least one peptide in each eighteen-member set that stabilized an HLA product covered by our library. These 9- and 10-mer peptides bound most frequently to the anticipated HLAs (**Appendix C.4**). Whilst no pattern in the position of these smaller fragments within the original 15-mer could be discerned, they often occurred at the N- or C-terminus of the parent peptide (**Figure 4.3**), underscoring the importance of the overlapping peptide screening approach.

4.3.B. HLA tetramer libraries for the identification of novel SARS-CoV, HBV and DENV epitopes

We subsequently confirmed the fine-specificity and HLA restriction for the curtailed list of peptides with tetramer staining (**Figure 4.3**). Employing a SARS NP44 T cell line (270), we found that both 9- and 10-mer peptides (i.e. GETALALLL and GETALALLLL) stabilize B*40:01 (**Figure 4.3**), yet the cell line stained positive only with B*40:01 tetramers loaded with the 10-mer GETALALLLL (**Appendix C.5**). Short-term expanded SARS-CoV NP65, DENV NS3-108, HBV Pol33, Env72 and

Figure 4.2

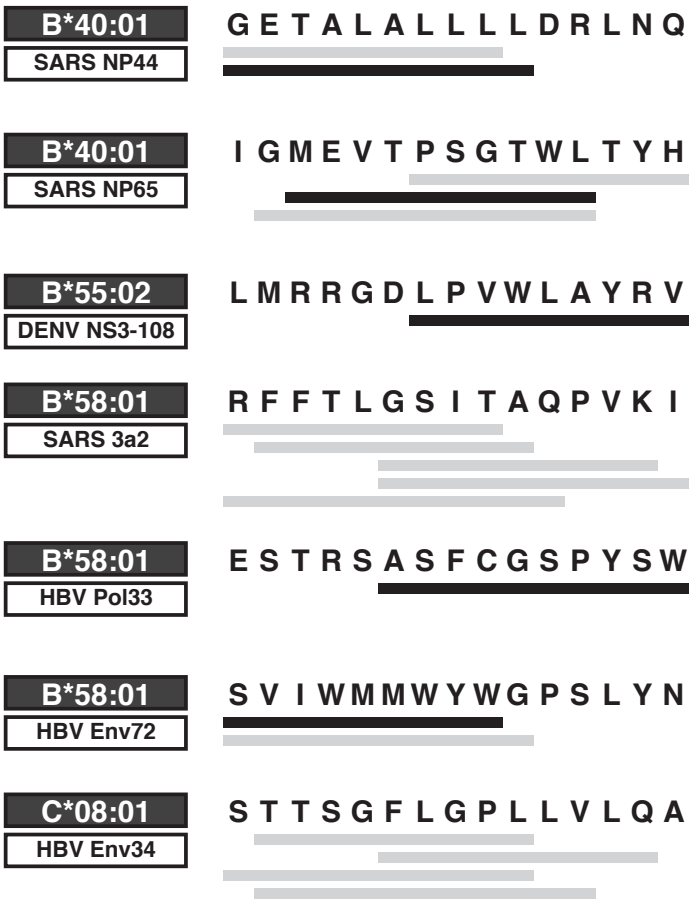
Derivation of 8-, 9-, 10-mer peptide libraries from 15-mer peptide



Overlapping 9-, 10-, 11-mer peptides were each tested for their capability to stabilize the corresponding HLA. A total of 18 different peptides were derived from each 15-mer peptide (sequence indicated). The truncations used to derive the peptide library were shown (thin black and grey bars). List of amino acid sequences of peptide fragments is provided in **Appendix C.3**.

Figure 4.3

Peptides binding to Asian HLA variants.



9-, 10-, or 11-mer truncated peptides imbedded in the 15-mer epitope that were tested positive for HLA binding in the HLA-stability ELISA are summarized from **Appendix C.4** (grey or black bars). Those that proved to be epitopes based on later HLA tetramer stainings (**Figure 4.4**) were indicated as black bars.

Env34-specific lines similarly stained with their corresponding B*40:01, B*55:02, B*58:01, or C*08:01 tetramers (**Appendix C.5A**). All of these tetramers did not to bind PBMCs from unrelated controls (**Appendix C.5B**). For the SARS 3a2 lines, the peptides pre-selected by ELISA failed to yield functional staining reagents. Comprehensive tetramer screens with all eighteen truncated peptides, however, revealed SITAQPVKI to be the optimal antigens when presented by B*58:01 (**Appendix C.5**). In HLA tetramer screenings with all eighteen truncated peptides for HBV Env34, we found that HLA-C*08:01 tetramers for FLGPLLVLQA gave a 4.3 fold higher percentage of HLA tetramers positively stained than the initial tetramers for FLGPLLVLQ (**Appendix C.5**). While the additional alanine residue in the C-terminus of the peptide abolished HLA-C*08:01 stabilizing effects at 37°C (**Figure 4.3**), the more superior HLA tetramer staining suggested TCR recognition was boosted (**Appendix C.5**). We had also found cross reactivity to HLA tetramers in SARS-CoV NP65 T cells (i.e MEVTPSGTWL and GMEVTPSGTWL). While the addition of a glycine residue in the N-terminus had no effect on HLA stability (**Figure 4.3**), it seemed to diminish TCR affinity as detected by HLA-tetramers (**Appendix C.5**). The extension or contraction of an epitope by a single amino was, in fact, frequently observed to be none detrimental to HLA binding (**Figure 4.3**), yet in most cases, only optimal epitopes were recognized by their corresponding TCRs. HLA tetramers were able to distinguish between similar peptides for SARS-CoV NP44 (i.e. GETALALLL and GETALALLLL), SARS-CoV 3a2 (i.e.SITAQPVKI, GSITAQPVKI, and GSITAQPVK), and HBV Env72 (SVIWMMWYW, and SVIWMMWYWG) (**Appendix C.5**). For the seven characterized epitopes, most T cells bound their HLA counterpart only with moderate affinity, as judged by the MFI shift for the individual tetramers, with only the B*58:01-presented HBV Pol33 9-mer

and HBV Env34 10-mer being displayed by C*08:01, such that antigen-specific cells were clearly resolved from the main $CD8^+$ T cell population. The functionality of all cell lines, established by IFN-γ production and CD107 surface expression following peptide stimulation, confirmed all our HLA-tetramer results (**Figure 4.4**). With increasing concentration of peptide used for stimulation, an increasing response in IFN-γ production and CD107 expression of the CD8 T cell population was observed (**Figure 4.4B**). The exception was for stimulation with ASFCGSPYSW, where the functional response of IFN-γ production or CD107 expression of the CD8 T cell declines beyond stimulation with 1ng/ml or 10ng/ml of the antigen respectively. Such a decline is a characteristic of overstimulation of T cell by high peptide concentration, which results in TCR downregulation and consequent decrease in IFN-γ production (271).

Collectively, we identified seven novel B*40:01, B*55:02, B*58:01 and C*08:01-restricted epitopes for SARS-CoV, HBV and DENV (**Table 4.2**), thus demonstrating the applicability of Asian HLA conditional ligands in T cell epitope mapping.

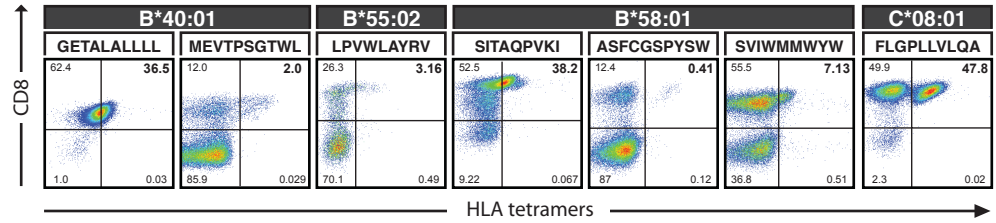
4.3.C. HLA tetramer libraries facilitate T cell cross-reactivity analysis

Mutations in the pathogen causing major or minor alterations to the expressed protein can modify the host's immune recognition of invading pathogens when the changes affect peptide epitopes. Modifications in the antigen may lead to dramatic differences in HLA presentation, T cell recognition, response, and even the undesirable immune surveillance escape of the mutated pathogen (201, 272-274). While T cells' cross recognition to variant epitopes have been extensively studied for a variety of viral

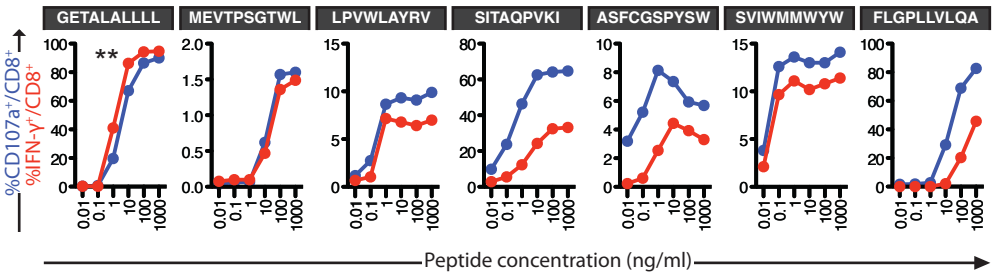
Figure 4.4

Identification of novel SARS-CoV, HBV and DENV epitopes with photocleavable tetramers.

A



B



(**A**) Summary of positive tetramer staining of SARS-CoV, HBV, or DENV-specific T cell lines to identify novel epitopes as shown in **Appendix C.5** (**B**) Dose response curves of the corresponding T cell lines to the optimal epitope following 5 hours peptide stimulation. The percentage of $CD107a^+$ (blue) or IFN-γ^+ (red) cells of the total $CD8^+$ T cell population is plotted against peptide concentration. The T cell function data for GETALALLLL-specific T cells has been previously reported (**) (270).

Table 4.2

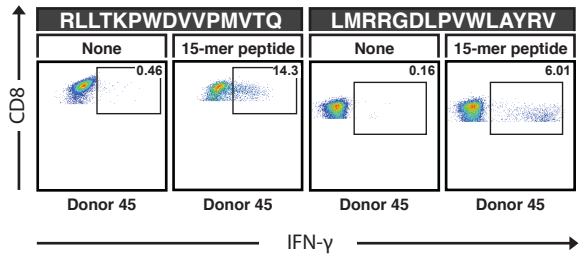
Sequences of newly identified epitopes.

Epitope Sequence	HLA	Source			Possible Strain (Accession no.)
		Organism	Protein	Location	
GETALALLLL	B*40:01	SARS Coronavirus	Nucleocapsid protein	216 to 225	Sin2774 (AY283798)
MEVTPSGTWL	B*40:01	SARS Coronavirus	Nucleocapsid protein	323 to 332	Sin2774 (AY283798)
SITAQPVKI	B*58:01	SARS Coronavirus	3a protein	12 to 20	Sin2774 (AY283798)
SVIWMMWYW	B*58:01	Hepatitis B virus	Envelope protein	356 to 364	Japan/Ry30/2002, Genotype B (Q8JXB9)
ASFCGSPYSW	B*58:01	Hepatitis B virus	Polymerase protein	161 to 170	Genotype B (AF121243)
LPVWLAYRV	B*55:02	Dengue Virus	Polyprotein (Non-structural 3 protein)	2017 to 2025 (542 to 550)	Dengue 2 (NC_001474)
FLGPLLVLQA	C*08:01	Hepatitis B virus	Envelope protein	171 to 180	Genotype C (AB112063)

Novel HLA-B*40:01, -B*55:02, -B*58:01 or C*08:01 restricted $CD8^+$ T cells epitopes and their HLA restriction element, organism and protein source of the peptide, location within the protein, and possible virus strain or genotype.

Figure 4.5

Dengue NS5-66 and NS3-108 $CD8^+$ T cell lines

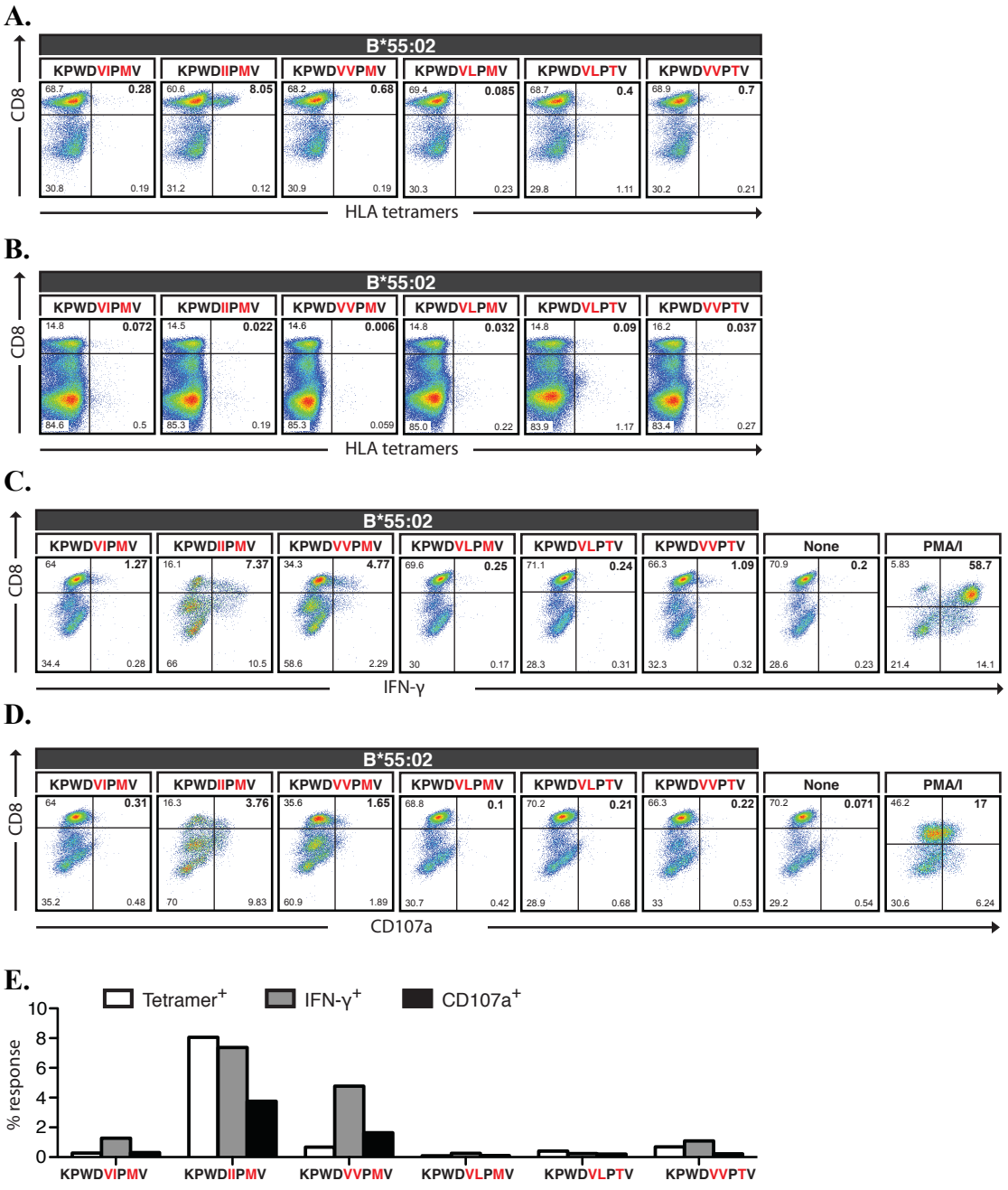


IFN-γ production of Dengue T cell lines to the corresponding 15-mer epitope following 5 hours peptide stimulation (15-mer). Both T cell lines were derived from donor No.45. As controls, the T cells were incubated for 5 hours in the absence of any peptide (none). The sequences of the 15-mer peptide used for stimulation to culture the T cell lines are shown. Cells shown were pre-gated on $CD8^+$, then gated for IFN-γ production (black box). The percentages of IFN-γ^+ cells out of the total $CD8^+$ T cell population are shown in the top right hand corner of each plot.

infections, the function of their cross-reactivity whether protective or pathological remains elusive (232, 275, 276). We aimed to demonstrate that the array of Asian HLAs with conditional ligands are ideal for probing TCR recognition of peptide variants independently from functional T cell responses. Dengue viruses are classified in four serotypes with significant amino acid variation between them, and provide an excellent collection of genotypically and antigenically distinct sequences. We employed two $CD8^+$ CTL clones with independent DENV-specificity that originated from a PBMC donor with a DENV infection history (Donor no.45) to examine potential cross-reactive epitope recognition (**Figure 4.5**). Our first B*55:02-restricted line showed reactivity towards 15-mer peptide RLLTKPWDVVPMVTQ (DENV NS5-66) and heterologous T cell responses to KPWD**VV**PMV (NS5 329-337) had been previously observed for several DENV clones (85). A panel of six B*55:02 tetramers was produced by peptide exchange, and differences in staining for the peptide variants were immediately evident. A clearly resolved antigen-specific population (8.05% of live $CD8^+$ cells) was seen with B*55:02 tetramers presenting KPWD**II**PMV, whereas all other variants stained significantly less with between 0.7% to background levels (**Figure 4.6A to 4.6B**). Functional responses of the $CD8^+$ line (i.e. IFN-γ production and surface upregulation of the degranulation marker CD107a) yielded comparable results upon stimulation with the various peptides (**Figure 4.6C to 4.6E**). Interestingly, the minimal 9-mer from the original 15-mer sequence used to develop the T cell line (KPWD**VV**PMV), showed only the second highest functional response and significantly lower tetramer staining compared to the optimal KPWD**II**PMV epitope. Therefore, functional T cell responses could be observed, yet the panel of HLA tetramers most clearly discriminated CTL specificity.

Figure 4.6

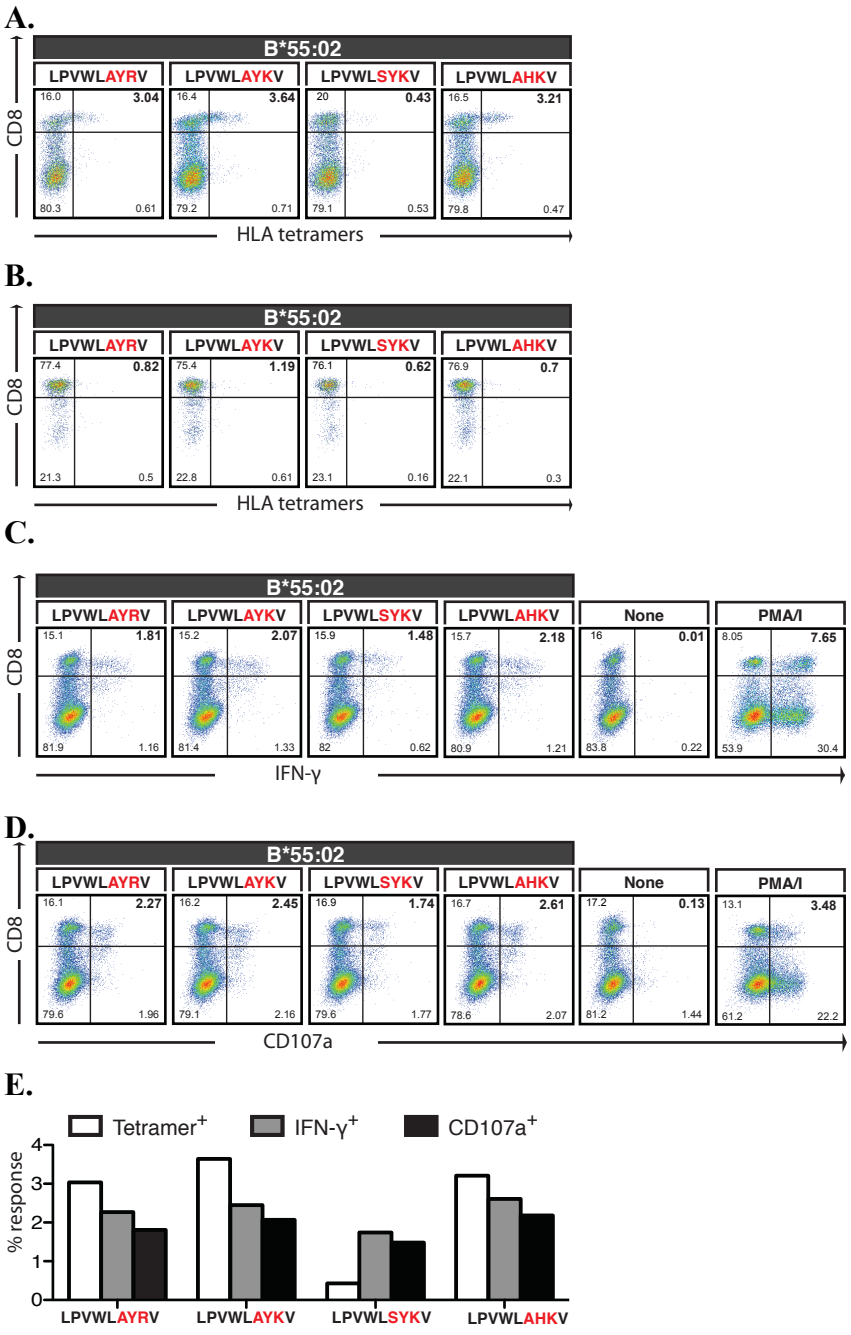
HLA tetramer stainings and function of NS5-66 T cell line.



(**A to B**) HLA tetramer staining of (**A**) DENV NS5-66-specific T cell line or (**B**) PBMCs from unrelated volunteers. The cells were stained with HLA-B*55:02 tetramers loaded with the indicated variants of $NS5_{329-337}$ peptide. Numbers shown in the flow cytometry analysis are percentages of cells for each gated quadrant. (**C**) IFN-γ production and (**D**) CD107a, also known as lysosomal-associated membrane protein-1 (LAMP-1) upregulation of DENV NS5-66-specific T cell line after 5 hours stimulation with the indicated peptide variant of $NS5_{329-337}$. As controls, the T cells were incubated for 5 hours in the absence of any peptide (none) or in the presence of phorbol myristate aceate and ionomycin(PMA/I). (**E**) % tetramer positive cells out of total live cells for each peptide variant, and their IFN-γ or CD107a response after 5 hours stimulation with the same peptide.

Figure 4.7

HLA tetramer stainings and function of NS3-108 T cell line.



(**A to B**) HLA tetramer staining of (**A**) DENV NS3-108-specific T cell line or (**B**) PBMCs from unrelated volunteers. The cells were stained with HLA-B*55:02 tetramers loaded with the indicated variants of $NS3_{542\text{-}550}$ peptide. Numbers shown in the flow cytometry analysis are percentages of cells for each gated quadrant. (**C**) IFN-γ production and (**D**) CD107a upregulation of DENV NS3-108-specific T cell line after 5 hours stimulation with the indicated peptide variant of $NS3_{542\text{-}550}$. As controls, the T cells were incubated for 5 hours in the absence of any peptide (none) or in the presence of phorbol myristate aceate and ionomycin(PMA/I). (**E**) % tetramer positive cells out of total live cells for each peptide variant, and their IFN-γ or CD107a response after 5 hours stimulation with the same peptide.

The second DENV-specific T cell line (NS3-108), developed from the same donor and reactive against newly identified epitope LPVWL**AYR**V restricted by B*55:02, was similarly probed with four peptide variants (**Appendix C.6**). HLA tetramer positive populations in NS3-108 were comparable for tetramers containing LPVWL**AYK**V (3.64% of live $CD8^+$ cells), LPVWL**AHK**V (3.21%) with the original LPVWL**AYR**V epitope staining (3.04%), whereas the LPVWL**SYK**V analogue (0.43%) showed diminished avidity (**Figure 4.7A to 4.7B**). CD107a upregulation and IFN-γ response in the T cells were, however, detected upon stimulation with either one of the four peptide variants, including LPVWL**SYK**V that yielded low tetramer stainings of NS3-108 (**Figure 4.7C to 4.7E**). The observed double population of CD8 stained T cells observed (**Figure 4.6 and 4.7**) is similar to published literature, which concluded that the intermediately stained $CD8^+$ T cells were CD161 high and share common differentiation patterns with T_H17 cells that play an important role in tissue inflammation (277). This can only be confirmed with further experiments, such as cell surface marker staining with CD161, on this subset of CD8 T cell.

The capacity of HLA tetramers to consistently highlight those antigens that CTLs have the highest affinity for, which was more demanding to uncover from cytokine production and degranulation possibly through the confounding effects of bystander activation, emphasized the benefits of HLA tetramer analysis over ICS. While we lacked appropriate clinical data in this study, such as viruses' serotype during each active infection or viral genotype, we believe that factors such as original antigenic sin, which is a documented phenomenon in Dengue infection, can also be explored with the supplementation of relevant informations and additional samples (232, 278).

4.3.D. Post-hoc analysis of T cell epitope discovery.

We deliberately chose not to employ bioinformatic epitope predictions to select our candidate antigens, even though their adoption is commonplace (149, 150, 279, 280), but were curious to assess whether state-of-the-art algorithms could faithfully reproduce the outcome of our discovery efforts. We selected *NetMHCpan* for this computational approach, which (A) has the ability to estimate the binding capability of peptides even for HLA products for which experimental data is scarce -as is the case for most HLAs reported here-, (B) can accommodate epitopes of variable length, and (C) is publicly accessible (281-284). We performed IC_{50} predictions for all possible 8, 9, 10 and 11-mer epitopes (encompassing 2716 SARS-CoV, 6256 HBV, and 12690 DENV peptides) of each HLA-A, -B, and -C variant of the appropriate donor's genotype (**Table 4.1**). For data inclusion, we imposed a 1000 nM affinity threshold, which is less stringent than the conventional IC_{50} cutoff of 500 nM, while providing an estimated specificity of 0.95 and sensitivity of 0.74 (282). All peptides were subsequently ranked by their IC_{50} value, and for each experimentally-derived epitope the rank was determined within the complete HLA constellation of the PBMC donor as well as for the correct restriction element (**Table 4.3**). Even this liberal strategy would instantly preclude the discovery of two novel epitopes. Peptides SITAQPVKI and FLGPLLVLQA were predicted not to bind HLA-B*58:01 and HLA-C*08:01 respectively, with IC_{50} values 10-20 fold above the cutoff. Moreover, predictions can become misguiding, where for instance FLGPLLVLQA was predicted to have an acceptable affinity for HLA-A*02:01 (20.44 nM), while for its true restriction element HLA-C*08:01 the IC_{50} of 22942.68 nM would have excluded it from consideration. Given that predictive methods aim to reduce the magnitude of the peptide collection under investigation, a full-proteome-based bioinformatics screen

should at least encompass no more peptides then a 15-mer overlapping peptide library. In this hypothetical scenario, the total number of peptides to screen with *NetMHCpan* will need to be equal to the collection of our ELISPOT screens (*i.e.* 313 peptides for HBV, 136 peptides for SARS-CoV, and 660 peptides for DENV). Furthermore, the tetramer reagents with conditional ligands developed here comprise 17 different HLA molecules, including the reported HLA-A*02:01 (141, 142), and therefore the number of highest-ranking peptides to screen per HLA will be 18 for HBV, 8 for SARS-CoV and 38 for DENV. Following these criteria, 3 of the novel epitopes out of the total of 7 would not have been discovered based on HLA-specific rankings (*i.e.* SITAQPVKI, SVIWMMWYW, FLGPLLVLQA, **Table 4.3**). If the *NetMHCpan* algorithm would have been employed to curtail the peptide collection to half the size of an ELISPOT screen, the peptides to screen per HLA would be 9 for HBV, 4 for SARS-CoV, and 19 for DENV, and subsequently the majority of novel antigens (*i.e.* MEVTPSGTWL, SITAQPVKI, SVIWMMWYW, FLGPLLVLQA) would have been missed. Although other methods might well outperform on individual allele-specific predictions, employing *netMHCpan* can be considered the most appropriate strategy for PBMC donor cohorts with high HLA polymorphism (281-284), particularly in populations with a preponderance of class I HLA variants for which limited experimentally verified data is available. This does, however, expose current limitations with the deployment of predictive algorithms and validates our approach of using 15-mer overlapping peptide libraries together with high-throughput tetramer screening to identify and characterize novel T cell epitopes.

Table 4.3

Bioinformatic prediction analysis

Identified Sequence	Identified HLA	NetMHCpan prediction				Remarks
		Predicted HLA	Predicted IC_{50} (nM)	Rank (pan-HLA)	Rank (HLA-specific)	
LPVWLAYRV	B*55:02	B*55:02	67.82	254 / 1329	10 / 64	Requires 254 peptide screen
GETALALLLL	B*40:01	B*40:01	8.28	9 / 160	3 / 13	Correct prediction
MEVTPSGTWL	B*40:01	B*40:01	42.57	58 / 263	7 / 13	Requires 58 peptide screen
SITAQPVKI	B*58:01	B*58:01	14252.45	-	-	Excluded by prediction
ASFCGSPYSW	B*58:01	B*58:01	3.63	4 / 908	2 / 181	Correct prediction
SVIWMMWYW	B*58:01	B*58:01	43.18	66 / 680	28 / 166	Requires 66 peptide screen
FLGPLLVLQA	C*08:01	C*08:01	22942.68	-	-	Excluded by prediction

The IC_{50} (nM) values for peptide-binding was predicted using the *NetMHCpan* version 2.3 algorithm for all possible 8, 9, 10, 11-mers within the viral proteins for each HLA-A, -B and -C alleles of the PBMC donor, and ranked accordingly. Peptides with IC_{50} >1000nM were excluded. The complete HLA genotyping of donors is listed in **Table 4.1**.

4.4 Discussion

For the identification of novel epitopes for Asian HLA variants, we exploited a combination of ELISPOT and conditional ligand-based assays (i.e. the HLA-stability ELISA and HLA tetramer arrays) thereby building on the key features each technology affords whilst circumventing their limitations. The ELISPOT assay, which quantifies the active secretion of cytokines, is recognized for its high detection sensitivity compared to assays such as ELISA, bead assays or intracellular cytokine staining (285, 286). The assay furthermore permits screening with pooled overlapping peptides, usage of clinical samples is economical, and lymphocytes maintain their viability, making it amenable for high-throughput proteome-wide antigen screening (285, 286). In ongoing epitope identification efforts for HBV, SARS-CoV and DENV infection, we identified seven 15-mer peptides capable of stimulating T cell responses in volunteer donors from Singapore (**Figure 4.1**). The HLA-restriction and fine-specificity of the epitopes would commonly be inferred from cellular assays. These, however, suffer from simultaneous expression of various HLA variants that complicates data interpretation. We drew on our caged Asian HLAs to measure their receptiveness to accommodate all shorter peptides embedded in the parent 15-mer antigen (**Figure 4.3**). This focused the search on the most likely antigenic sequences while sparing precious sample material, accelerated the identification process, and utilizes low amounts of HLA reagents. Ultimate confirmation of the peptide fine-specificity and HLA restriction element responsible for engendering the T cell response was provided by staining with arrays of HLA tetramers, which allow rapid visualization and quantification of *in vitro* cultured, cryopreserved samples or freshly isolated *ex vivo* antigen-specific $CD8^+$ T cells with high specificity and sensitivity (**Figure 4.4**). This highlights that HLA tetramers not only permit the monitoring and

isolation of antigen-specific T lymphocytes (134, 137, 243), but their usage extends to screening for novel $CD8^+$ T cell antigens (124, 138, 141, 145-147, 149, 150). Notably, in seven out of seven times, the antigenic epitope was correctly identified. The construction of the 16-membered peptide-exchangeable HLA library, excluding HLA-A*02:01, gives sufficient coverage for epitope screening efforts in Asia (24) where the probability of matching the HLA profile of a lymphocyte donor with at least one HLA in our panel is high for SE Asia (HLA-A: 93%, -B: 63%, -C: 79%) as well as the closely related populations from Oceania (HLA-A: 88%, -B:46%, -C:82%). This contrasts with the European (HLA-A: 30%, -B:11%, -C:54%) and North American regions (HLA-A: 74%, -B:5%, -C:74%) for which equivalent molecules have already been designed (124, 141, 142), whereas similar efforts in Sub-Saharan Africa (HLA-A: 11%, -B:23%, -C:43%) or South West Asia (HLA-A: 53%, -B:10%, -C:38%), where HLA diversity is more pronounced, would need a suite of reagents that is proportionally extended. The use of these reagents to distinguish closely related epitopes from different dengue virus variants have also been demonstrated. Finally, we highlighted through retrospective analysis that employing bioinformatic strategies for the prediction of HLA-restricted epitopes would likely have failed to identify a considerable portion of the novel CD8 T cell epitopes described herein.

Advances in cytometry compel HLA tetramer technology to keep pace with equivalent improvements. Higher-dimensional staining for distinct antigen-specific T cell subsets has, for example, become feasible by labeling HLA multimers with unique combinatorial codes (138, 148), or passing the cells through microarray platforms for on-surface inspection (194-196). An emerging technology that labels

cell surface markers with heavy metal ions for CyTOF analysis which, amongst its advantages, eliminates the need for fluorescence compensation, enables simultaneous immunophenotypic tracking of large cellular subsets (139, 243, 287). Furthermore, for various pathologies where T cell immunity is strongly correlated with protection, the exploration of therapeutic approaches is both warranted and increasing met with clinical success (288-290). Re-directed T cells that target virally infected or malignant cells may be generated either by *in vitro* expansion of autologous lymphocytes , or by genetic modification to drive TCR expression of the desired specificity, before being infused back into the patient (288, 291). All such developments vitally rely on accurate definition of antigen-specific T lymphocyte responses, and the HLA libraries that can be constructed from the conditional ligands for Asian variants reported here are essential for the characterization and enumeration of clinically relevant antigen-specific $CD8^+$ T cells. This will enable us to conquer the extensive diversity imposed by the allelic variation in HLA, coupled with their ability to present distinct repertoires of peptides (267).

5

Conclusion and future work

UV-light sensitive conditional ligands have proved to be exceptionally useful in aiding the search of T cell antigenic peptides in mouse models (145-147) and humans (124, 149, 150). They allow for high-throughput production of fluorochrome-linked MHC tetramer libraries that can be employed for screening. CD8 T cell epitopes are determined by identifying MHC tetramers capable of staining antigen-specific T cell populations. In comparison with many available techniques for T cell antigen screening, the selection of MHC tetramer libraries for this task is highly recommended as it (A) is the only technique that provides direct visualization and enumeration of antigen-specific T cell populations by staining the cells for flow cytometry analysis, (B) accurately defines the fine specificity of the epitope based on peptide-MHC and TCR interaction, (C) directly furnishes the MHC restriction of the epitope, and (D) allows for sorting of antigen-specific T cells for further analysis as MHC tetramer staining is non-destructive to samples. Furthermore, epitope screening using MHC tetramers has the advantage of (E) can be conducted with other cell stainings for cell surface or intracellular markers, and also (F) enables high-throughput screening for novel epitopes (244, 292). In view of these attractive features of MHC tetramers as an epitope identification tool, we successfully adopted this technology for the identification of novel dengue virus associated epitopes in a

mouse model by using methodologies aligned with published literature (**Chapter 2**) and presented the following key findings:

1. For D2Y98P-PP1, a virulent dengue virus 2 strain for AG129 mice (IFNα/βR$^{-/-}$ and IFNγR$^{-/-}$) there are 2 H-K^{b} restricted epitopes (NS2A$_{36-44}$ and NS4A$_{109-116}$) and 7 H-2D^{b} restricted epitopes (prM$_{49-59}$, E$_{4-12}$, E$_{262-270}$, NS2A$_{178-186}$, NS4B$_{66-74}$, NS4B$_{99-107}$, NS5$_{237-245}$)
2. The dominant antigen-specific CD8^{+} T cell response is recognized by H-2D^{b} tetramers with NS4B$_{99-107}$ peptide (i.e. YSQVNPITL), and has an increasing response of up to 39.2% of total CD8^{+} T cells towards day 20 post-infection around the time of death.
3. NS4B$_{99-107}$ antigen-specific CD8^{+} T cells are CD44^{high} CXCR3^{high} PD1^{high} CD122^{high} LAG3^{high} CD62L^{low} CD25^{low} CD69^{low} in comparison with the entire T cell population including both CD8^{+} and CD4^{+} cells. The TCRs of the same population of cells displayed a preferential usage of TCRV-β7, -Vβ8 and -Vβ9.

Future work for this dengue mouse model have been described in detail (**Chapter 2.4**). One interesting observation to pursue is to better understand the role of the dominant YSQVNPITL-specific CD8^{+} T cell response.

While discovery of epitopes in mice using MHC tetramers and conditional ligands has been successful, the adoption of the technology for human T cell is posed with challenges and limitations. Unlike mouse models where additional samples with satisfactory uniformity can be easily obtained, the attainments of human clinical isolates are frequently limited. Furthermore, intra-human variations with correspondingly different immune state are common. As the antigen-specific T cell

population is often low, moderate amounts of cells are usually required for each staining (i.e. approximately 200 000 cells per stain) and this restricts the use of MHC tetramers for a full proteomic scan (244). To further complicate matters, the human HLA is highly polymorphic and reagents developed for a particular HLA are unlikely to be useful for other HLA variants, especially those from a different HLA supertype family. With over 5800 class I HLA alleles identified to date (22), it seems daunting to generate enough reagents to cover the entire human population (244). We have demonstrated how these problems could be overcome such that MHC tetramers and conditional ligands may still be applied for epitope identification in the human population, specifically those residing in Asia. We identified the prevalent HLA alleles in the South East Asian (SEA) population and synthesized reagents to target this population group (**Chapter 3**). This addresses the lack of conditional ligands usable in the SEA population, as those previously developed were largely catered towards the Western population (124). Although conditional ligands for all known HLA molecules are not designed, we have provided further proof of concept that our approach of designing conditional ligand is applicable for any HLA of interest, and this is congruent with published literature (124, 141). The core capabilities of our synthesized conditional ligands (**Chapter 3**) are summarized as follows:

1. The probability of matching the HLA profile of a lymphocyte donor from South East Asia with at least one HLA for which we had synthesized conditional ligands for, while excluding HLA-A*02:01, is 93%, 63% and 79% for HLA-A, -B and -C respectively.
2. HLA molecules refolded with the conditional ligands may be adopted for the HLA-stability ELISA, where the binding capability of a peptide to HLA-A*02:01, A*02:03, A*02:06, A*02:07, A*02:11, A*11:01, A*24:02,

A*33:03, B*15:02, B*40:01, B*46:01, B*55:02, B*58:01, C*03:04, C*04:01, C*07:02 and C*08:01 may be tested.

3. Conditional ligands in tetramers may be exchanged with an appropriate peptide using UV-light to form HLA tetramers, and it was shown that tetramers for HLA-A*02:01, A*02:06, A*02:07, A*11:01, A*24:02, A*33:03, B*40:01, B*55:02, B*58:01 and C*08:01 are capable of staining the appropriate antigen-specific T cells.

With our developed conditional ligands and recombinant HLA molecules, we proceeded to identify novel Dengue virus, Hepatitis B and SARS coronavirus epitopes for HLA variants common in Asia using these tools. Some highlights in our epitope identification efforts in the Asian population (**Chapter 4**) are.

1. For overcoming the limitation in the amount of human clinical samples, a full proteome scan of overlapping 15-mer peptides by ELISPOT based on IFN-γ was conducted to identify T cell antigens, which were used to cultivate T cell lines for further epitope fine mapping.
2. Another immunological technique conducted before HLA screenings to save on the consumption of patient samples was the HLA-stability ELISA. All possible 9-, 10- and 11-mer peptides within the 15-mer antigen were scanned for HLA binding capability using this assay.
3. With the HLA tetramers screening synthesized by the peptide-exchange strategy, we successfully confirmed the epitope fine specificity and HLA restriction of 1 novel HLA-B*40:01 restricted epitope for SARS-CoV, 3 HLA-B*58:01 restricted epitopes for SARS-CoV and HBV, 1 HLA-B*55:02 epitope for DENV and 1 HLA-C*08:01 epitope for HBV. The identified

antigenic peptides were capable of stimulating the relevant T cell lines to produce IFN-γ and upregulate CD107a.

4. HLA tetramers of variant peptides from different viral strains were used to facilitate T cell cross-reactivity analysis.
5. Bioinformatic predictions were excluded in our epitope mapping pipeline. While such tools are useful for epitope identification in mouse models, their use is limited for epitopes restricted by Asian HLA variants and especially so for less common HLA such as HLA-B*55:02.

In this study, we demonstrated how the conditional ligands for Asian HLA may be adopted in epitope mapping. By obtaining appropriate clinical samples, epitope identification for the other HLAs for which we have developed conditional ligands should also be possible. The application of our designed conditional ligands may be extended to other uses of the MHC multimeric reagents including (A) characterization of T cells following vaccination against cancer (293, 294) or various viruses such as HIV (295), influenza (296), yellow fever virus (297), (B) assessment of the T cell responses in patients suffering from various diseases such as cancer (298, 299), diabetes (150, 260), hepatitis B (300), hepatitis C (301) or dengue hemorrhagic fever (232), (C) purification of T cells for adoptive transfer (238-240, 302, 303), or (D) the selective deletion of antigen-specific T cell population (241, 242). Recombinant HLA molecules with a conditional ligand are certainly useful as they can expand our current understanding of $CD8^+$ T cell immunology. Future work pertaining to the application of our conditional ligands will largely depend on the availability of suitable clinical samples. Having said that, additional epitope idenfication efforts using our conditional ligand is advantageous not only to fill in the gap of knowledge in antigen-specific T cell responses restricted by these HLAs variants, but to provide a

platform for the correlation of antigen-specific T cell with disease-states and possibly an avenue for future development of T cell based immunotherapies or peptide-based vaccination.

A

Appendices

Appendix A

Appendix A.1

Selected candidate epitopes from D2Y98P and D2Y98-PP1

A. Selected H-2K^{b} candidate epitopes from D2Y98P

No.	Peptide	Position	arb-rank	smm-rank	uda-rank	park-rank	Median Rank
1	MGVTYLAL	1197	20	1	67	2	11
2	RALIFILL	265	14	9	15	11	12.5
3	VSILASSL	1359	31	4	2	215	17.5
4	**VAVSFVTL**	1163	18	3	17	19	17.5
5	ILLWYAQI	2186	15	48	9	23	19
6	ITYKCPLL	163	29	10	5	31	19.5
7	WSYYCGGL	2578	5	5	100	34	19.5
8	RTLRVLNL	2651	23	16	12	109	19.5
9	AVVSVSPL	1300	4	37	1	154	20.5
10	VVTLYLGV	764	36	23	4	78	29.5
11	VSLVLVGV	757	33	22	32	1128	32.5
12	RNTPFNML	9	115	53	13	10	33
13	TAWDFGSL	698	75	30	37	16	33.5
14	**IILEFFLI**	2202	93	18	36	35	35.5
15	VTLITGNM	1168	17	12	55	72	36
16	AGLLFSIM	2476	58	15	120	8	36.5
17	TTFVTPML	2288	35	43	16	39	37
18	VVPMVTQM	2825	32	91	10	42	37
19	RAIWYMWL	2963	168	24	54	3	39
20	CSHHFHEL	3200	233	62	20	18	41
21	VGVVTLYL	762	16	14	214	70	43
22	IQMSSGNL	550	28	34	58	145	46
23	RSCTLPPL	1089	42	47	47	132	47
24	VSVSPLFL	1302	8	19	77	251	48
25	SILASSLL	1360	2	29	79	112	54
26	IGVEPGQL	660	22	13	116	91	56.5
27	LGKSYAQM	3245	60	54	169	29	57

28	FSLGVLGM	1135	40	20	85	250	62.5
29	RFLEFEAL	2973	99	31	244	7	65
30	RHPGFTIM	243	88	42	461	1	65
31	SVTRLENL	834	41	107	29	94	67.5
32	IGMGVTYL	1195	62	8	73	114	67.5
33	IITASILL	2181	1	61	76	113	68.5
34	GILGYSQI	1504	104	102	46	21	74
35	VIITWIGM	741	45	6	119	105	75
36	LNPTAIFL	1328	123	161	41	33	82
37	DIPMTGPL	1370	21	138	513	27	82.5
38	VSWKNKEL	781	67	70	95	177	82.5
39	VSRGSAKL	2546	48	35	121	531	84.5
40	MTMRCIGM	279	27	25	211	142	84.5
41	YIYMGEPL	1947	92	79	203	30	85.5
42	LITEMGRL	2099	39	71	170	101	86
43	MNSRSTSL	748	69	218	52	104	86.5
44	MWLGARFL	2968	85	88	80	286	86.5
45	TILIRTGL	1441	26	33	141	788	87
46	YGAAFSGV	724	50	45	815	127	88.5
47	SGVEGEGL	2996	82	63	656	95	88.5

B. Selected H-2D^{b} 9-mer candidate epitopes from D2Y98P

No.	Peptide	Position	arb-rank	smm-rank	uda-rank	park-rank	Median rank
1	**RMLINRFTM**	2728	2	1	1	5	1.5
2	**TAIANQATV**	2309	1	4	3	25	3.5
3	LMMRTTWAL	2425	5	7	14	297	10.5
4	TMAKNKPTL	313	9	28	12	10	11
5	LQMENKAWL	479	6	42	4	17	11.5
6	**CSQVNPITL**	2342	17	73	11	11	14
7	DALDNLAVL	2116	26	23	10	2	16.5
8	SLLKNDIPM	1365	33	33	2	3	18
9	VMVMVGATM	1183	3	3	55	73	29
10	TAVAPSMTM	273	34	25	35	70	34.5
11	SALTALNDM	3167	194	14	25	46	35.5
12	ASIAARGYI	1767	230	29	36	65	50.5
13	GMLQGRGPL	36	366	47	5	55	51
14	VAATMANEM	2238	7	2	95	235	51
15	**TALTGATEI**	542	163	32	20	71	51.5
16	AAGRKSLTL	2089	18	89	19	874	54
17	**IGMSNRDFV**	284	79	182	7	36	57.5
18	LAVVSVSPL	1299	60	79	6	56	58
19	VPCRNQDEL	3217	95	284	9	21	58
20	YGLNTFTNM	3098	189	49	62	58	60
21	SSPILSITI	1415	55	310	15	68	61.5
22	VNMCTLMAM	145	40	84	40	268	62
23	ASIILEFFL	2200	53	258	61	67	64
24	GMGVTYLAL	1196	27	40	93	210	66.5
25	RAVQTKPGL	1582	82	9	159	57	69.5

26	FQPESPSKL	809	63	356	33	80	71.5
27	FPQSNAPIM	1800	67	266	80	18	73.5
28	SWPLNEAIM	1346	81	155	79	6	80
29	VPNYNLIIM	1750	131	236	31	14	81
30	TSLSVSLVL	753	145	83	89	78	86
31	VQPENLEYT	410	375	150	8	23	86.5
32	VATTFVTPM	2286	104	69	17	240	86.5
33	WALCEALTL	2431	86	48	88	97	87
34	FLLVAHYAI	2355	374	93	23	83	88
35	GMALFLEEM	1141	20	108	70	211	89
36	VPMVTQMAM	2826	11	116	63	247	89.5
37	VALVPHVGM	207	512	134	18	48	91
38	IVLLSQSTI	1235	58	241	16	127	92.5
39	ALVAFLRFL	49	66	338	97	88	92.5
40	LTVWNRVWI	3301	173	82	106	8	94
41	VAVSFVTLI	1163	324	120	48	75	97.5
42	VAAEMEEAL	1702	31	10	169	189	100
43	VSILASSLL	1359	1165	136	30	66	101
44	LMMTTIGIV	1228	38	43	160	1021	101.5
45	VILQNAWKV	1286	89	121	129	28	105
46	GGLKNVREV	2583	214	175	37	24	106
47	LAVTIMAIL	1272	13	292	121	93	107

C. Selected H-2K^{b} 8-mer candidate epitopes from D2Y98P-PP1

No.	Peptide	Position	arb-rank	smm-rank	uda-rank	park-rank	Median Rank
1	RVLIFILL	265	8	8	7	11	8
2	MGVTYLAL	1197	22	1	71	2	12
3	**ITYNCPLL**	163	25	9	8	31	17
4	**VAVSFVTL**	1163	20	3	16	19	17.5
5	VSILASSL	1359	32	4	2	214	18
6	ILLWYAQI	2186	15	46	10	23	19
7	WSYYCGGL	2578	6	5	106	35	20.5
8	AVVSVSPL	1300	5	38	1	157	21.5
9	RTLRVLNL	2651	26	17	12	109	21.5
10	VIPMVTQM	2825	9	24	22	36	23
11	VVTLYLGV	764	36	23	4	81	29.5
12	ITLTAALL	2348	45	13	15	117	30
13	VSLVLVGV	757	33	21	32	1138	32.5
14	VSVSPLLL	1302	19	28	38	270	33
15	RNTPFNML	9	116	53	13	10	33
16	AGLLFSIM	2476	54	16	123	8	35
17	TAWDFGSL	698	76	31	39	15	35
18	**IILEFFLI**	2202	96	18	36	37	36.5
19	VTLITGNM	1168	18	11	61	75	39.5
20	CSHHFHEL	3200	234	62	19	17	40.5
21	RAIWYMWL	2963	168	25	59	3	42
22	VGVVTLYL	762	17	14	218	73	45
23	RSCTLPPL	1089	41	45	50	134	47.5

24	IQMSSGNL	550	30	35	64	147	49.5
25	ITPMFRHS	2291	762	15	53	50	51.5
26	AIGCYSQV	2338	55	136	28	48	51.5
27	SILASSLL	1360	4	30	82	114	56
28	LGKSYAQM	3245	58	54	170	29	56
29	IGVEPGQL	660	24	12	120	93	58.5
30	FSLGVLGM	1135	39	19	89	247	64
31	RHPGFTIM	243	90	42	464	1	66
32	RFLEFEAL	2973	101	32	248	7	66.5
33	SVTRLENL	834	40	113	29	96	68
34	IGMGVTYL	1195	60	7	77	116	68.5
35	IITASILL	2181	3	61	80	115	70.5
36	IRVPNYNL	1748	1	22	127	420	74.5
37	GILGYSQI	1504	105	107	48	21	76.5
38	LNPTAIFL	1328	125	168	43	33	84
39	VSRGSAKL	2546	44	36	126	532	85
40	VSWKNKEL	781	66	71	100	180	85.5
41	DIPMTGPL	1370	23	146	518	27	86.5
42	LITEMGRL	2099	38	72	171	101	86.5
43	MNSRSTSL	748	69	229	57	104	86.5
44	KIVQPENL	408	42	64	124	110	87
45	YIYMGEPL	1947	95	81	206	30	88
46	YGAAFSGV	724	47	43	827	129	88
47	SGVEGEGL	2996	83	63	661	97	90

D. Selected H-2D^{b} 9-mer candidate epitopes from D2Y98P-PP1

No	Peptide	Position	arb-rank	smm-rank	uda-rank	park-rank	Median rank
1	**RMLINRFTM**	2728	2	1	1	4	1.5
2	**TAIANQATV**	2309	1	4	5	27	4.5
3	NALDNLAVL	2116	17	15	4	1	9.5
4	TMAKNKPTL	313	8	27	12	10	11
5	LMMRTTWAL	2425	5	7	16	304	11.5
6	LQMENKAWL	479	6	41	6	19	12.5
7	SLLKNDIPM	1365	29	32	2	3	16
8	**YSQVNPITL**	2342	22	76	3	11	16.5
9	VMVMVGATM	1183	3	3	52	78	27.5
10	TAVAPSMTM	273	31	23	34	74	32.5
11	SALTALNDM	3167	194	12	25	50	37.5
12	VAATMANEM	2238	7	2	93	242	50
13	**IGISNRDFV**	284	93	73	11	28	50.5
14	ASIAARGYI	1767	228	28	35	69	52
15	GMLQGRGPL	36	365	46	7	59	52.5
16	**TALTGATEI**	542	163	31	21	75	53
17	AAGRKSLTL	2089	16	90	20	874	55
18	VPCRNQDEL	3217	94	278	10	23	58.5
19	YGLNTFTNM	3098	188	48	59	62	60.5
20	LAVVSVSPL	1299	62	79	8	60	61
21	VNMCTLMAM	145	41	85	39	276	63
22	SSPILSITI	1415	55	304	17	72	63.5

23	ASIILEFFL	2200	53	253	58	71	64.5
24	GMGVTYLAL	1196	26	39	91	217	65
25	RAVQTKPGL	1582	80	8	160	61	70.5
26	FPQSNAPIM	1800	68	264	77	21	72.5
27	FQPESPSKL	809	65	347	32	85	75
28	SWPLNEAIM	1346	79	157	76	5	77.5
29	VPNYNLIIM	1750	131	230	30	15	80.5
30	WALCEALTL	2431	84	47	86	100	85
31	TSLSVSLVL	753	144	84	87	83	85.5
32	VQPENLEYT	410	372	150	9	25	87.5
33	GMALFLEEM	1141	19	108	68	218	88
34	VATTFITPM	2286	95	87	14	231	91
35	LTVWNRVWI	3301	173	83	102	8	92.5
36	VALVPHVGM	207	512	133	19	52	92.5
37	ALVAFLRFL	49	67	332	95	92	93.5
38	IVLLSQSTI	1235	58	235	18	133	95.5
39	VAAEMEEAL	1702	28	9	170	197	99
40	VAVSFVTLI	1163	322	118	47	80	99
41	LMMTTIGIV	1228	36	42	161	1017	101.5
42	VSILASSLL	1359	1167	135	29	70	102.5
43	VILQNAWKV	1286	88	119	130	31	103.5
44	AKLTNTTTA	343	308	178	13	35	106.5
45	LAVTIMAIL	1272	12	286	120	96	108
46	GVLGMALFL	1138	72	16	150	183	111
47	YMPSVIEKM	2677	174	124	89	99	111.5

All possible (**A**) H-2K^b 8-mer peptides or (**B**) H-2D^b 9-mer peptides from the proteome of the prominent variant strain of D2Y98P as well as all possible (**C**) H-2K^b 8-mer peptides or (**D**) H-2D^b 9-mer peptides from the proteome of the plaque purified clone of D2Y98P (D2Y98P-PP1) were predicted using four scoring matrices and ranked for each matrix. The median values of all four ranks were obtained, and the peptides were ranked according to their median ranks. The top 47 peptides, the identified epitope (red), and identified epitope variants (blue) based on later experiments are indicated here.

Appendix A.2

D2Y98P 8-mer and 9-mer peptides containing serine residue at 178 in NS2A and phenylalanine at 47 in NS4B

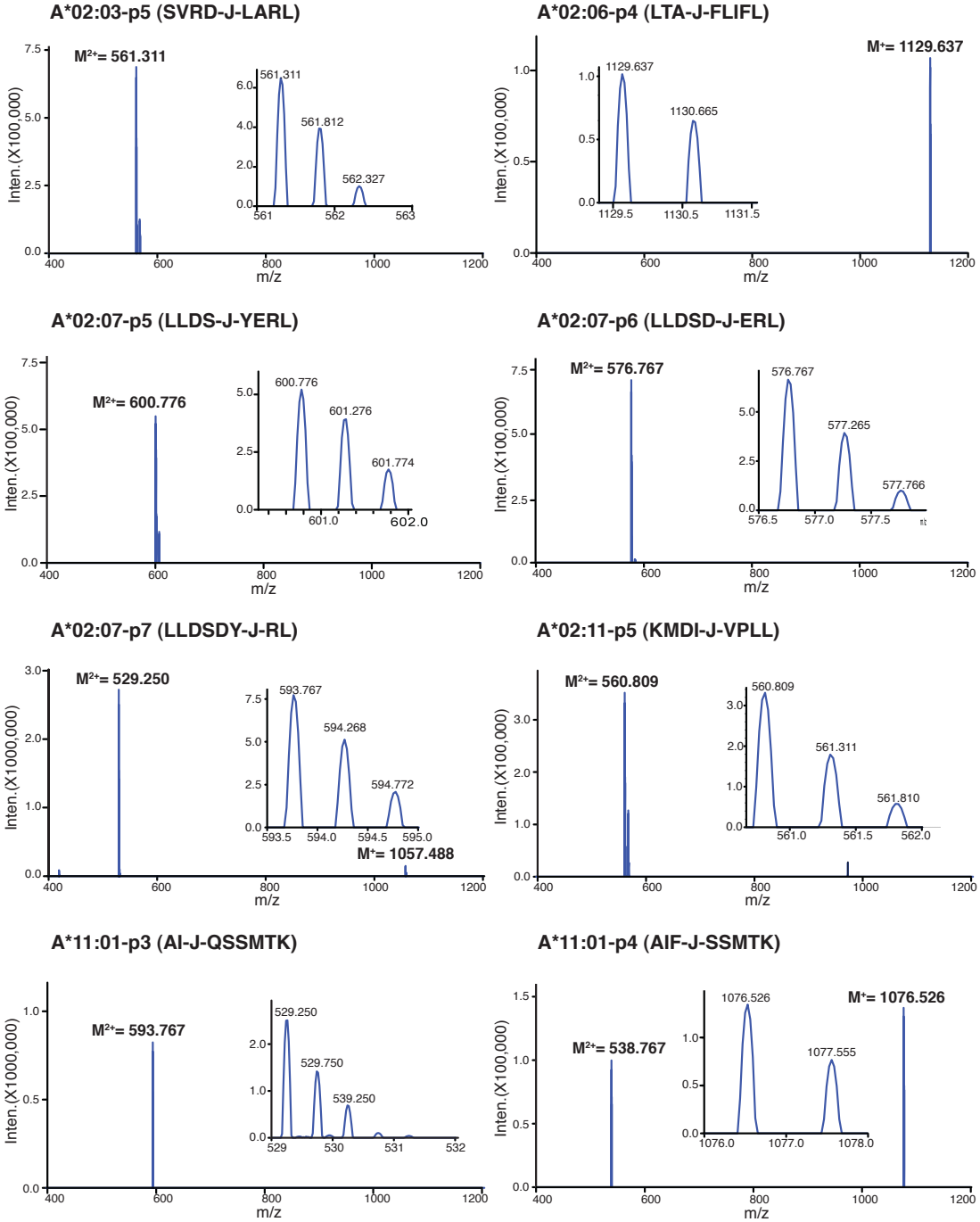
No	Peptide	Protein	Start	End	Length
1	ILAVVSV**S**	NS2A	171	178	8-mer
2	LAVVSV**S**P	NS2A	172	179	8-mer
3	AVVSV**S**PL	NS2A	173	180	8-mer
4	VVSV**S**PLF	NS2A	174	181	8-mer
5	VSV**S**PLFL	NS2A	175	182	8-mer
6	SV**S**PLFLT	NS2A	176	183	8-mer
7	V**S**PLFLTS	NS2A	177	184	8-mer
8	**S**PLFLTSS	NS2A	178	185	8-mer
9	LYAVATT**F**	NS4B	40	47	8-mer
10	YAVATT**F**V	NS4B	41	48	8-mer
11	AVATT**F**VT	NS4B	42	49	8-mer
12	VATT**F**VTP	NS4B	43	50	8-mer
13	ATT**F**VTPM	NS4B	44	51	8-mer
14	TT**F**VTPML	NS4B	45	52	8-mer
15	T**F**VTPMLR	NS4B	46	53	8-mer
16	**F**VTPMLRH	NS4B	47	54	8-mer
17	TILAVVSV**S**	NS2A	170	178	9-mer
18	ILAVVSV**S**P	NS2A	171	179	9-mer
19	LAVVSV**S**PL	NS2A	172	180	9-mer
20	AVVSV**S**PLF	NS2A	173	181	9-mer
21	VVSV**S**PLFL	NS2A	174	182	9-mer
22	VSV**S**PLFLT	NS2A	175	183	9-mer
23	SV**S**PLFLTS	NS2A	176	184	9-mer
24	V**S**PLFLTSS	NS2A	177	185	9-mer
25	**S**PLFLTSSQ	NS2A	178	186	9-mer
26	TLYAVATT**F**	NS4B	39	47	9-mer
27	LYAVATT**F**V	NS4B	40	48	9-mer
28	YAVATT**F**VT	NS4B	41	49	9-mer
29	AVATT**F**VTP	NS4B	42	50	9-mer
30	VATT**F**VTPM	NS4B	43	51	9-mer
31	ATT**F**VTPML	NS4B	44	52	9-mer
32	TT**F**VTPMLR	NS4B	45	53	9-mer
33	T**F**VTPMLRH	NS4B	46	54	9-mer
34	**F**VTPMLRHS	NS4B	47	55	9-mer

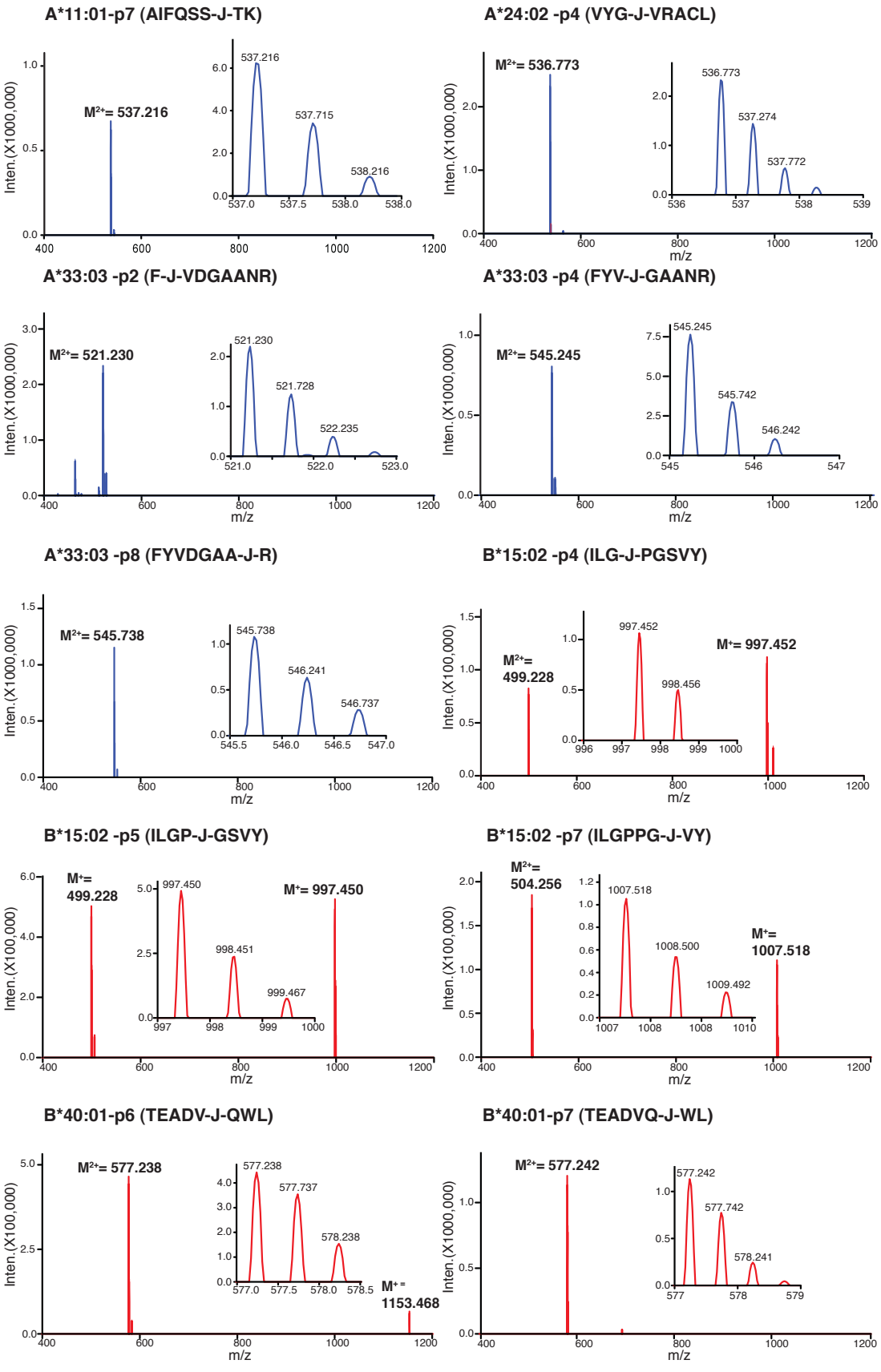
The D2Y98P virus strain differ from its earlier predecessor, the P13 strain, by mutation F178S in NS2A and a I47F substitution in NS4B protein. All possible 8-mer and 9-mer peptides containing the mutated residue (red), and the position of the mutated residue in the protein are shown above.

Appendix B

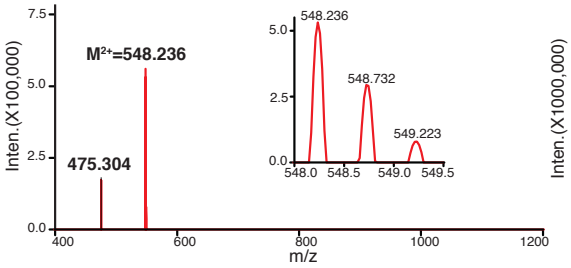
Appendix B.1

Mass spectrometry (MS) analysis of ligands

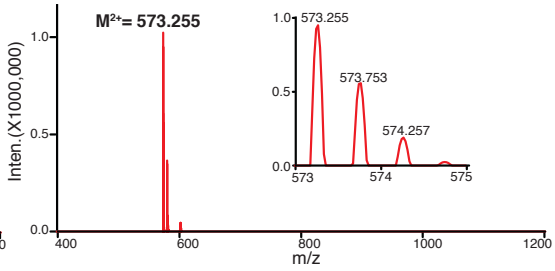



A*11:01-p7 (AIFQSS-J-TK)
M²⁺= 537.216
537.216
537.715
538.216
A*24:02 -p4 (VYG-J-VRACL)
M²⁺= 536.773
536.773
537.274
537.772
A*33:03 -p2 (F-J-VDGAANR)
M²⁺= 521.230
521.230
521.728
522.235
A*33:03 -p4 (FYV-J-GAANR)
M²⁺= 545.245
545.245
545.742
546.242
A*33:03 -p8 (FYVDGAA-J-R)
M²⁺= 545.738
545.738
546.241
546.737
B*15:02 -p4 (ILG-J-PGSVY)
M²⁺= 499.228
997.452
998.456
M⁺= 997.452
B*15:02 -p5 (ILGP-J-GSVY)
M⁺= 499.228
997.450
998.451
999.467
M⁺= 997.450
B*15:02 -p7 (ILGPPG-J-VY)
M²⁺= 504.256
1007.518
1008.500
1009.492
M⁺= 1007.518
B*40:01-p6 (TEADV-J-QWL)
M²⁺= 577.238
577.238
577.737
578.238
M⁺ = 1153.468
B*40:01-p7 (TEADVQ-J-WL)
M²⁺= 577.242
577.242
577.742
578.241
Inten.(X1000,000)
Inten.(X100,000)
m/z

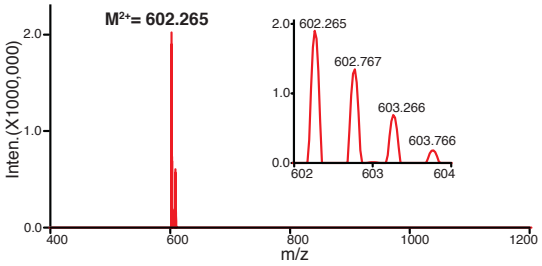
B*40:01-p8 (TEADVQQ-J-L)



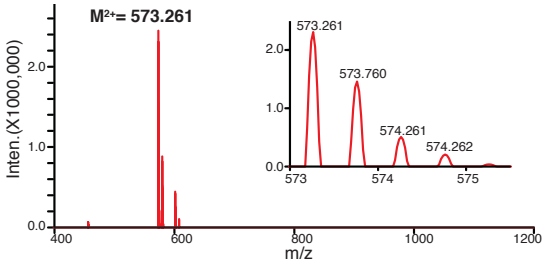
B*46:01-p4 (KMK-J-IAEAY)



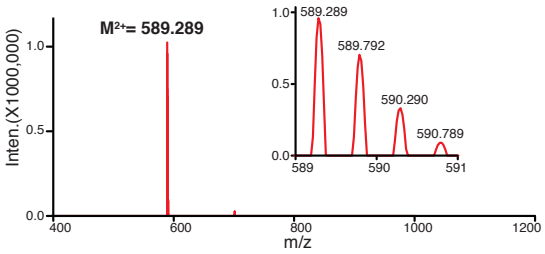
B*46:01-p6 (KMKEI-J-EAY)



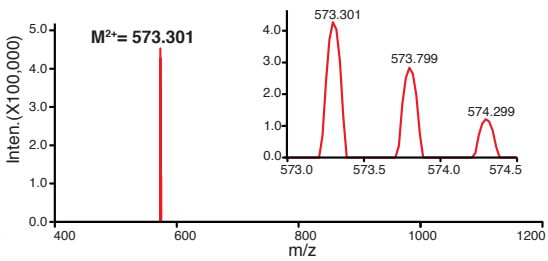
B*46:01-p7 (KMKEIA-J-AY)



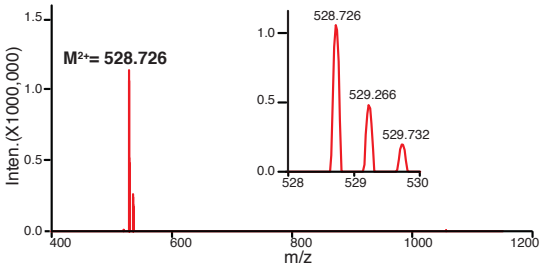
B*55:02-p5 (KPWD-J-IPMV)



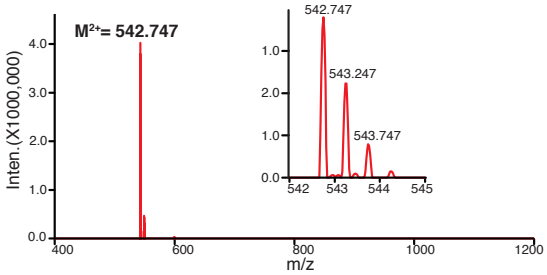
B*55:02-p8 (KPWDVIP-J-V)



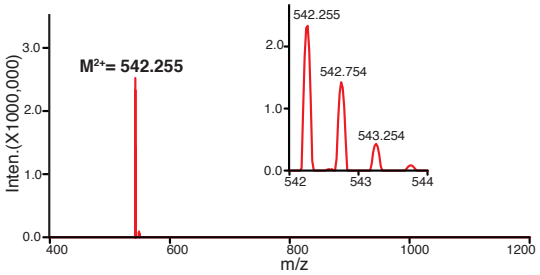
B*58:01-p4 (ISA-J-GQELF)



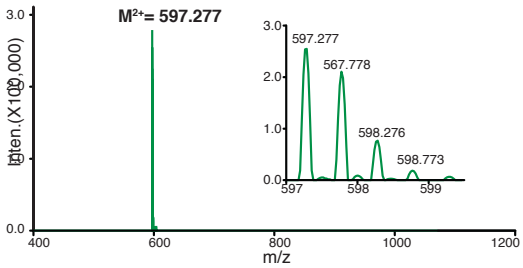
B*58:01-p6 (ISARG-J-ELF)

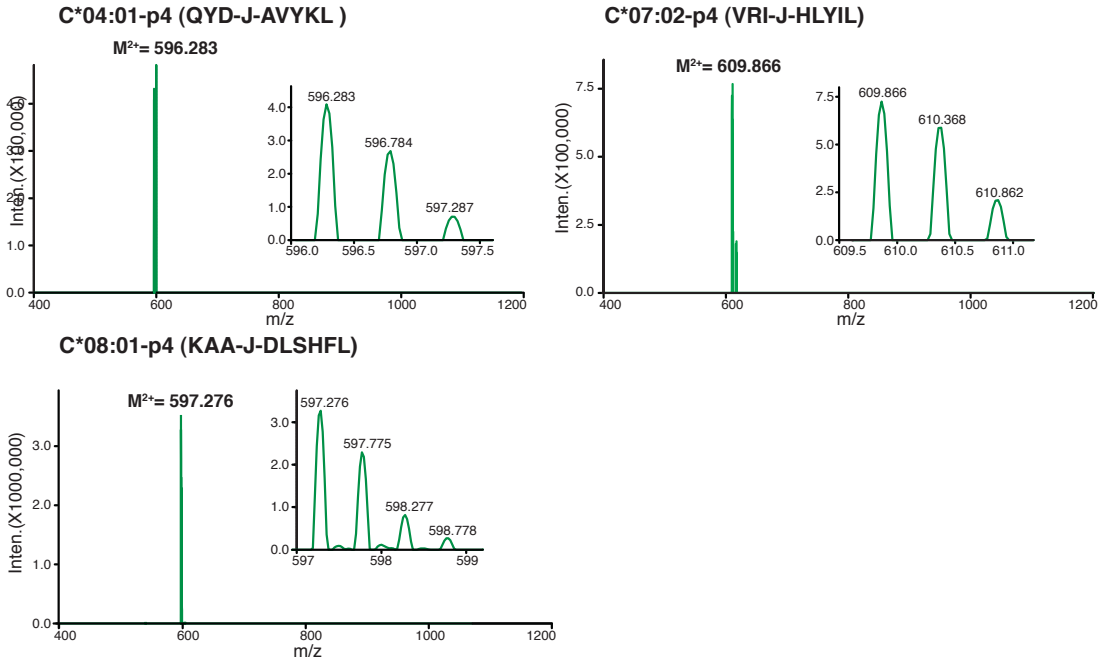


B*58:01-p7 (ISARGQ-J-LF)



C*03:04-p5 (FVYG-J-SKTSL)





Mass spectrometry (MS) analysis with isotopic distribution of all synthesized conditional ligands was conducted by LCMS-IT- TOF. The mass to charge ratios (m/z) of the peptides were measured after liquid chromatography (LC) separation on a C18 column. The amino acid sequences of conditional ligands analyzed was indicated, where the letter J represents the photocleavable 3-amino-3-(2-nitro) phenyl-propanoic acid (Anp) residue. As both isomers of Anp were used during solid phase peptide synthesis of the conditional ligand, 2 diasteroisomers could usually be identified. In some cases, however, the two isomers failed to separate during chromatography. The data displayed is for one of the two isomers, mostly for doubly charged molecules (M2+).

Appendix B.2
Gene and protein sequences of the pertinent heavy chains.

HLA	Type	Sequence
A*02:03	Gene	AGGAGATATACCATGGGCTCGCACTCGATGCGTTACTTCTTTACCTCTGTGTCGCGTC CGGGTCGTGGTGAACCGCGTTTCATTGCCGTTGGCTATGTCGATGACACCCAGTTTGT TCGTTTCGATAGCGACGCGGCCTCTCAACGTATGGAACCGCGCGCGCCGTGGATTGAA CAGGAAGGCCCGGAATATTGGGATGGTGAAACCCGCAAAGTCAAAGCCCATTCCCAGA CCCACCGTGTGGATCTGGGTACGCTGCGTGGTTATTACAACCAGAGTGAAGCAGGCTC CCATACCGTGCAACGTATGTACGGCTGCGATGTTGGTAGCGACTGGCGTTTTCTGCGC GGCTATCACCAGTATGCATACGATGGTAAAGACTACATCGCTCTGAAAGAAGATCTGC GTTCTTGGACCGCAGCTGACATGGCGGCCCAAACCACGAAACATAAATGGGAAACCGC GCACGAAGCCGAACAGTGGCGCGCATATCTGGAAGGCACGTGCGTTGAATGGCTGCGT CGCTACCTGGAAAATGGTAAAGAAACCCTGCAGCGTACGGATGCTCCGAAAACCCATA TGACGCATCACGCAGTCTCAGATCACGAAGCTACCCTGCGCTGTTGGGCGCTGTCGTT TTATCCGGCCGAAATTACCCTGACGTGGCAGCGTGATGGTGAAGACCAGACCCAAGAT ACGGAACTGGTTGAAACCCGTCCGGCAGGTGACGGCACGTTCCAAAAATGGGCAGCTG TGGTTGTCCCGAGCGGTCAGGAACAACGTTATACCTGTCATGTGCAGCACGAAGGTCT GCCGAAACCGCTGACGCTGCGCCTGCCGGAAACGGGCGGTCTGAACGACATCTTTGAA GCACAGAAAATTGAATGGCATGAATGAGGATCCGAATTCG
	Protein	MGSHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQEGP EYWDGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQ YAYDGKDYIALKEDLRSWTAADMAAQTTKHKWETAHEAEQWRAYLEGTCVEWLRRYLE NGKETLQRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGEDQTQDTELV ETRPAGDGTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRLPETGGLNDIFEAQKI EWHE
A*02:06	Gene	AGGAGATATACCATGGGCAGTCACTCAATGCGTTACTTTTACACCAGCGTGTCTCGTC CGGGTCGTGGCGAACCGCGTTTCATCGCCGTGGGCTATGTCGATGACACCCAGTTTGT CCGTTTCGATAGCGACGCGGCCTCTCAACGTATGGAACCGCGCGCACCGTGGATTGAA CAGGAAGGCCCGGAATATTGGGATGGTGAAACCCGCAAAGTCAAAGCTCATTCCCAGA CCCACCGTGTGGATCTGGGTACGCTGCGTGGTTATTACAACCAGAGTGAAGCGGGCTC CCATACCGTGCAACGTATGTACGGCTGCGATGTTGGTAGCGACTGGCGTTTTCTGCGC GGCTATCACCAGTATGCGTACGATGGTAAAGACTACATCGCCCTGAAAGAAGATCTGC GTTCTTGGACCGCAGCTGACATGGCGGCCCAAACCACGAAACATAAATGGGAAGCAGC TCACGTTGCAGAACAGCTGCGCGCTTATCTGGAAGGCACGTGCGTCGAATGGCTGCGT CGCTACCTGGAAAATGGTAAAGAAACCCTGCAGCGTACGGATGCACCGAAAACCCATA TGACGCATCACGCGGTGTCAGATCACGAAGCCACCCTGCGCTGTTGGGCACTGTCGTT TTATCCGGCTGAAATTACCCTGACGTGGCAGCGTGATGGTGAAGACCAGACCCAAGAT ACGGAACTGGTGGAAACCCGTCCGGCAGGTGACGGCACGTTCCAAAAATGGGCAGCAG TGGTTGTCCCGAGCGGTCAGGAACAACGTTATACCTGTCATGTTCAGCACGAAGGTCT GCCGAAACCGCTGACGCTGCGCCTGCCGGAAACGGGCGGTCTGAACGACATCTTTGAA GCACAAAAAATTGAATGGCATGAATGAGGATCCGAATTCG
	Protein	MGSHSMRYFYTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQEGP EYWDGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQ YAYDGKDYIALKEDLRSWTAADMAAQTTKHKWEAAHVAEQLRAYLEGTCVEWLRRYLE NGKETLQRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGEDQTQDTELV ETRPAGDGTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRLPETGGLNDIFEAQKI EWHE
A*02:07	Gene	CCATGGGCAGCCATAGCATGCGTTACTTTTTCACCAGCGTGAGCCGTCCGGGCCGTGG CGAACCGCGTTTTATTGCGGTGGGCTATGTGGATGATACCCAGTTTGTGCGTTTTGAT AGCGATGCGGCGAGCCAGCGTATGGAACCGCGTGCGCCGTGGATTGAACAGGAAGGCC CGGAATATTGGGATGGCGAAACCCGTAAAGTGAAAGCGCATAGCCAGACCCATCGTGT GGATCTGGGCACCCTGCGTGGCTATTATAACCAGAGCGAAGCGGGCAGCCATACCGTG CAGCGTATGTGCGGCTGCGATGTGGGCAGCGATTGGCGTTTTCTGCGTGGCTATCATC AGTATGCGTATGATGGCAAAGATTATATTGCGCTGAAAGAAGATCTGCGTAGCTGGAC CGCGGCGGATATGGCGGCGCAGACCACCAAACATAAATGGGAAGCGGCGCATGTGGCG GAACAGCTGCGTGCGTATCTGGAAGGCACCTGCGTGGAATGGCTGCGTCGTTATCTGG AAAACGGCAAAGAAACCCTGCAGCGTACCGATGCGCCGAAAACCCATATGACCCATCA TGCCGTGAGCGATCATGAAGCGACCCTGCGTTGCTGGGCGCTGAGCTTTTATCCGGCG GAAATTACCCTGACCTGGCAGCGTGATGGCGAAGATCAGACCCAGGATACCGAACTGG TGGAAACCCGTCCGGCAGGTGATGGTACCTTTCAGAAATGGGCAGCAGTGGTGGTGCC GAGCGGTCAGGAACAGCGTTATACCTGCCATGTGCAGCATGAAGGCCTGCCGAAACCG CTGACCCTGCGTCTGCCGGAAACCGGCGGCCTGAACGATATTTTTGAAGCGCAGAAAA TTGAATGGCATGAATAAGGATCC
	Protein	MGSHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQEGP EYWDGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMCGCDVGSDWRFLRGYHQ YAYDGKDYIALKEDLRSWTAADMAAQTTKHKWEAAHVAEQLRAYLEGTCVEWLRRYLE NGKETLQRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGEDQTQDTELV

		ETRPAGDGTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRLPETGGLNDIFEAQKI EWHE
A*02:11	Gene	AGGAGATATACCATGGGCTCGCACTCGATGCGTTACTTCTTTACCTCAGTTTCCCGTC CGGGTCGTGGTGAACCGCGTTTTATTGCCGTTGGCTATGTCGATGACACCCAGTTTGT CCGTTTCGATAGCGACGCGGCCTCTCAACGTATGGAACCGCGCGCACCGTGGATTGAA CAGGAAGGCCCGGAATATTGGGATGGTGAAACCCGCAAAGTCAAAGCTCATTCCCAGA TCGATCGTGTGGACCTGGGTACCCTGCGTGGTTATTACAACCAGAGTGAAGCGGGCTC CCATACGGTGCAACGTATGTACGGCTGCGATGTTGGTAGCGACTGGCGTTTTCTGCGC GGCTATCACCAGTATGCGTACGATGGTAAAGACTACATTGCCCTGAAAGAAGATCTGC GTTCTTGGACCGCAGCTGACATGGCGGCCCAAACCACGAAACATAAATGGGAAGCAGC TCACGTTGCAGAACAGCTGCGCGCTTATCTGGAAGGCACGTGCGTCGAATGGCTGCGT CGCTACCTGGAAAATGGTAAAGAAACCCTGCAGCGTACGGATGCACCGAAAACCCATA TGACGCATCACGCGGTGTCAGATCACGAAGCCACCCTGCGCTGTTGGGCACTGTCGTT TTATCCGGCTGAAATCACCCTGACGTGGCAGCGTGATGGTGAAGACCAGACCCAAGAT ACGGAACTGGTGGAAACCCGTCCGGCAGGTGACGGCACGTTCCAAAAATGGGCAGCAG TGGTTGTCCCGAGCGGTCAGGAACAACGTTATACCTGTCATGTTCAGCACGAAGGTCT GCCGAAACCGCTGACGCTGCGCCTGCCGGAAACGGGTGGCCTGAACGACATCTTTGAA GCCCAGAAAATTGAATGGCATGAATGAGGATCCGAATTCG
	Protein	MGSHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQEGP EYWDGETRKVKAHSQIDRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQ YAYDGKDYIALKEDLRSWTAADMAAQTTKHKWEAAHVAEQLRAYLEGTCVEWLRRYLE NGKETLQRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGEDQTQDTELV ETRPAGDGTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRLPETGGLNDIFEAQKI EWHE
A*11:01	Gene	-
	Protein	MRYFYTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQEGPEYWDQ ETRNVKAQSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGPDGRFLRGYRQDAYDG KDYIALNEDLRSWTAADMAAQITKRKWEAAHAAEQQRAYLEGRCVEWLRRYLENGKET LQRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTELVETRPA GDGTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWELSSQPGT
A*24:02	Gene	-
	Protein	MRYFSTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQEGPEYWDE ETGKVKAHSQTDRENLRIALRYYNQSEAGSHTLQMMFGCDVGSDGRFLRGYHQYAYDG KDYIALKEDLRSWTAADMAAQITKRKWEAAHVAEQQRAYLEGTCVDGLRRYLENGKET LQRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTELVETRPA GDGTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWEPSSQPGSLHHILDAQKMVW NHR
A*33:03	Gene	AGGAGATATACCATGGGCAGCCATTCTATGCGTTATTTTACCACGAGTGTGTCCCGCC CGGGTCGTGGCGAACCGCGCTTCATTGCGGTTGGTTACGTCGATGACACCCAGTTTGT GCGTTTCGATTCAGACGCCGCATCGCAACGCATGGAACCGCGTGCTCCGTGGATCGAA CAGGAAGGCCCGGAATATTGGGATCGCAACACGCGTAATGTTAAAGCGCACAGCCAAA TTGACCGCGTCGATCTGGGTACCCTGCGTGGCTACTATAACCAGTCTGAAGCCGGTAG TCATACGATCCAAATGATGTACGGCTGCGACGTGGGTTCCGATGGCCGCTTTCTGCGT GGTTATCAGCAAGACGCATACGATGGCAAAGACTATATTGCTCTGAATGAAGATCTGC GCTCATGGACCGCGGCCGACATGGCAGCTCAGATCACGCAACGTAAATGGGAAGCGGC CCGCGTTGCAGAACAGCTGCGTGCTTACCTGGAAGGTACCTGTGTCGAATGGCTGCGC CGTTATCTGGAAAACGGCAAAGAAACGCTGCAACGCACCGATCCGCCGAAAACGCACA TGACCCATCACGCGGTGTCGGACCATGAAGCCACGCTGCGTTGCTGGGCACTGAGCTT CTACCCGGCTGAAATTACCCTGACGTGGCAGCGCGATGGTGAAGATCAAACCCAGGAT ACGGAACTGGTTGAAACCCGTCCGGCAGGCGATGGTACCTTTCAGAAATGGGCCTCTG TCGTGGTTCCGTCAGGCCAGGAACAGCGCTATACCTGTCACGTGCAGCATGAAGGTCT GCCGAAACCGCTGACCCTGCGTCTGCCGGAAACCGGCGGTCTGAATGATATCTTCGAA GCGCAGAAAATTGAATGGCACGAATAAGGATCCGAATTCG
	Protein	MGSHSMRYFTTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQEGP EYWDRNTRNVKAHSQIDRVDLGTLRGYYNQSEAGSHTIQMMYGCDVGSDGRFLRGYQQ DAYDGKDYIALNEDLRSWTAADMAAQITQRKWEAARVAEQLRAYLEGTCVEWLRRYLE NGKETLQRTDPPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGEDQTQDTELV ETRPAGDGTFQKWASVVVPSGQEQRYTCHVQHEGLPKPLTLRLPETGGLNDIFEAQKI EWHE
B*15:02	Gene	AGGAGATATACCATGGGCAGCCATTCTATGCGTTATTTTTACACCGCGATGAGTCGCC CGGGTCGTGGCGAACCGCGCTTCATTGCCGTGGGTTATGTTGATGACACGCAGTTTGT CCGTTTCGATTCCGACGCAGCTTCACCGCGCATGGCGCCGCGTGCACCGTGGATCGAA CAAGAAGGTCCGGAATACTGGGATCGCAACACCCAGATTTCGAAAACGAATACCCAAA CGTATCGTGAAAGCCTGCGCAACCTGCGTGGTTACTATAATCAGTCTGAAGCAGGCAG TCACATCATTCAACGCATGTACGGTTGCGACGTGGGCCCGGATGGTCGTCTGCTGCGC GGCTATGACCAGTCCGCTTACGATGGTAAAGACTATATCGCGCTGAACGAAGATCTGT CATCGTGGACCGCCGCAGACACGGCTGCGCAAATTACCCAGCGTAAATGGGAAGCCGC ACGCGAAGCTGAACAACTGCGTGCGTACCTGGAAGGCCTGTGTGTTGAATGGCTGCGC CGTTATCTGGAAAATGGTAAAGAAACGCTGCAGCGCGCCGATCCGCCGAAAACCCATG

		TCACGCACCATCCGATCAGTGACCACGAAGCAACCCTGCGTTGCTGGGCTCTGGGCTT TTACCCGGCAGAAATTACGCTGACCTGGCAACGCGATGGTGAAGACCAGACGCAAGAT ACCGAACTGGTGGAAACGCGTCCGGCCGGCGATCGCACCTTCCAGAAATGGGCGGCCG TTGTCGTTCCGTCTGGTGAAGAACAGCGTTATACCTGTCATGTGCAGCACGAAGGCCT GCCGAAACCGCTGACCCTGCGCCTGCCGGAAACCGGTGGCCTGAACGATATCTTTGAA GCGCAGAAAATTGAATGGCATGAATAAGGATCCGAATTCG
	Protein	MGSHSMRYFYTAMSRPGRGEPRFIAVGYVDDTQFVRFDSDAASPRMAPRAPWIEQEGP EYWDRNTQISKTNTQTYRESLRNLRGYYNQSEAGSHIIQRMYGCDVGPDGRLLRGYDQ SAYDGKDYIALNEDLSSWTAADTAAQITQRKWEAAREAEQLRAYLEGLCVEWLRRYLE NGKETLQRADPPKTHVTHHPISDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTELV ETRPAGDRTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRLPETGGLNDIFEAQKI EWHE
B*40:01	Gene	AGGAGATATACCATGGGCAGCCATTCTATGCGCTATTTTCATACCGCTATGAGCCGTC CGGGTCGCGGTGAACCGCGTTTCATTACGGTGGGTTACGTTGATGACACCCTGTTTGT TCGTTTCGATAGTGACGCTACCTCCCCGCGTAAAGAACCGCGTGCCCCGTGGATTGAA CAGGAAGGCCCGGAATATTGGGATCGCGAAACCCAGATCAGCAAAACCAACACGCAAA CCTACCGTGAATCTCTGCGTAACCTGCGCGGCTATTACAATCAGTCAGAAGCAGGTTC GCATACCCTGCAACGCATGTATGGTTGCGATGTTGGTCCGGATGGTCGTCTGCTGCGT GGTCACAACCAGTATGCCTACGATGGTAAAGACTACATTGCACTGAATGAAGATCTGC GTAGTTGGACCGCCGCGGACACCGCAGCTCAGATCTCCCAACGTAAACTGGAAGCCGC ACGTGTTGCCGAACAGCTGCGTGCATATCTGGAAGGCGAATGCGTGGAATGGCTGCGT CGCTACCTGGAAAATGGTAAAGATAAACTGGAACGTGCGGACCCGCCGAAAACGCATG TGACCCATCACCCGATTAGCGATCACGAAGCAACCCTGCGCTGTTGGGCTCTGGGCTT TTATCCGGCGGAAATCACGCTGACCTGGCAGCGTGATGGTGAAGACCAGACGCAAGAT ACCGAACTGGTTGAAACCCGTCCGGCGGGTGACCGTACCTTCCAAAAATGGGCAGCTG TGGTTGTCCCGTCTGGCGAAGAACAGCGCTATACCTGTCATGTCCAACACGAAGGTCT GCCGAAACCGCTGACCCTGCGTCTGCCGGAAACCGGCGGTCTGAACGATATTTTTGAA GCACAGAAAATCGAATGGCATGAATAAGGATCCGAATTCG
	Protein	MGSHSMRYFHTAMSRPGRGEPRFITVGYVDDTLFVRFDSDATSPRKEPRAPWIEQEGP EYWDRETQISKTNTQTYRESLRNLRGYYNQSEAGSHTLQRMYGCDVGPDGRLLRGHNQ YAYDGKDYIALNEDLRSWTAADTAAQISQRKLEAARVAEQLRAYLEGECVEWLRRYLE NGKDKLERADPPKTHVTHHPISDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTELV ETRPAGDRTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRLPETGGLNDIFEAQKI EWHE
B*46:01	Gene	AGGAGATATACCATGGGCAGTCATTCCATGCGCTATTTTTATACCGCTATGAGCCGTC CGGGTCGTGGTGAACCGCGTTTCATTGCAGTGGGTTATGTTGATGACACGCAGTTTGT GCGTTTCGATTCAGATGCCGCAAGCCCGCGTATGGCACCGCGTGCTCCGTGGATCGAA CAGGAAGGCCCGGAATATTGGGATCGCGAAACCCAAAAATACAAACGTCAGGCCCAAA CGGACCGCGTTAGTCTGCGTAACCTGCGCGGCTATTACAATCAGAGCGAAGCGGGTTC TCATACCCTGCAACGTATGTATGGTTGCGATGTTGGTCCGGATGGTCGTCTGCTGCGT GGTCATGATCAGTCCGCATATGATGGTAAAGACTACATTGCACTGAACGAAGATCTGA GCAGCTGGACCGCAGCTGACACCGCCGCACAGATCACCCAACGTAAATGGGAAGCAGC TCGCGAAGCGGAACAGTGGCGTGCCTATCTGGAAGGCCTGTGCGTCGAATGGCTGCGT CGCTACCTGGAAAATGGTAAAGAAACCCTGCAGCGTGCCGATCCGCCGAAAACCCATG TGACGCATCACCCGATTAGCGATCACGAAGCAACGCTGCGCTGTTGGGCTCTGGGCTT TTATCCGGCGGAAATCACCCTGACGTGGCAGCGTGATGGTGAAGACCAGACCCAAGAT ACGGAACTGGTTGAAACCCGTCCGGCGGGTGACCGTACCTTCCAAAAATGGGCCGCGG TGGTTGTCCCGTCTGGTGAAGAACAGCGCTACACCTGTCATGTGCAACACGAAGGTCT GCCGAAACCGCTGACCCTGCGTCTGCCGGAAACCGGTGGTCTGAACGATATTTTTGAA GCGCAGAAAATCGAATGGCACGAATAAGGATCCGAATTCG
	Protein	MGSHSMRYFYTAMSRPGRGEPRFIAVGYVDDTQFVRFDSDAASPRMAPRAPWIEQEGP EYWDRETQKYKRQAQTDRVSLRNLRGYYNQSEAGSHTLQRMYGCDVGPDGRLLRGHDQ SAYDGKDYIALNEDLSSWTAADTAAQITQRKWEAAREAEQWRAYLEGLCVEWLRRYLE NGKETLQRADPPKTHVTHHPISDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTELV ETRPAGDRTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRLPETGGLNDIFEAQKI EWHE
B*55:02	Gene	AGGAGATATACCATGGGCTCGCACTCTATGCGTTACTTTTACACGGCTATGTCGCGCC CGGGTCGTGGTGAACCGCGTTTCATCGCAGTCGGCTACGTCGATGACACCCAGTTTGT TCGCTTCGATAGTGACGCAGCATCCCCGCGTGAAGAACCGCGTGCTCCGTGGATTGAA CAGGAAGGCCCGGAATATTGGGATCGCAACACCCAAATCTACAAAGCACAGGCTCAAA CGGATCGTGAATCCCTGCGTAACCTGCGCGGCTATTACAATCAGTCAGAAGCAGGTTC GCATACCTGGCAAACGATGTATGGTTGCGATCTGGGTCCGGACGGTCGTCTGCTGCGT GGTCACAACCAGCTGGCATATGATGGTAAAGACTACATTGCACTGAATGAAGATCTGA GCTCTTGGACCGCAGCTGACACGGCAGCACAGATCACCCAACGTAAATGGGAAGCAGC TCGTGTGGCAGAACAGCTGCGTGCATATCTGGAAGGCACCTGCGTTGAATGGCTGCGT CGCTACCTGGAAAATGGTAAAGAAACGCTGCAGCGTGCCGATCCGCCGAAAACCCATG TCACGCATCACCCGATTAGCGATCACGAAGCAACCCTGCGCTGTTGGGCTCTGGGCTT TTATCCGGCGGAAATCACCCTGACGTGGCAGCGTGATGGTGAAGACCAGACCCAAGAT ACGGAACTGGTCGAAACCCGTCCGGCAGGTGACCGTACGTTCCAAAAATGGGCAGCAG

		TGGTTGTCCCGTCTGGTGAAGAACAGCGCTACACCTGTCATGTGCAACACGAAGGTCT GCCGAAACCGCTGACGCTGCGTCTGCCGGAAACGGGCGGTCTGAACGACATCTTTGAA GCACAGAAAATTGAATGGCATGAATGAGGATCCGAATTCG
	Protein	MGSHSMRYFYTAMSRPGRGEPRFIAVGYVDDTQFVRFDSDAASPREEPRAPWIEQEGP EYWDRNTQIYKAQAQTDRESLRNLRGYYNQSEAGSHTWQTMYGCDLGPDGRLLRGHNQ LAYDGKDYIALNEDLSSWTAADTAAQITQRKWEAARVAEQLRAYLEGTCVEWLRRYLE NGKETLQRADPPKTHVTHHPISDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTELV ETRPAGDRTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRLPETGGLNDIFEAQKI EWHE
B*58:01	Gene	AGGAGATATACCATGGGCAGTCATTCCATGCGTTATTTTTATACCGCCATGAGCCGTC CGGGTCGTGGTGAACCGCGTTTCATTGCCGTGGGTTATGTTGATGACACCCAGTTTGT TCGTTTTGATTCAGATGCCGCAAGCCCGCGTACCGAACCGCGTGCCCCGTGGATTGAA CAGGAAGGCCCGGAATATTGGGATGGTGAAACCCGTAACATGAAAGCAAGTGCTCAAA CGTACCGTGAAAACCTGCGCATCGCTCTGCGTTATTACAATCAGAGCGAAGCAGGCTC TCATATTATCCAACGCATGTATGGTTGCGATCTGGGTCCGGATGGTCGTCTGCTGCGT GGTCATGATCAGTCCGCATATGATGGTAAAGACTACATTGCACTGAACGAAGATCTGA GCAGCTGGACCGCAGCTGACACCGCCGCACAGATCACCCAACGCAAATGGGAAGCAGC TCGTGTTGCGGAACAGCTGCGCGCCTATCTGGAAGGCCTGTGCGTCGAATGGCTGCGT CGCTACCTGGAAAATGGTAAAGAAACCCTGCAGCGTGCCGATCCGCCGAAAACCCATG TCACGCATCACCCTGTGAGCGACCACGAAGCAACGCTGCGTTGTTGGGCTCTGGGCTT TTATCCGGCGGAAATCACCCTGACGTGGCAGCGCGATGGTGAAGACCAGACCCAAGAT ACGGAACTGGTTGAAACCCGCCCGGCAGGCGACCGTACGTTCCAAAAATGGGCCGCGG TGGTTGTCCCGTCTGGTGAAGAACAGCGTTACACCTGTCATGTGCAACACGAAGGTCT GCCGAAACCGCTGACCCTGCGTCTGCCGGAAACCGGTGGTCTGAATGATATTTTTGAA GCGCAGAAAATCGAATGGCACGAATAAGGATCCGAATTCG
	Protein	MGSHSMRYFYTAMSRPGRGEPRFIAVGYVDDTQFVRFDSDAASPRTEPRAPWIEQEGP EYWDGETRNMKASAQTYRENLRIALRYYNQSEAGSHIIQRMYGCDLGPDGRLLRGHDQ SAYDGKDYIALNEDLSSWTAADTAAQITQRKWEAARVAEQLRAYLEGLCVEWLRRYLE NGKETLQRADPPKTHVTHHPVSDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTELV ETRPAGDRTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRLPETGGLNDIFEAQKI EWHE
C*03:04	Gene	AGGAGATATACCATGGGCAGTCATTCTATGCGTTATTTCTATACGGCAGTTAGTCGTC CGGGTCGTGGCGAACCGCACTTCATTGCTGTGGGCTATGTGGATGACACCCAGTTTGT CCGTTTCGATAGCGACGCAGCATCTCCGCGTGGTGAACCGCGTGCA CCGTGGGTGGAACAGGAAGGTCCGGAATATTGGGATCGTGAAACCCAAAAATACAAAC GTCAGGCCCAAACGGATCGCGTTTCACTGCGTAACCTGCGCGGCTATTACAATCAGAG TGAAGCAGGTTCCCATATTATCCAACGCATGTATGGTTGCGATGTCGGTCCGGACGGT CGTCTGCTGCGTGGTTATGATCAGTATGCTTACGATGGTAAAGACTACATTGCGCTGA ACGAAGATCTGCGTTCGTGGACCGCAGCTGACACGGCAGCACAGATCACCCAACGTAA ATGGGAAGCAGCTCGCGAAGCAGAACAGCTGCGTGCTTATCTGGAAGGCCTGTGCGTT GAATGGCTGCGTCGCTACCTGAAAAATGGTAAAGAAACCCTGCAGCGCGCTGAACATC CGAAAACCCACGTGACGCATCACCCGGTTAGCGATCACGAAGCGACCCTGCGTTGTTG GGCCCTGGGCTTTTATCCGGCAGAAATTACCCTGACGTGGCAGTGGGATGGTGAAGAC CAGACCCAAGATACGGAACTGGTTGAAACCCGCCCGGCAGGCGACGGCACGTTCCAAA AATGGGCAGCAGTGGTTGTCCCGAGTGGTGAAGAACAGCGCTACACCTGTCATGTGCA ACACGAAGGTCTGCCGGAACCGCTGACGCTGCGTCTGCCGGAAACGGGTGGCCTGAAC GACATCTTTGAAGCACAGAAAATTGAATGGCATGAATAAGGATCCGAATTCG
	Protein	MGSHSMRYFYTAVSRPGRGEPHFIAVGYVDDTQFVRFDSDAASPRGEPRAPWVEQEGP EYWDRETQKYKRQAQTDRVSLRNLRGYYNQSEAGSHIIQRMYGCDVGPDGRLLRGYDQ YAYDGKDYIALNEDLRSWTAADTAAQITQRKWEAAREAEQLRAYLEGLCVEWLRRYLK NGKETLQRAEHPKTHVTHHPVSDHEATLRCWALGFYPAEITLTWQWDGEDQTQDTELV ETRPAGDGTFQKWAAVVVPSGEEQRYTCHVQHEGLPEPLTLRLPETGGLNDIFEAQKI EWHE
C*04:01	Gene	AGGAGATATACCATGGGCTCTCACTCTATGCGTTACTTTTCTACCTCAGTCAGTTGGC CGGGTCGTGGCGAACCGCGTTTCATTGCTGTTGGCTACGTTGATGACACCCAGTTTGT CCGTTTCGATAGCGACGCAGCATCTCCGCGTGGTGAACCGCGCGAA CCGTGGGTGGAACAGGAAGGTCCGGAATACTGGGATCGCGAAACCCAAAAATATAAAC GTCAGGCACAAGCTGATCGCGTTAACCTGCGTAAACTGCGCGGCTATTACAATCAGAG TGAAGATGGTTCCCATACGCTGCAACGTATGTTTGGTTGCGATCTGGGTCCGGACGGT CGTCTGCTGCGTGGTTATAACCAGTTCGCATACGATGGTAAAGACTATATTGCGCTGA ATGAAGATCTGCGTTCTTGGACCGCAGCTGACACGGCAGCACAGATCACCCAACGTAA ATGGGAAGCAGCTCGTGAAGCAGAACAGCGTCGCGCATACCTGGAAGGCACCTGCGTG GAATGGCTGCGTCGCTATCTGGAAAACGGTAAAGAAACGCTGCAGCGTGCCGAACATC CGAAAACCCACGTGACGCATCACCCGGTTTCAGATCACGAAGCAACCCTGCGCTGTTG GGCTCTGGGCTTTTACCCGGCCGAAATTACCCTGACGTGGCAGTGGGATGGTGAAGAC CAGACCCAAGATACGGAACTGGTTGAAACCCGTCCGGCGGGCGACGGCACGTTCCAAA AATGGGCAGCAGTGGTTGTCCCGTCGGGTGAAGAACAGCGTTATACCTGTCATGTCCA ACACGAAGGTCTGCCGGAACCGCTGACGCTGCGCCTGCCGGAAACGGGTGGCCTGAAC GACATCTTTGAAGCCCAGAAAATTGAATGGCATGAATAAGGATCCGAATTCG

	Protein	MGSHSMRYFSTSVSWPGRGEPRFIAVGYVDDTQFVRFDSDAASPRGEPREPWVEQEGP EYWDRETQKYKRQAQADRVNLRKLRGYYNQSEDGSHTLQRMFGCDLGPDGRLLRGYNQ FAYDGKDYIALNEDLRSWTAADTAAQITQRKWEAAREAEQRRAYLEGTCVEWLRRYLE NGKETLQRAEHPKTHVTHHPVSDHEATLRCWALGFYPAEITLTWQWDGEDQTQDTELV ETRPAGDGTFQKWAAVVVPSGEEQRYTCHVQHEGLPEPLTLRLPETGGLNDIFEAQKI EWHE
C*07:02	Gene	AGGAGATATACCATGGGCTCACACTCGATGCGTTACTTTGATACGGCAGTCAGTCGTC CGGGTCGTGGCGAACCGCGTTTCATTTCTGTCGGCTATGTTGATGACACCCAGTTTGT GCGTTTCGATAGCGACGCAGCATCTCCGCGTGGTGAACCGCGTGCA CCGTGGGTTGAACAGGAAGGTCCGGAATACTGGGATCGTGAAACCCAAAAATATAAAC GTCAGGCGCAAGCCGACCGCGTTAGCCTGCGTAACCTGCGCGGCTATTACAATCAGAG TGAAGATGGTTCCCATACGCTGCAACGCATGTCTGGTTGCGATCTGGGTCCGGACGGT CGTCTGCTGCGTGGTTATGATCAGTCAGCCTACGATGGTAAAGACTATATTGCACTGA ACGAAGATCTGCGTTCGTGGACCGCAGCTGACACGGCAGCACAGATCACCCAACGTAA ACTGGAAGCAGCTCGTGCAGCAGAACAGCTGCGTGCTTACCTGGAAGGCACCTGCGTC GAATGGCTGCGTCGCTATCTGGAAAATGGTAAAGAAACGCTGCAGCGCGCAGAACCGC CGAAAACCCATGTGACGCATCACCCGCTGAGTGATCACGAAGCAACCCTGCGTTGTTG GGCTCTGGGCTTTTACCCGGCGGAAATTACCCTGACGTGGCAGCGCGATGGTGAAGAC CAGACCCAAGATACGGAACTGGTCGAAACCCGTCCGGCAGGTGACGGCACGTTCCAGA AATGGGCAGCTGTGGTTGTCCCGTCCGGTCAGGAACAACGCTATACCTGTCATATGCA GCACGAAGGTCTGCAAGAACCGCTGACGCTGAGCCTGCCGGAAACGGGTGGCCTGAAC GACATCTTTGAAGCACAGAAAATTGAATGGCATGAATGAGGATCCGAATTCG
	Protein	MGSHSMRYFDTAVSRPGRGEPRFISVGYVDDTQFVRFDSDAASPRGEPRAPWVEQEGP EYWDRETQKYKRQAQADRVSLRNLRGYYNQSEDGSHTLQRMSGCDLGPDGRLLRGYDQ SAYDGKDYIALNEDLRSWTAADTAAQITQRKLEAARAAEQLRAYLEGTCVEWLRRYLE NGKETLQRAEPPKTHVTHHPLSDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTELV ETRPAGDGTFQKWAAVVVPSGQEQRYTCHMQHEGLQEPLTLSLPETGGLNDIFEAQKI EWHE
C*08:01	Gene	AGGAGATATACCATGGGCAGCCACTCGATGCGTTACTTTTACACGGCAGTTTCGCGTC CGGGTCGTGGTGAACCGCGTTTCATTGCAGTCGGCTATGTCGATGACACCCAGTTTGT CCAATTCGATAGCGACGCAGCATCTCCGCGTGGCGAACCGCGTGCACCGTGGGTGGAA CAGGAAGGTCCGGAATACTGGGATCGCGAAACCCAAAAATATAAACGTCAGGCCCAAA CGGATCGCGTTTCACTGCGTAACCTGCGCGGCTATTACAATCAGAGTGAAGCAGGTTC CCATACCCTGCAACGCATGTATGGTTGCGATCTGGGTCCGGACGGTCGTCTGCTGCGT GGTTATAACCAGTTTGCTTACGATGGTAAAGACTATATTGCGCTGAATGAAGATCTGC GTTCGTGGACCGCAGCTGACACGGCAGCACAGATCACCCAACGTAAATGGGAAGCAGC TCGCACGGCAGAACAGCTGCGTGCTTACCTGGAAGGCACCTGCGTGGAATGGCTGCGT CGCTATCTGGAAAACGGCAAGAAAACCCTGCAGCGCGCTGAACATCCGAAAACCCACG TGACGCATCACCCGGTTAGCGATCACGAAGCGACCCTGCGTTGTTGGGCCCTGGGCTT TTACCCGGCAGAAATTACCCTGACGTGGCAGCGCGATGGTGAAGACCAGACCCAAGAT ACGGAACTGGTTGAAACCCGTCCGGCAGGCGACGGCACGTTCCAGAAATGGGCAGCAG TGGTTGTCCCGAGTGGTGAAGAACAGCGCTATACCTGTCATGTCCAACACGAAGGTCT GCCGGAACCGCTGACGCTGCGTCTGCCGGAAACGGGCGGTCTGAACGACATCTTTGAA GCACAAAAAATTGAATGGCATGAATGAGGATCCGAATTCG
	Protein	MGSHSMRYFYTAVSRPGRGEPRFIAVGYVDDTQFVQFDSDAASPRGEPRAPWVEQEGP EYWDRETQKYKRQAQTDRVSLRNLRGYYNQSEAGSHTLQRMYGCDLGPDGRLLRGYNQ FAYDGKDYIALNEDLRSWTAADTAAQITQRKWEAARTAEQLRAYLEGTCVEWLRRYLE NGKKTLQRAEHPKTHVTHHPVSDHEATLRCWALGFYPAEITLTWQRDGEDQTQDTELV ETRPAGDGTFQKWAAVVVPSGEEQRYTCHVQHEGLPEPLTLRLPETGGLNDIFEAQKI EWHE

The HLA genes were cloned into pET-28a(+) vector at the NcoI (CCATGG) and BamHI (GGATCC) sites. The restriction sites (underline) and cloning overhangs (red text) are indicated. These genes were optimized for *E. coli* expression by GenScript (Piscataway, NJ). Plamids containing the gene for HLA-A*11:01 (PTCF48) and HLA-A*24:02 (PTCF6) were gifts from the NIH tetramer core facility, and the plasmid sequences can be found in http://tetramer.yerkes.emory.edu/client/protocols#1. The translated products are indicated below each optimized gene sequence.

Appendix B.3

List of known CTL epitopes used in this study.

#	Epitope Sequence	HLA	Organism	Protein	Location	ID/Ref
1	CLGGLLTMV	A*02:01 A*02:07	Epstein–Barr virus	Latent Membrane Proteins 2	426 to 434	6568
2	NLVPMVATV	A*02:01	Human Cytomegalovirus	65 kDa lower matrix phosphoprotein	495 to 503	44920
3	GILGFVFTL	A*02:01 A*02:06	Influenza A virus	Matrix protein 1	58 to 66	20354
4	FLPSDFFPSV	A*02:01 A*02:03 A*02:06 A*02:07	Heptatitis B Virus	Core protein	18 to 27	16833
5	FLLTKILTI	A*02:01	Hepatitis B virus	Pre-S/S protein	194 to 202	16753
6	GLSRYVARL	A*02:01	Hepatitis B virus	Polymerase	453 to 461	21145
7	AVFDRKSDAK	A*11:01	Epstein–Barr virus	EBNA4-NP	399 to 408	5316
8	SIIPSGPLK	A11	Influenza A virus	Matrix Protein	13 to 21	58567
9	ATIGTAMYK	A*11:01	Epstein–Barr virus	Transcription activator BRLF1	134 to 142	5002
10	TYGPVFMSL	A24	Epstein–Barr virus	Latent Membrane Proteins 2	419 to 427	67350
11	QYDPVAALF	A*24:02	Human Cytomegalovirus	65 kDa lower matrix phosphoprotein	341 to 349	52886
12	SVNVHNPTGR	A*33:03	Human Cytomegalovirus	65 kDa lower matrix phosphoprotein	91 to 100	62266
13	DTPVLPHETR	A33	Human Cytomegalovirus	65 kDa lower matrix phosphoprotein	31 to 40	10453
14	SVNVHNPTG	A*33:03	Human Cytomegalovirus	65 kDa lower matrix phosphoprotein	91 to 99	(-)
15	LEKARGSTY	B62	Epstein–Barr virus	Epstein–Barr nuclear antigen-3B	406 to 414	35533
16	GQGGSPTAM	B62	Epstein–Barr virus	Epstein-Barr nuclear antigen 4	831-839	21870
17	LLDFVRFMGV	B62	Epstein–Barr virus	Epstein-Barr nuclear antigen 6	284 to 293	37153
18	HERNGFTVL	B*40:01	Human Cytomegalovirus	65 kDa lower matrix phosphoprotein	267 to 275	23732
19	KEVNSQLSL	B*40:01	Human Cytomegalovirus	Immediate Early Protein 1	42 to 50	30660
20	CEDVPSGKL	B*40:01	Human Cytomegalovirus	65 kDa lower matrix phosphoprotein	232 to 240	6156
21	HLAAQGMAY	B*46:01	Epstein–Barr virus	Epstein-Barr nuclear antigen 3	318 to 326	24170
22	CSAEWPTF	B*46:01	Friend murine leukemia virus	Gag polyprotein	39 to 46	6962
23	VQPPQLTQV	B*46:01	Epstein–Barr virus	Epstein–Barr nuclear antigen-3A	617 to 625	70624
24	KPWDVIPMV	B*55:02	Dengue virus 2	Polyprotein (Non-structural 5 protein)	2822 to 2830	32940
25	KPWDIIPMV	not known	Dengue virus 2	Polyprotein (Non-structural 5 protein)	2822 to 2830	32939
26	KPWDVVPMV	not known	Dengue virus 2	Polyprotein (Non-structural 5 protein)	2822 to 2830	32943
27	KPWDVLPMV	not known	Dengue virus 2	Polyprotein (Non-structural 5 protein)	2822 to 2830	32941
28	RGINDRNFW	B*58:01	Influenza A virus	Nucleoprotein	199 to 207	53890
29	IALYLQQNW		Epstein–Barr virus	Latent Membrane Proteins 1	156 to 164	25351
30	VSFIEFVGW	B*58:01	Epstein–Barr virus	Epstein–Barr nuclear antigen-3B	279 to 287	70932

31	AYAAQGYKVL	C*03:04	Hepatitis C virus	Polyprotein	1243 to 1252	5714
32	PYFVRAQGLI	C*03:04	Hepatitis C virus	Polyprotein	910 to 919	50089
33	FVYGGSKTSL	C*03:04	Epstein–Barr virus	Epstein–Barr nuclear antigen-1	508 to 517	18328
34	KDVALRHVV	C*04:01	Human Cytomegalovirus	65 kDa lower matrix phosphoprotein	181 to 189	122808
35	PYKRIEELL	C*04:01	Human T cell lymphotrophic virus type 1	Protein Tax-1	187 to 195	50124
36	VYPEYVIQY	C*07:02	Homosapien	A904L tumor cell line, cDNA clone 31.2	659 to 685	(304)
37	KRQEILDLWVY	C*07:02	Human immunodeficiency virus 1	Nef protein	105 to 115	(305)
38	VRIGHLYIL	C*07:02	Homosapien	Human bladder carcinoma line LB831-BLC, MAGE-A12 protein	170 to 178	(253)
39	KAAVDLSHFL	C*08:01	Human immunodeficiency virus 1	Nef protein	80 to 89	101766
40	VVCAHELVC	C*08:01	Human Cytomegalovirus	65 kDa lower matrix phosphoprotein	198 to 206	71613

Previously discovered antigenic peptides were employed to generate HLAs monomers or tetramers by UV-mediated peptide exchange (253, 304, 306, 307). The arbritarily allocated peptide number used in this study (#), epitope sequence, HLA restriction element, organism and protein source of the peptide, its location within the protein and the literarture references are given. ID stands for the epitope identification number in the immune epitope database and analysis resource (254). For (-), further information is not available

Appendix B.4

Assays used to identify known CTL epitopes that are used in this study.

#	Epitope Sequence	HLA	Assays used for identification*					
			[3H]-TdR	^{51}Cr	ELISA	ELISPOT	ICS	Tetramer
1	CLGGLLTMV	A*02:01	+	+	+	+	+	+
		A*02:07		+		+		
2	NLVPMVATV	A*02:01	+	+	+	+	+	+
3	GILGFVFTL	A*02:01	+	+	+	+	+	+
		A*02:06				+		
4	FLPSDFFPSV	A*02:01	+	+	+	+	+	+
		A*02:03		+				
		A*02:06		+				
		A*02:07		+				
5	FLLTKILTI	A*02:01				+		+
6	GLSRYVARL	A*02:01		+	+	+		
7	AVFDRKSDAK	A*11:01		+	+	+		+
8	SIIPSGPLK	A11		+				
9	ATIGTAMYK	A*11:01		+		+		+
10	TYGPVFMSL	A24				+		
11	QYDPVAALF	A*24:02		+	+	+	+	+
12	SVNVHNPTGR	A*33:03					+	
13	DTPVLPHETR	A33				+	+	
14	SVNVHNPTG	A*33:03						
15	LEKARGSTY	B62					+	
16	GQGGSPTAM	B62					+	
17	LLDFVRFMGV	B62		+				
18	HERNGFTVL	B*40:01		+		+	+	+
19	KEVNSQLSL	B*40:01				+	+	
20	CEDVPSGKL	B*40:01		+		+	+	+
21	HLAAQGMAY	B*46:01		+		+		+
22	CSAEWPTF	B*46:01				+		
23	VQPPQLTQV	B*46:01				+		
24	KPWDVIPMV	B*55:02		+		+		
25	KPWDIIPMV	not known		+		+		
26	KPWDVVPMV	not known		+		+		
27	KPWDVLPMV	not known		+		+		
28	RGINDRNFW	B*58:01				+		
29	IALYLQQNW	B58				+		
30	VSFIEFVGW	B*58:01		+		+		
31	AYAAQGYKVL	C*03:04		+		+		
32	PYFVRAQGLI	C*03:04		+		+		
33	FVYGGSKTSL	C*03:04				+		+
34	KDVALRHVV	C*04:01			+			
35	PYKRIEELL**	C*04:01						
36	VYPEYVIQY	C*07:02			+			
37	KRQEILDLWVY	C*07:02				+		
38	VRIGHLYIL	C*07:02		+				
39	KAAVDLSHFL	C*08:01				+		
40	VVCAHELVC	C*08:01		+		+		

(*) Assays used to identify antigenic peptides previously (253, 304, 306, 307) were indicated (+) and includes 3H-thymidine cell proliferation ([3H]-TdR), ^{51}Cr release killing (^{51}Cr), IFN-γ based ELISA (ELISA), IFN-γ based ELISPOT (ELISPOT), Intracellular cytokine staining for IFN-γ (ICS), HLA multimer staining (Tetramer). (**) Peptide PYKRIEELL was identified by nonradioactive cytotoxicity assay for lactate dehydrogenase (LDH) release.

Appendix B.5

HLA typing of PBMCs donors

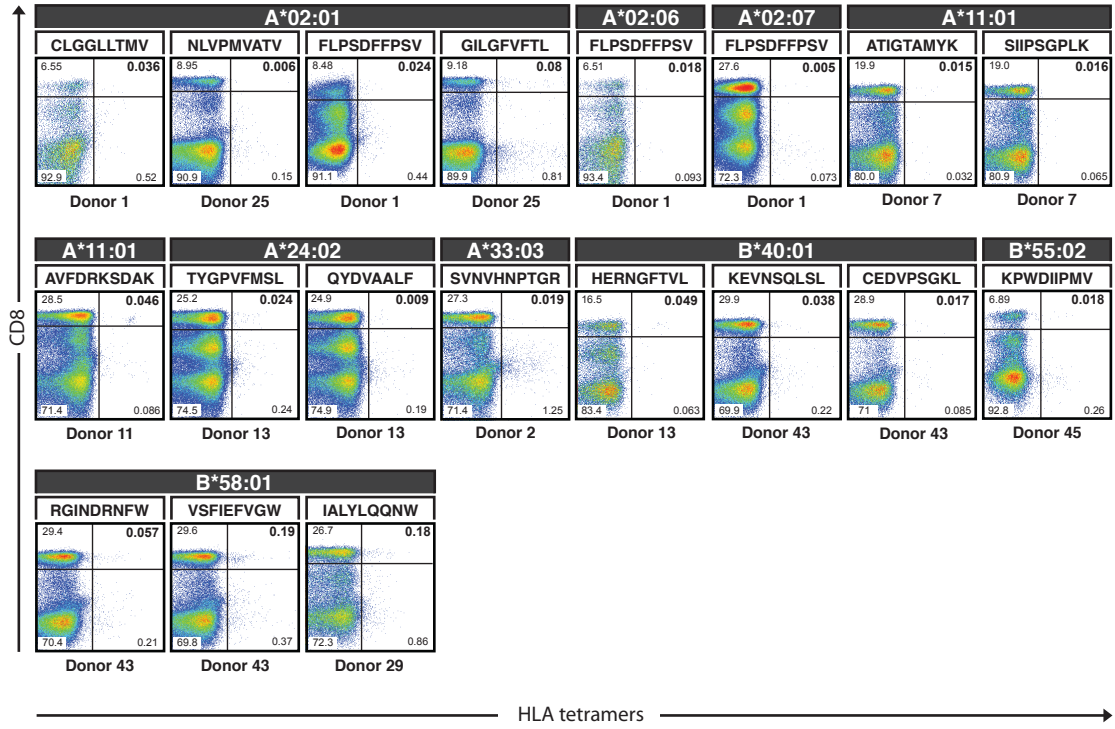
Donor	Epitope	HLA-A	HLA-A	HLA-B	HLA-B	HLA-C	HLA-C
1	CLGGLLTMV	A*02:01	A*02:06	B*40:01	B*27:05	C*03:04	C*03:03
2	SVNVHNPTGR TYGPVFMSL QYDPVAALF	A*33:03	A*24:02	B*58:01	B*40:01	C*07:02	C*03:02
3	-	A*11:01	A*11:01	B*13:06	B*48:03	C*03:04	C*08:06
4	-	A*02:03	A*33:03	B*15:02	B*35:03	C*04:01	C*08:01
5	-	A*24:77	A*26:01	B*07:02	B*39:01	C*07:02	C*07:02
6	-	A*02:01	A*02:06	B*56:01	B*95:21	C*07:02	C*08:01
7	ATIGTAMYK SIIPSGPLK	A*11:01	A*32:01	B*51:07	B*46:01	C*14:02	C*01:02
8	AVFDRKSDAK	A*33:03	A*11:01	B*15:02	B*51:01	C*08:01	C*14:02
9	-	A*02:03	A*11:01	B*40:01	B*39:01	C*04:01	C*07:02
10	RGINDRNFW IALYLQQNW	A*02:01	A*33:03	B*58:01	B*58:01	C*03:02	C*03:02
11	AVFDRKSDAK	A*11:01	A*33:03	B*58:01	B*13:01	C*05:01	C*01:02
12		A*11:01	A*11:01	B*44:02	B*27:05	C*03:02	C*03:04
13	TYGPVFMSL QYDPVAALF HERNGFTVL KEVNSQLSL	A*24:02	A*02:07	B*40:01	B*40:01	C*04:03	C*04:01
14	-	A*02:01	A*24:02	B*35:02	B*35:01	C*04:01	C*16:02
15	-	A*24:02	A*24:02	B*40:01	B*15:01	C*07:02	C*03:04
16	-	A*02:07	A*02:06	B*46:01	B*51:01	C*01:02	C*14:02
17	-	A*02:01	A*26:01	B*40:01	B*40:01	C*07:02	C*03:04
18	-	A*02:03	A*24:02	B*58:01	B*40:06	C*03:04	C*03:02
19	-	A*11:02	A*02:07	B*46:01	B*51:02	C*15:02	C*01:02
20	-	A*11:01	A*02:01	B*40:01	B*13:01	C*07:02	C*03:04
21	-	A*24:02	A*24:02	B*40:01	B*40:47	C*03:04	C*04:03
22	-	A*24:02	A*11:01	B*15:02	B*38:02	C*08:01	C*07:02
23	-	A*24:02	A*11:01	B*40:53	B*40:75	C*08:03	C*01:02
24	-	A*24:02	A*33:03	B*58:01	B*27:04	C*03:02	C*12:02
25	NLVPMVATV GILGFVFTL	A*02:01	A*33:03	B*40:06	B*35:03	C*04:01	C*01:02
26	-	A*02:03	A*24:02	B*40:01	B*51:01	C*04:03	C*15:02
27	-	A*02:01	A*33:03	B*58:01	B*13:01	C*03:04	C*03:02
28	-	A*33:03	A*33:03	B*58:01	B*58:01	C*03:02	C*03:02
29	TYGPVFMSL QYDPVAALF VSFIEFVGW IALYLQQNW	A*33:03	A*24:02	B*58:01	B*13:01	C*03:04	C*03:02
30	SIIPSGPLK	A*11:01	A*33:03	B*40:01	B*13:01	C*07:02	C*03:04
31	-	A*11:01	A*11:02	B*40:01	B*15:27	C*03:04	C*04:01
32	-	A*24:02	A*24:02	B*40:01	B*40:02	C*03:04	C*03:03
33	GILGFVFTL	A*02:01	A*02:01	B*46:01	B*40:01	C*15:02	C*01:02
34	-	A*02:03	A*02:03	B*38:02	B*13:01	C*03:04	C*07:02
35	-	A*02:06	A*02:01	B*15:02	B*51:01	C*08:01	C*14:02
36	-	A*11:01	A*31:01	B*54:01	B*51:01	C*14:02	C*01:02
37	-	A*02:06	A*02:06	B*55:02	B*15:02	C*08:01	C*01:02
38	TYGPVFMSL	A*24:02	A*02:03	B*46:01	B*15:01	C*01:02	C*01:02
39	-	A*31:01	A*11:01	B*51:01	B*51:01	C*14:02	C*14:02
40	-	A*11:02	A*02:01	B*54:01	B*27:04	C*01:02	C*12:02
41	QYDPVAALF	A*24:02	A*11:01	B*54:01	B*15:02	C*08:01	C*01:02
43	-	A*11:01	A*26:01	B*51:01	B*39:01	C*07:02	C*15:02

43	HERNGFTVL KEVNSQLSL CEDVPSGKL RGINDRNFW VSFIEFVGW IALYLQQNW	A*33:03	A*02:03	B*58:01	B*40:01	C*04:03	C*03:02
44	-	A*11:01	A*02:01	B*40:01	B*40:01	C*15:17	C*07:01
45	KPWDIIPMV	A*02:06	A*24:07	B*55:02	B*27:06	C*03:04	C*03:03

The HLA-A, -B and -C genotyping of healthy volunteers , together with the sequences of the HCMV, EBV, influenza, or dengue virus antigens that could elicit a response in the PBMCs are shown. HLA alleles for which we report conditional ligands for have been indicated in red.

Appendix B.6

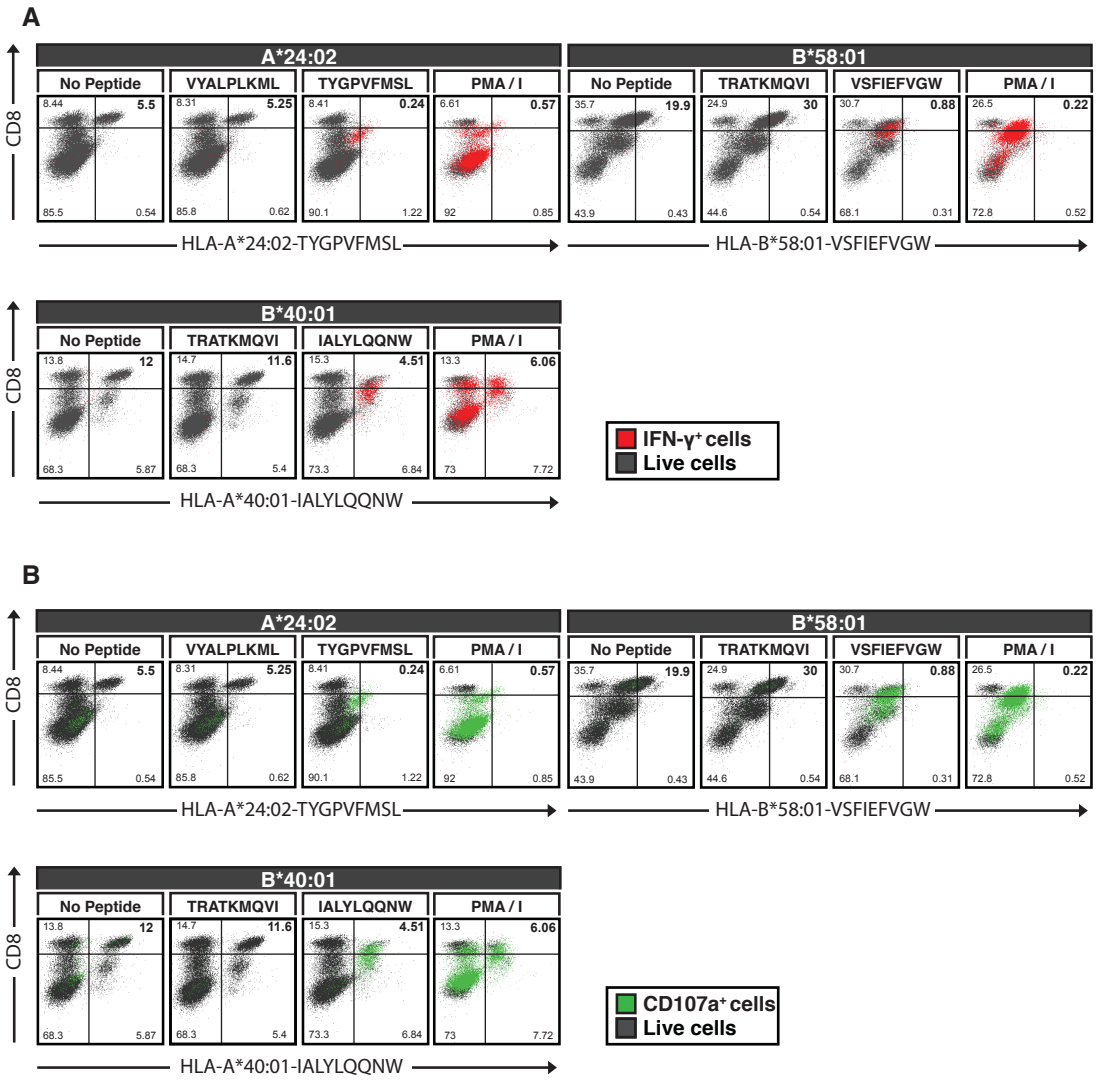
Direct *ex-vivo* HLA tetramer staining of PBMCs from various HLA-matched healthy volunteers.



PBMCs were isolated and stained with HLA tetramers, followed by flow cytometry analysis. Peptides originating from EBV, HCMV, DENV or Influenza A were used in the construction of tetramers for HLA-A*02:01, -A*02:06, -A*02:07, -A*11:01, -A*24:02, -A*33:03, -B*40:01, -B*55:02 or -B*58:01 as indicated. Cells were also stained for viability and co-stained with anti-human CD8a antibodies. The results shown here were gated on viable cells. For the generation of HLA-A*33:03 tetramers, conditional ligands with the photocleavable moiety at position 8 instead of position 4 was used. Numbers shown in the flow cytometry analysis are percentages of total live cells for each gated quadrant. Cells in the top-right quadrant represent HLA tetramer and CD8 double-positive cells. The data was obtained from independent experiments.

Appendix B.7

IFN- γ production and CD107 expression of antigen-specific T cells



Surface cell marker staining for CD107 (green), CD8a and intracellular cytokine staining for IFN-γ (red), together with HLA-tetramer and viability dye stainings were performed on antigen-specific cells. These cells were PBMCs of donor 29 expanded for 14 with their respective peptide. Numbers shown in the flow cytometry analysis are percentages of total live cells for each gated quadrant. Cells in the top-right quadrant represent HLA tetramer and CD8 double-positive cells. The data shown here were pre-gated for viable cells.

Appendix C

Appendix C.1

Peptide-binding motifs of HLAs

A*02:01

Position

1	2	3	4	5	6	7	8	9
L			E		V		K	**V**
M			K		L			L
			P		I			

A*02:06

Position

1	2	3	4	5	6	7	8	9
	V	I	E		L			V
	Q	L	P					L

A*02:07

Position

1	2	3	4	5	6	7	8	9
	L	**D**	E		I			**L**
		P						

A*11:01

Position

1	2	3	4	5	6	7	8	9
	V	M				L		**K**
	I	L				I		
	F	F				Y		
	Y	Y				V		
	T	I				F		
	L	A						
	M							
	S							
	A							
	G							

A*24:02

Position

1	2	3	4	5	6	7	8	9
	Y							**F**
	F							**W**
								I
								L

A*33:03

Position

1	2	3	4	5	6	7	8	9
	A							**R**
	I							
	L							
	F							
	Y							
	V							

B*15:02

Position								
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
	L	Y	P	F	D	V		**Y**
	V	R	E	H				**F**
	Q	F	D					M
	P	K						
		N						
		W						

B*40:01

Position								
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
	E					I		**L**
						V		

B*46:01

Position								
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
	M	K	D	P	S	E	V	**Y**
		R	E	I			A	**F**
		N	V					

B*55:02

Position								
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
	P	R					M	A
		M						
		Y						
		F						
		H						
		K						
		L						

B*58:01

Position								
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
	A	L	P	V			W	**W**
	T	N	E	M				**F**
	S	F	K	I				
		Y	Q	L				
			R	F				

C*03:04

Position								
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
	A	V	P		F	M		**L**
		I	E		Y	E		**M**
		P						
		Y						
		M						

C*04:01

Position								
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
	Y	D	D	A	V	A	K	L
	P		E		I			F
			P		L			

C*07:02

Position								
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
	Y	P	D	V	V			Y
	P	G	E	Y	I			
		A						

The peptide-binding motifs of HLA complexes for which we had designed conditional ligands are compiled as shown (255). The amino acid identity for the anchors, including dominant anchor residue (bold), and their position within the peptide starting from the N to C-terminus (1 to 9) are indicated. For HLAs, A*02:03, A*02:11 and C*08:01, the information is unavailable.

Appendix C.2

Possible HLA restriction of HBV and DENV 15-mer peptides based on peptide-binding motif

T cell line	Epitope	Predicted HLA restriction	Possible peptides	
HBV Pol 33	ESTRSASFCGSPYSW	B*58:01	SASFCGSPY SASFCGSPYS SASFCGSPYSW STRSASFCG STRSASFCGS STRSASFCGSP	ESTRSASFC ESTRSASFCG ESTRSASFCGS RSASFCGSP RSASFCGSPY RSASFCGSPYS
HBV Env 72	SVIWMMWYWGPSLYN	B*58:01	SVIWMMWYW SVIWMMWYWG SVIWMMWYWGP	MWYWGPSLY MWYWGPSLYN VIWMMWYWG VIWMMWYWGP
DENV NS3-108	LMRRGDLPVWLAYRV	B*55:02	LPVWLAYRV LMRRGDLPV LMRRGDLPVW LMRRGDLPVWL MRRGDLPVW MRRGDLPVWL MRRGDLPVWLA	GDLPVWLAY GDLPVWLAYR GDLPVWLAYRV RGDLPVWLA RRGDLPVWLA MRRGDLPVWLA
DENV NS3-108	LMRRGDLPVWLAYRV	B*27:06	RRGDLPVWL RRGDLPVWLA RRGDLPVWLAY LMRRGDLPVWL MRRGDLPVW MRRGDLPVWL MRRGDLPVWLA	GDLPVWLAY GDLPVWLAYR GDLPVWLAYRV LPVWLAYRV
DENV NS3-108	LMRRGDLPVWLAYRV	C*03:04	RRGDLPVWL MRRGDLPVWL LMRRGDLPVWL DLPVWLAYR DLPVWLAYRV	GDLPVWLAY GDLPVWLAYR GDLPVWLAYRV

The peptide-binding motifs of all possible HLAs of the T cell lines were compared to the 15-mer peptides. The possible dominant amino acid residue (red), and auxillary anchor residues (blue) of the 9, 10, 11-mer peptides imbedded within the 15-mer for the predicted HLAs are indicated (90, 255). Only peptides that contain dominant residues for anchoring to HLAs found in the donors are shown. Information for HLA-A*24:07, B*38:02, C*03:02, C*03:03 are not available and was obmitted in the analysis.

Appendix C.3

List of SARS-CoV, HBV or DENV peptides screened

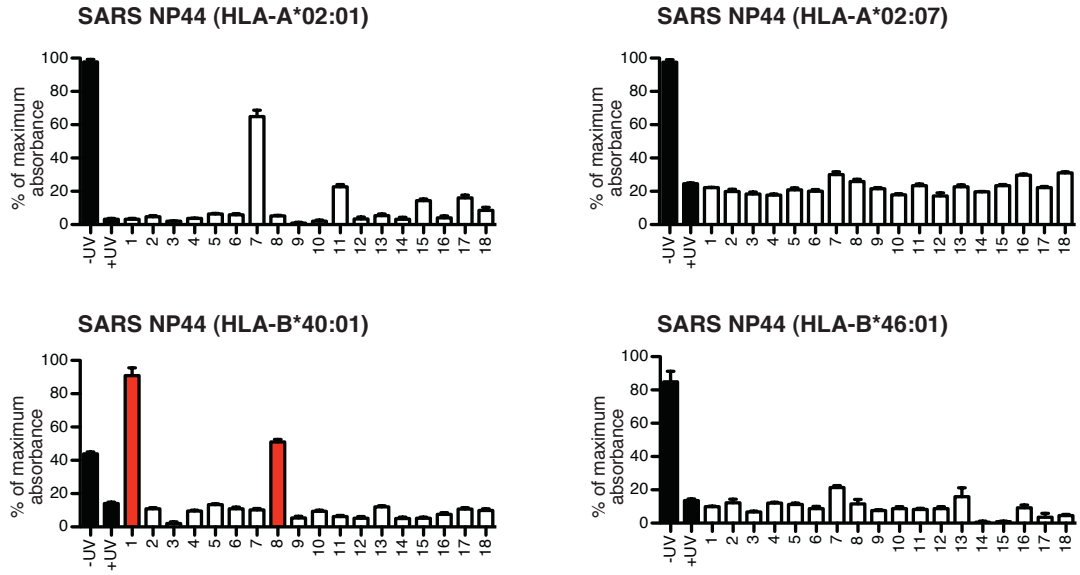
T cell line	9-mer		10-mer		11-mer	
	No.	Sequence	No.	Sequence	No.	Sequence
SARS NP44	1	GETALALLL	8	GETALALLLL	14	GETALALLLLD
	2	ETALALLLL	9	ETALALLLLD	15	ETALALLLLDR
	3	TALALLLLD	10	TALALLLLDR	16	TALALLLLDRL
	4	ALALLLLDR	11	ALALLLLDRL	17	ALALLLLDRLN
	5	LALLLLDRL	12	LALLLLDRLN	18	LALLLLDRLNQ
	6	ALLLLDRLN	13	ALLLLDRLNQ		
	7	LLLLDRLNQ				
SARS NP65	1	IGMEVTPSG	8	IGMEVTPSGT	14	IGMEVTPSGTW
	2	GMEVTPSGT	9	GMEVTPSGTW	15	GMEVTPSGTWL
	3	MEVTPSGTW	10	MEVTPSGTWL	16	MEVTPSGTWLT
	4	EVTPSGTWL	11	EVTPSGTWLT	17	EVTPSGTWLTY
	5	VTPSGTWLT	12	VTPSGTWLTY	18	VTPSGTWLTYH
	6	TPSGTWLTY	13	TPSGTWLTYH		
	7	PSGTWLTYH				
DENV-NS3-108	1	LMRRGDLPV	8	LMRRGDLPVW	14	LMRRGDLPVWL
	2	MRRGDLPVW	9	MRRGDLPVWL	15	MRRGDLPVWLA
	3	RRGDLPVWL	10	RRGDLPVWLA	16	RRGDLPVWLAY
	4	RGDLPVWLA	11	RGDLPVWLAY	17	RGDLPVWLAYR
	5	GDLPVWLAY	12	GDLPVWLAYR	18	GDLPVWLAYRV
	6	DLPVWLAYR	13	DLPVWLAYRV		
	7	LPVWLAYRV				
SARS 3a2	1	RFFTLGSIT	8	RFFTLGSITA	14	RFFTLGSITAQ
	2	FFTLGSITA	9	FFTLGSITAQ	15	FFTLGSITAQP
	3	FTLGSITAQ	10	FTLGSITAQP	16	FTLGSITAQPV
	4	TLGSITAQP	11	TLGSITAQPV	17	TLGSITAQPVK
	5	LGSITAQPV	12	LGSITAQPVK	18	LGSITAQPVKI
	6	GSITAQPVK	13	GSITAQPVKI		
	7	SITAQPVKI				
HBV Pol-33	1	ESTRSASFC	8	ESTRSASFCG	14	ESTRSASFCGS
	2	STRSASFCG	9	STRSASFCGS	15	STRSASFCGSP
	3	TRSASFCGS	10	TRSASFCGSP	16	TRSASFCGSPY
	4	RSASFCGSP	11	RSASFCGSPY	17	RSASFCGSPYS
	5	SASFCGSPY	12	SASFCGSPYS	18	SASFCGSPYSW
	6	ASFCGSPYS	13	ASFCGSPYSW		
	7	SFCGSPYSW				
HBV Env72	1	SVIWMMWYW	8	SVIWMMWYWG	14	SVIWMMWYWGP
	2	VIWMMWYWG	9	VIWMMWYWGP	15	VIWMMWYWGPS
	3	IWMMWYWGP	10	IWMMWYWGPS	16	IWMMWYWGPSL
	4	WMMWYWGPS	11	WMMWYWGPSL	17	WMMWYWGPSLY
	5	MMWYWGPSL	12	MMWYWGPSLY	18	MMWYWGPSLYN
	6	MWYWGPSLY	13	MWYWGPSLYN		
	7	WYWGPSLYN				
HBV Env34	1	STTSGFLGP	8	STTSGFLGPL	14	STTSGFLGPLL
	2	TTSGFLGPL	9	TTSGFLGPLL	15	TTSGFLGPLLV
	3	TSGFLGPLL	10	TSGFLGPLLV	16	TSGFLGPLLVL
	4	SGFLGPLLV	11	SGFLGPLLVL	17	SGFLGPLLVLQ
	5	GFLGPLLVL	12	GFLGPLLVLQ	18	GFLGPLLVLQA
	6	FLGPLLVLQ	13	FLGPLLVLQA		
	7	LGPLLVLQA				

Seven 15-mer SARS-CoV, HBV, or DENV antigens capable of stimulating the respective T cells were discovered by ELISPOT screening. Sequences of all possible eighteen 9-, 10-, and 11-mer truncated peptides imbedded in the 15-mer epitope are indicated. This system of numbering peptide fragments is used in **Appendix C.4**.

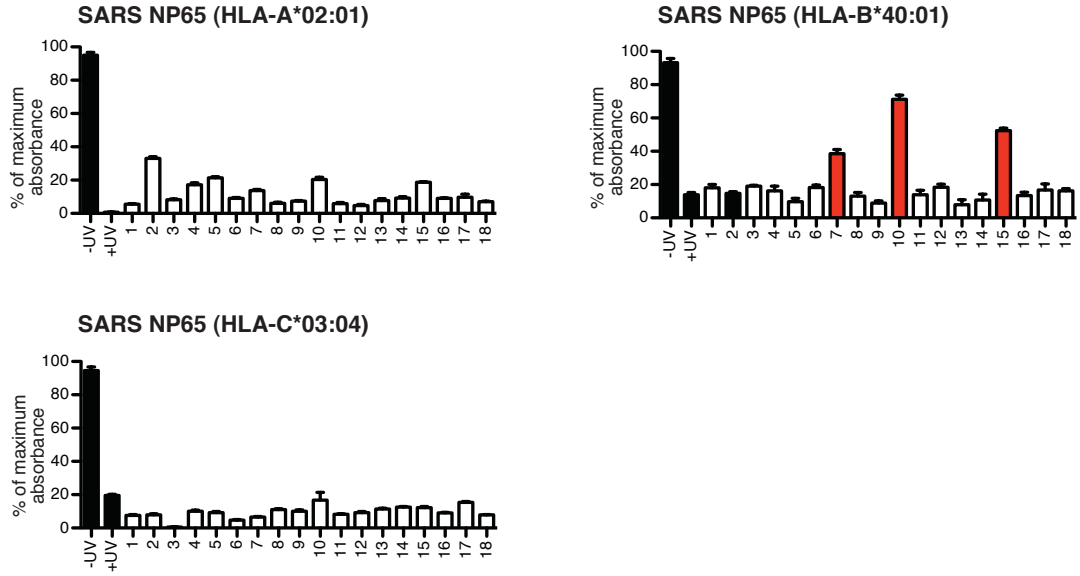
Appendix C.4

HLA-stability ELISA screening

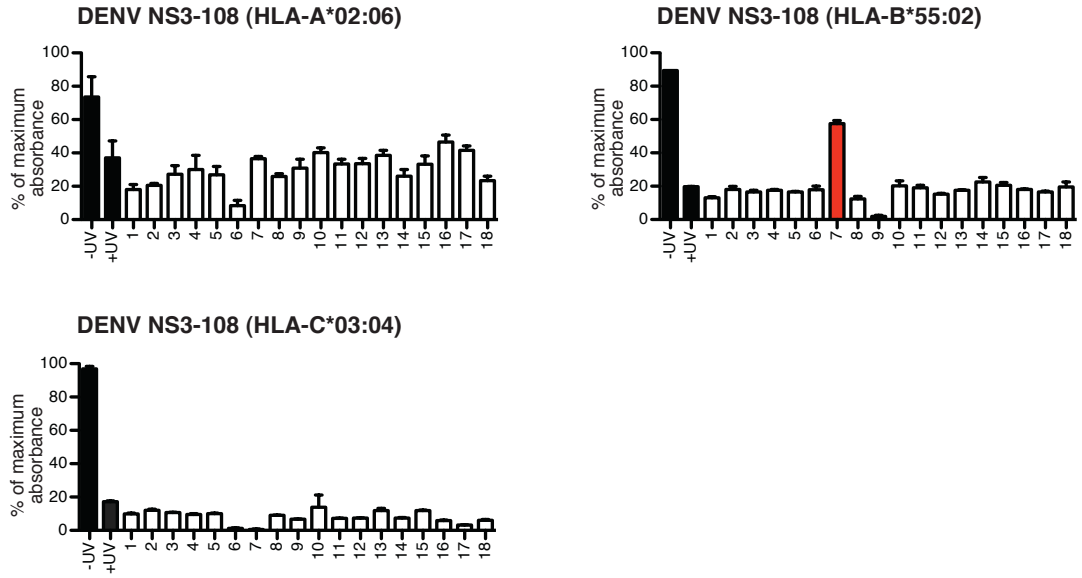
A.



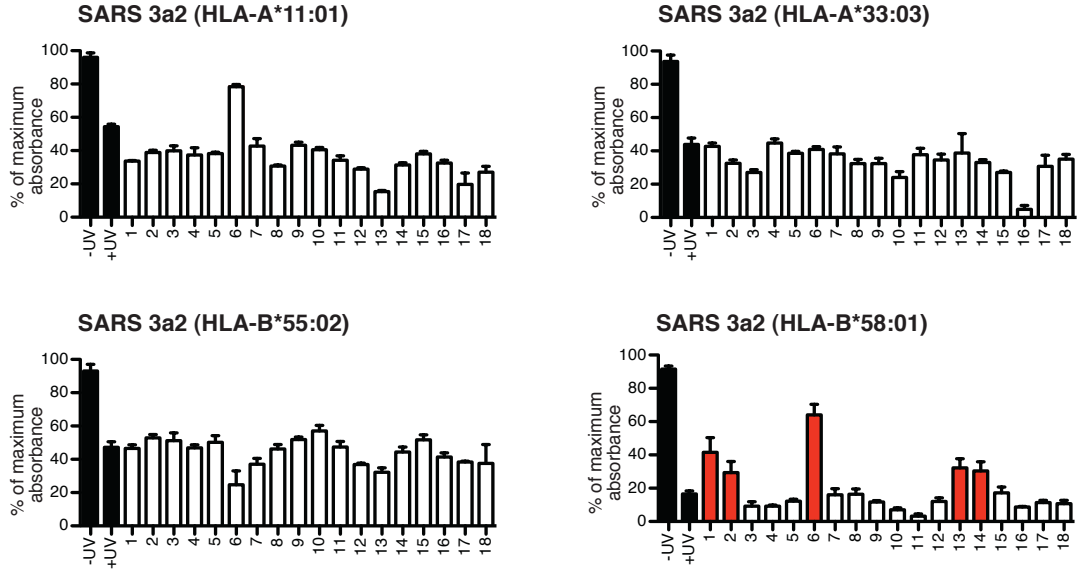
B.



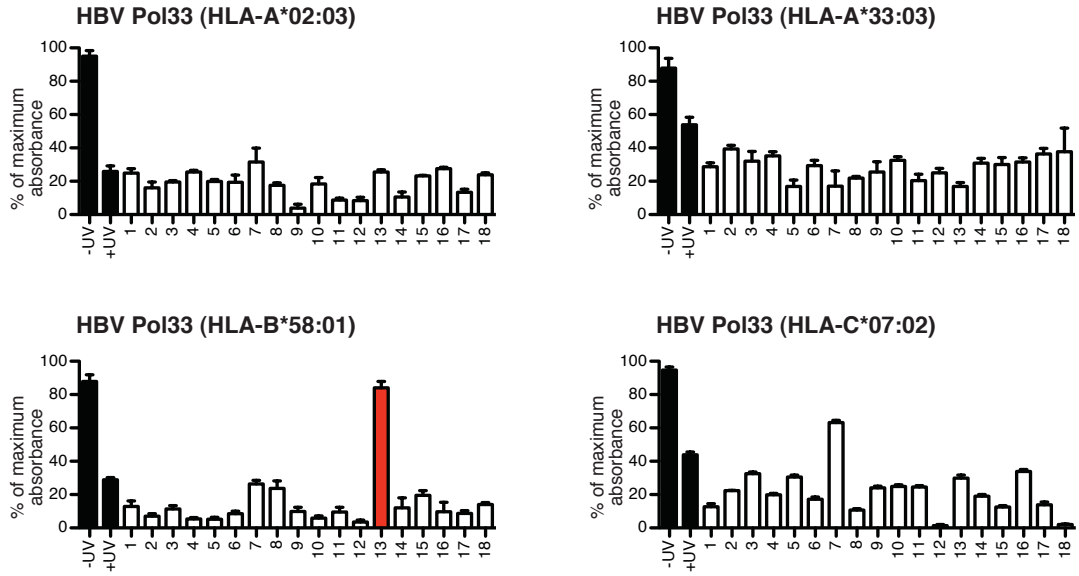
C.


DENV NS3-108 (HLA-A*02:06)
DENV NS3-108 (HLA-B*55:02)
DENV NS3-108 (HLA-C*03:04)
% of maximum absorbance
0
20
40
60
80
100
-UV
+UV
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

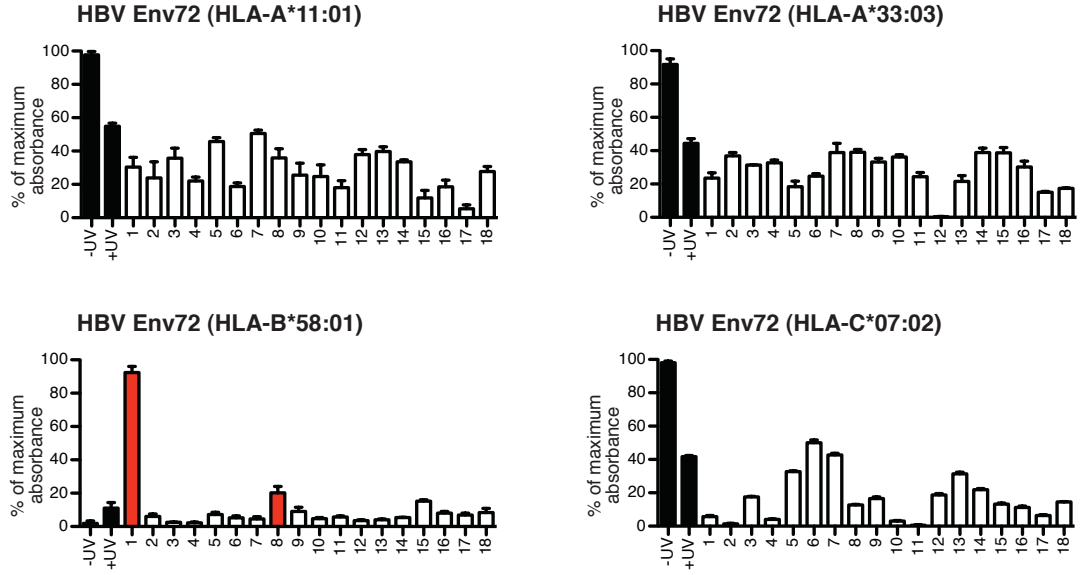
D.


SARS 3a2 (HLA-A*11:01)
SARS 3a2 (HLA-A*33:03)
SARS 3a2 (HLA-B*55:02)
SARS 3a2 (HLA-B*58:01)
% of maximum absorbance
0
20
40
60
80
100
-UV
+UV
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

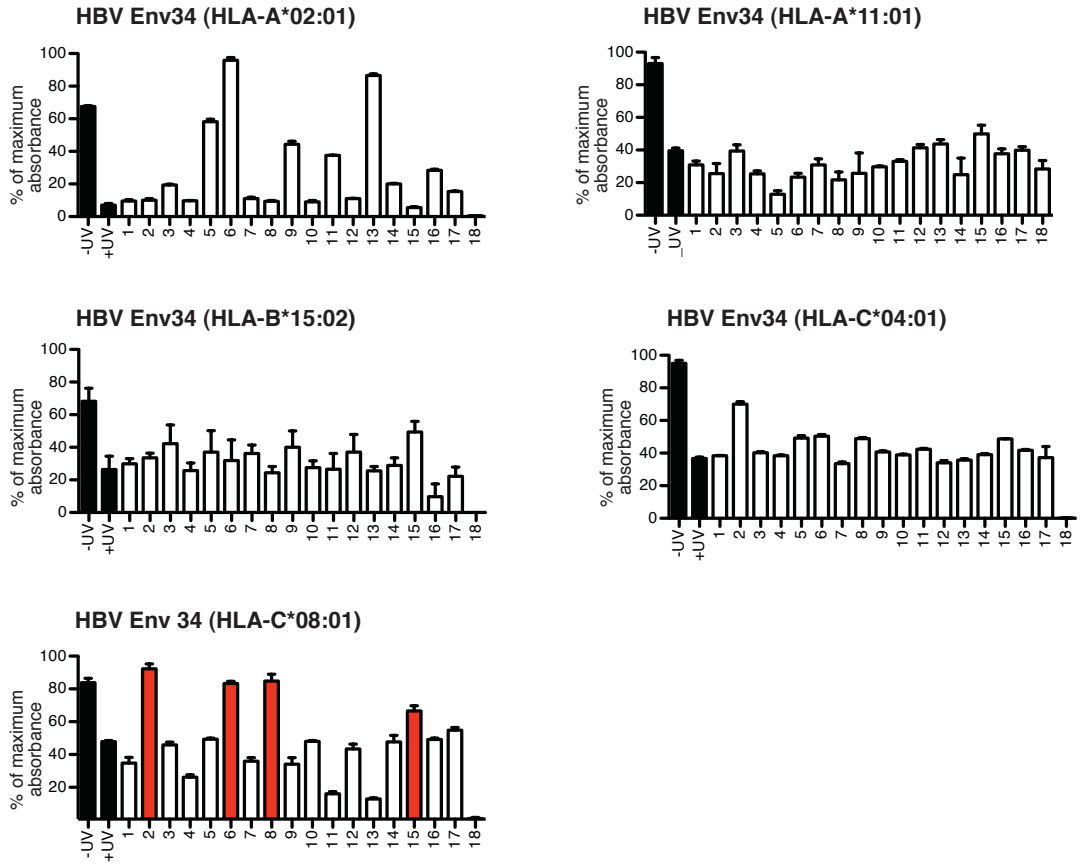
E.


HBV Pol33 (HLA-A*02:03)
HBV Pol33 (HLA-A*33:03)
HBV Pol33 (HLA-B*58:01)
HBV Pol33 (HLA-C*07:02)
% of maximum absorbance
0
20
40
60
80
100
-UV
+UV
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

F.


HBV Env72 (HLA-A*11:01)
HBV Env72 (HLA-A*33:03)
HBV Env72 (HLA-B*58:01)
HBV Env72 (HLA-C*07:02)
% of maximum absorbance
0
20
40
60
80
100
-UV
+UV
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

G.

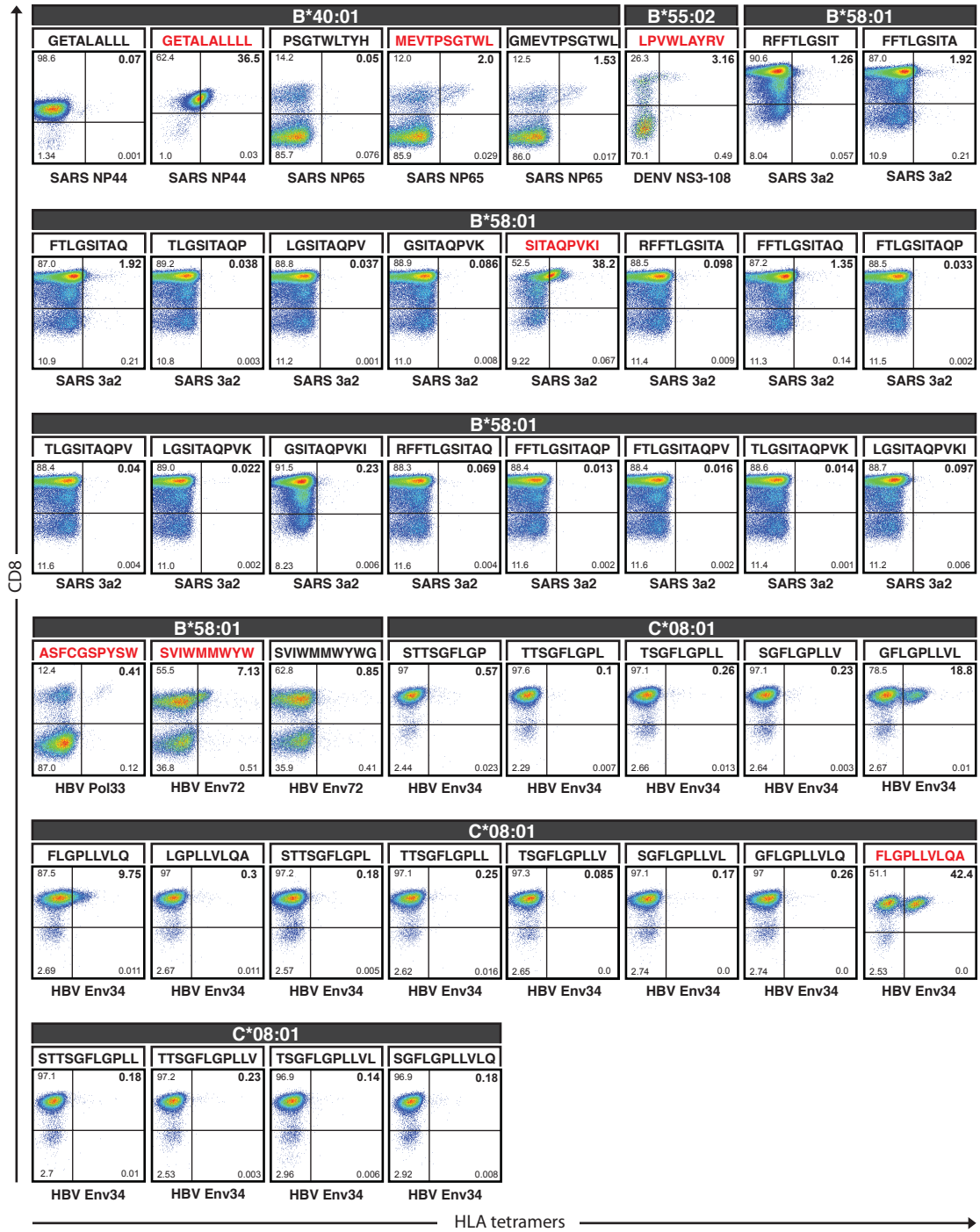


Soluble HLA molecules that correspond to the PBMC donors HLA constellation were ligand-exchanged with truncated peptides from (**A**) SARS NP44, (**B**) SARS NP65, (**C**) DENV NS3-108, (**D**) SARS 3a2, (**E**) HBV Pol 33, (**F**) HBV Env72 and (**G**) HBV Env34. The peptides screened are 9-, 10-, and 11-mer peptides imbedded in the 15-mer epitopes and their exact sequences can be found in **Appendix C.3**. The resulting HLA complexes were captured on streptavidin-coated plates and probed for β2m as marker for HLA complex stability. As controls, the photocleavable HLA molecules were treated with (+) or without (-) UV irradiation in the absence of rescue peptide (black bars). For the correct HLA restriction of the identified epitope (**Table 4.2**), peptides are labeled as capable (red bars) or incapable (white bars) of stabilizing the corresponding HLA.

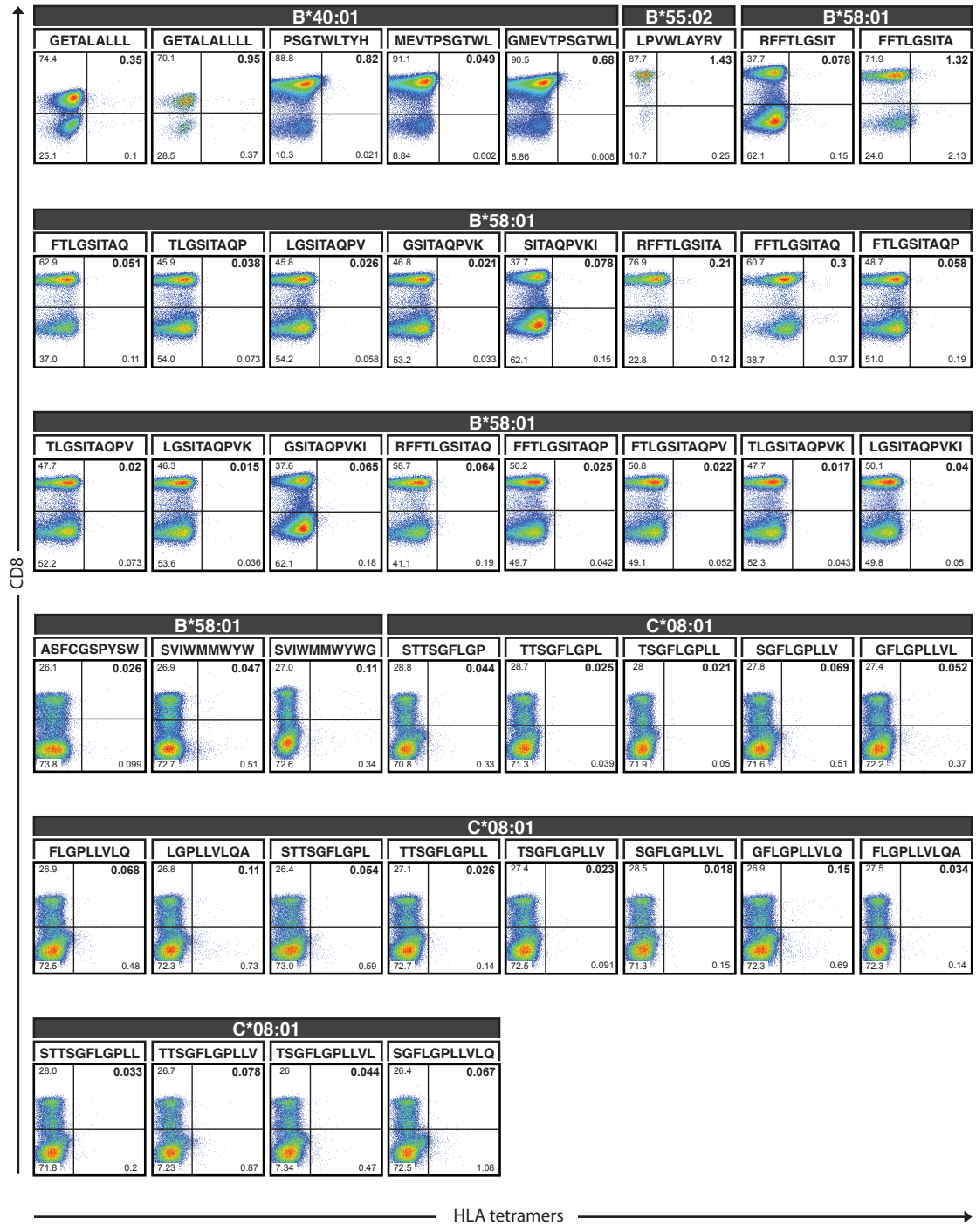
Appendix C.5

HLA tetramers screening for novel epitope

A



B



HLA Tetramer staining of (**A**) SARS-CoV, HBV, or DENV-specific T cell lines or (**B**) PBMCs from uninfected volunteers. The T cell lines were cultivated from PBMCs from Donor no.45 (DENV NS3-108), no.46 (SARS NP44), no.47 (SARS NP65), no.48 (SARS 3a2), no.49 (HBV Pol33), no.50 (HBV Env72), no.51 (HBV Env34). The cells were stained with HLA-B*40:01, -B*55:02, -B*58:01 or -C*08:01 tetramers loaded with the indicated peptide to identify novel epitopes (red). HLA-C*08:01-GFLGPLLVLQA tetramer stainings are omitted due to high background. Numbers shown in the flow cytometry analysis are percentages of cells for each gated quadrant. The data were from independent experiments.

Appendix C.6

Variations in Dengue $NS3_{542-550}$ from different strains

NS3 Epitope	Strain	Country, Region	Year	Dengue Serotype	Source Accession
LPVWLAYRV	TSV01	Australia, Townsville	1993	2	AAK67712
	BR64022	Brazil	1998	2	AAL96681
LPVWLAYKV	ThNH-p11/93	Thailand, Nakhon Phanom	1993	2	AAC40835
	Puerto Rico/PR159-S1/1969	USA, Puerto Rico	1969	2	P12823
	DENV-3/CO/BID-V2988/2007	Colombia, Santander	2007	3	ACQ44484
LPVWLSYKV	HawO3758	USA, Hawaii	2001	1	ABG75765
	DENV-1/VN/BID-V3906/2008	Vietnam	2008	1	ADK26437
	Brazil/97-11/1997	Brazil	1997	1	P27909
	Singapore/S275/1990	Singapore	1990	1	P33478
	Thailand/0476/1997	Thailand	1997	4	Q2YHF2
	Singapore/8976/1995	Singapore	1995	4	Q5UCB8
LPVWLAHKV	TB16	Indonesia, Jakarta	2004	3	AAW51418
	DENV-3/NI/BID-V4796/2009	Nicaragua, Managua	2009	3	ADK26436
	InJ-16-82	Indonesia	1982	3	ABD65874
	DENV-3/US/BID-V1732/2002	USA, Puerto Rico	2002	3	FJ390373
	Singapore/8120/1995	Singapore	1995	3	Q5UB51
	China/80-2/1980	China	1980	3	Q99D35

4 common dengue variations of the $NS3_{542-550}$ epitope were found using protein blast and tblastn provided by NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The epitope sequence with the variable region (red), and dengue strain's name, country and region of origin, year of data collection, predicted dengue serotype and source accession are given.

B

Bibliography

1. Delves PJ, Roitt IM. The immune system. Second of two parts. *N Engl J Med*. 2000;343(2):108-117.
2. Williams MA, Bevan MJ. Effector and memory CTL differentiation. *Annu Rev Immunol*. 2007;25:171-192.
3. Smith-Garvin JE, Koretzky GA, Jordan MS. T cell activation. *Annu Rev Immunol*. 2009;27:591-619.
4. Kaech SM, Wherry EJ, Ahmed R. Effector and memory T-cell differentiation: implications for vaccine development. *Nat Rev Immunol*. 2002;2(4):251-262.
5. Arstila TP, Casrouge A, Baron V, et al. A direct estimate of the human alphabeta T cell receptor diversity. *Science*. 1999;286(5441):958-961.
6. Casrouge A, Beaudoing E, Dalle S, et al. Size estimate of the alpha beta TCR repertoire of naive mouse splenocytes. *J Immunol*. 2000;164(11):5782-5787.
7. Nikolich-Zugich J, Slifka MK, Messaoudi I. The many important facets of T-cell repertoire diversity. *Nat Rev Immunol*. 2004;4(2):123-132.
8. Turner SJ, Doherty PC, McCluskey J, Rossjohn J. Structural determinants of T-cell receptor bias in immunity. *Nat Rev Immunol*. 2006;6(12):883-894.
9. Rudolph MG, Stanfield RL, Wilson IA. How TCRs bind MHCs, peptides, and coreceptors. *Annu Rev Immunol*. 2006;24:419-466.
10. Venturi V, Price DA, Douek DC, Davenport MP. The molecular basis for public T-cell responses? *Nat Rev Immunol*. 2008;8(3):231-238.
11. Sidney J, Peters B, Frahm N, Brander C, Sette A. HLA class I supertypes: a revised and updated classification. *BMC Immunol*. 2008;9:1.
12. Vyas JM, Van der Veen AG, Ploegh HL. The known unknowns of antigen processing and presentation. *Nat Rev Immunol*. 2008;8(8):607-618.
13. Shastri N, Schwab S, Serwold T. Producing nature's gene-chips: the generation of peptides for display by MHC class I molecules. *Annu Rev Immunol*. 2002;20:463-493.
14. Trombetta ES, Mellman I. Cell biology of antigen processing in vitro and in vivo. *Annu Rev Immunol*. 2005;23:975-1028.
15. Jensen PE. Recent advances in antigen processing and presentation. *Nat Immunol*. 2007;8(10):1041-1048.
16. Shiina T, Hosomichi K, Inoko H, Kulski JK. The HLA genomic loci map: expression, interaction, diversity and disease. *J Hum Genet*. 2009;54(1):15-39.
17. Horton R, Wilming L, Rand V, et al. Gene map of the extended human MHC. *Nat Rev Genet*. 2004;5(12):889-899.

18. Zinkernagel RM, Doherty PC. Restriction of in vitro T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. *Nature*. 1974;248(450):701-702.
19. Hraber P, Kuiken C, Yusim K. Evidence for human leukocyte antigen heterozygote advantage against hepatitis C virus infection. *Hepatology*. 2007;46(6):1713-1721.
20. Thursz MR, Thomas HC, Greenwood BM, Hill AV. Heterozygote advantage for HLA class-II type in hepatitis B virus infection. *Nat Genet*. 1997;17(1):11-12.
21. Carrington M, Nelson GW, Martin MP, et al. HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. *Science*. 1999;283(5408):1748-1752.
22. Robinson J, Mistry K, McWilliam H, et al. The IMGT/HLA database. *Nucleic Acids Res*. 2011;39(Database issue):D1171-1176.
23. Solberg OD, Mack SJ, Lancaster AK, et al. Balancing selection and heterogeneity across the classical human leukocyte antigen loci: a meta-analytic review of 497 population studies. *Hum Immunol*. 2008;69(7):443-464.
24. National Center for Biotechnology Information. dbMHC Anthropology Resources Alleles and Haplotype Frequencies Website. http://www.ncbi.nlm.nih.gov/gv/mhc/ihwg.cgi?cmd=page&page=AnthroMain. Updated 2007. Accessed June, 2011.
25. Stephens HA. HLA and other gene associations with dengue disease severity. *Curr Top Microbiol Immunol*. 2010;338:99-114.
26. Carrington M, O'Brien SJ. The influence of HLA genotype on AIDS. *Annu Rev Med*. 2003;54:535-551.
27. Ghodke Y, Joshi K, Chopra A, Patwardhan B. HLA and disease. *Eur J Epidemiol*. 2005;20(6):475-488.
28. Tan AT, Loggi E, Boni C, et al. Host ethnicity and virus genotype shape the hepatitis B virus-specific T-cell repertoire. *J Virol*. 2008;82(22):10986-10997.
29. Martin M, Carrington M. Immunogenetics of viral infections. *Curr Opin Immunol*. 2005;17(5):510-516.
30. Sidney J, del Guercio M, Southwood S, et al. Several HLA alleles share overlapping peptide specificities. *J Immunol*. 1995;154(1):247-259.
31. Sidney J, Grey H, Kubo R, Sette A. Practical, biochemical and evolutionary implications of the discovery of HLA class I supermotifs. *Immunol Today*. 1996;17(6):261-266.
32. Fremont DH, Matsumura M, Stura EA, Peterson PA, Wilson IA. Crystal structures of two viral peptides in complex with murine MHC class I H-2Kb. *Science*. 1992;257(5072):919-927.
33. Guo HC, Jardetzky TS, Garrett TP, et al. Different length peptides bind to HLA-Aw68 similarly at their ends but bulge out in the middle. *Nature*. 1992;360(6402):364-366.
34. Madden DR, Gorga JC, Strominger JL, Wiley DC. The three-dimensional structure of HLA-B27 at 2.1 A resolution suggests a general mechanism for tight peptide binding to MHC. *Cell*. 1992;70(6):1035-1048.
35. Matsumura M, Fremont DH, Peterson PA, Wilson IA. Emerging principles for the recognition of peptide antigens by MHC class I molecules. *Science*. 1992;257(5072):927-934.

36. Saper MA, Bjorkman PJ, Wiley DC. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 A resolution. *J Mol Biol*. 1991;219(2):277-319.
37. Sette A, Sidney J. Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. *Immunogenetics*. 1999;50(3-4):201-212.
38. Yewdell JW, Hickman HD. New lane in the information highway: alternative reading frame peptides elicit T cells with potent antiretrovirus activity. *J Exp Med*. 2007;204(11):2501-2504.
39. Ho O, Green WR. Alternative translational products and cryptic T cell epitopes: expecting the unexpected. *J Immunol*. 2006;177(12):8283-8289.
40. Ireland J, Herzog J, Unanue ER. Cutting edge: unique T cells that recognize citrullinated peptides are a feature of protein immunization. *J Immunol*. 2006;177(3):1421-1425.
41. Shastri N, Cardinaud S, Schwab SR, Serwold T, Kunisawa J. All the peptides that fit: the beginning, the middle, and the end of the MHC class I antigen-processing pathway. *Immunol Rev*. 2005;207:31-41.
42. Robbins PF, El-Gamil M, Li YF, et al. The intronic region of an incompletely spliced gp100 gene transcript encodes an epitope recognized by melanoma-reactive tumor-infiltrating lymphocytes. *J Immunol*. 1997;159(1):303-308.
43. Coulie PG, Lehmann F, Lethé B, et al. A mutated intron sequence codes for an antigenic peptide recognized by cytolytic T lymphocytes on a human melanoma. *Proc Natl Acad Sci USA*. 1995;92(17):7976-7980.
44. Uenaka A, Hirano Y, Hata H, et al. Cryptic CTL epitope on a murine sarcoma Meth A generated by exon extension as a novel mechanism. *J Immunol*. 2003;170(9):4862-4868.
45. Bullock TN, Eisenlohr LC. Ribosomal scanning past the primary initiation codon as a mechanism for expression of CTL epitopes encoded in alternative reading frames. *J Exp Med*. 1996;184(4):1319-1329.
46. Bullock TN, Patterson AE, Franlin LL, Notidis E, Eisenlohr LC. Initiation codon scanthrough versus termination codon readthrough demonstrates strong potential for major histocompatibility complex class I-restricted cryptic epitope expression. *J Exp Med*. 1997;186(7):1051-1058.
47. Malarkannan S, Horng T, Shih PP, Schwab S, Shastri N. Presentation of out-of-frame peptide/MHC class I complexes by a novel translation initiation mechanism. *Immunity*. 1999;10(6):681-690.
48. Schwab SR, Li KC, Kang C, Shastri N. Constitutive display of cryptic translation products by MHC class I molecules. *Science*. 2003;301(5638):1367-1371.
49. Shastri N, Nguyen V, Gonzalez F. Major histocompatibility class I molecules can present cryptic translation products to T-cells. *J Biol Chem*. 1995;270(3):1088-1091.
50. Uenaka A, Ono T, Akisawa T, et al. Identification of a unique antigen peptide pRL1 on BALB/c RL male 1 leukemia recognized by cytotoxic T lymphocytes and its relation to the Akt oncogene. *J Exp Med*. 1994;180(5):1599-1607.
51. Saeterdal I, Bjørheim J, Lislerud K, et al. Frameshift-mutation-derived peptides as tumor-specific antigens in inherited and spontaneous colorectal cancer. *Proc Natl Acad Sci USA*. 2001;98(23):13255-13260.

52. Saulquin X, Scotet E, Trautmann L, et al. +1 Frameshifting as a novel mechanism to generate a cryptic cytotoxic T lymphocyte epitope derived from human interleukin 10. *J Exp Med*. 2002;195(3):353-358.
53. Berger CT, Carlson JM, Brumme CJ, et al. Viral adaptation to immune selection pressure by HLA class I-restricted CTL responses targeting epitopes in HIV frameshift sequences. *J Exp Med*. 2010;207(1):61-75.
54. Cardinaud S, Moris A, Février M, et al. Identification of cryptic MHC I-restricted epitopes encoded by HIV-1 alternative reading frames. *J Exp Med*. 2004;199(8):1053-1063.
55. Walsh CT, Garneau-Tsodikova S, Gatto GJ. Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew Chem Int Ed Engl*. 2005;44(45):7342-7372.
56. Anderton SM. Post-translational modifications of self antigens: implications for autoimmunity. *Curr Opin Immunol*. 2004;16(6):753-758.
57. Engelhard VH, Altrich-Vanlith M, Ostankovitch M, Zarling AL. Post-translational modifications of naturally processed MHC-binding epitopes. *Curr Opin Immunol*. 2006;18(1):92-97.
58. Petersen J, Purcell AW, Rossjohn J. Post-translationally modified T cell epitopes: immune recognition and immunotherapy. *J Mol Med*. 2009;87(11):1045-1051.
59. Yagüe J, Alvarez I, Rognan D, et al. An N-acetylated natural ligand of human histocompatibility leukocyte antigen (HLA)-B39. Classical major histocompatibility complex class I proteins bind peptides with a blocked NH(2) terminus in vivo. *J Exp Med*. 2000;191(12):2083-2092.
60. Meadows L, Wang W, den Haan JM, et al. The HLA-A*0201-restricted H-Y antigen contains a posttranslationally modified cysteine that significantly affects T cell recognition. *Immunity*. 1997;6(3):273-281.
61. Van den Steen PE, Proost P, Brand DD, et al. Generation of glycosylated remnant epitopes from human collagen type II by gelatinase B. *Biochemistry*. 2004;43(33):10809-10816.
62. Yagüe J, Vázquez J, López de Castro JA. A post-translational modification of nuclear proteins, N(G),N(G)-dimethyl-Arg, found in a natural HLA class I peptide ligand. *Protein Sci*. 2000;9(11):2210-2217.
63. Herzog J, Maekawa Y, Cirrito TP, Illian BS, Unanue ER. Activated antigen-presenting cells select and present chemically modified peptides recognized by unique CD4 T cells. *Proc Natl Acad Sci USA*. 2005;102(22):7928-7933.
64. van Stipdonk MJ, Willems AA, Amor S, et al. T cells discriminate between differentially phosphorylated forms of alphaB-crystallin, a major central nervous system myelin antigen. *Int Immunol*. 1998;10(7):943-950.
65. McAdam SN, Fleckenstein B, Rasmussen IB, et al. T cell recognition of the dominant I-A(k)-restricted hen egg lysozyme epitope: critical role for asparagine deamidation. *J Exp Med*. 2001;193(11):1239-1246.
66. Skipper JC, Hendrickson RC, Gulden PH, et al. An HLA-A2-restricted tyrosinase antigen on melanoma cells results from posttranslational modification and suggests a novel pathway for processing of membrane proteins. *J Exp Med*. 1996;183(2):527-534.
67. Cao L, Goodin R, Wood D, Moscarello MA, Whitaker JN. Rapid release and unusual stability of immunodominant peptide 45-89 from citrullinated myelin basic protein. *Biochemistry*. 1999;38(19):6157-6163.

68. Chen W, Ede NJ, Jackson DC, Mccluskey J, Purcell AW. CTL recognition of an altered peptide associated with asparagine bond rearrangement. Implications for immunity and vaccine design. *J Immunol*. 1996;157(3):1000-1005.
69. Mannering SI, Harrison LC, Williamson NA, et al. The insulin A-chain epitope recognized by human T cells is posttranslationally modified. *J Exp Med*. 2005;202(9):1191-1197.
70. Hill JA, Southwood S, Sette A, et al. Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis-associated HLA-DRB1*0401 MHC class II molecule. *J Immunol*. 2003;171(2):538-541.
71. Arentz-Hansen H, Körner R, Molberg O, et al. The intestinal T cell response to alpha-gliadin in adult celiac disease is focused on a single deamidated glutamine targeted by tissue transglutaminase. *J Exp Med*. 2000;191(4):603-612.
72. Moss CX, Matthews SP, Lamont DJ, Watts C. Asparagine deamidation perturbs antigen presentation on class II major histocompatibility complex molecules. *J Biol Chem*. 2005;280(18):18498-18503.
73. Berkers CR, de Jong A, Ovaa H, Rodenko B. Transpeptidation and reverse proteolysis and their consequences for immunity. *Int J Biochem Cell Biol*. 2009;41(1):66-71.
74. Paulus H. Protein splicing and related forms of protein autoprocessing. *Annu Rev Biochem*. 2000;69:447-496.
75. Hanada K-I, Yewdell JW, Yang JC. Immune recognition of a human renal cancer antigen through post-translational protein splicing. *Nature*. 2004;427(6971):252-256.
76. Vigneron N, Stroobant V, Chapiro J, et al. An antigenic peptide produced by peptide splicing in the proteasome. *Science*. 2004;304(5670):587-590.
77. Warren EH, Vigneron NJ, Gavin MA, et al. An antigen produced by splicing of noncontiguous peptides in the reverse order. *Science*. 2006;313(5792):1444-1447.
78. Fleischmann RD, Adams MD, White O, et al. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science*. 1995;269(5223):496-512.
79. Fuller CW, Middendorf LR, Benner SA, et al. The challenges of sequencing by synthesis. *Nat Biotechnol*. 2009;27(11):1013-1023.
80. Shendure J, Ji H. Next-generation DNA sequencing. *Nat Biotechnol*. 2008;26(10):1135-1145.
81. Okeke IN, Wain J. Post-genomic challenges for collaborative research in infectious diseases. *Nat Rev Microbiol*. 2008;6(11):858-864.
82. Wren BW. Microbial genome analysis: insights into virulence, host adaptation and evolution. *Nat Rev Genet*. 2000;1(1):30-39.
83. Zhang R, Zhang C-T. The impact of comparative genomics on infectious disease research. *Microbes Infect*. 2006;8(6):1613-1622.
84. Bashyam HS, Green S, Rothman AL. Dengue virus-reactive CD8+ T cells display quantitative and qualitative differences in their response to variant epitopes of heterologous viral serotypes. *J Immunol*. 2006;176(5):2817-2824.
85. Imrie A, Meeks J, Gurary A, et al. Differential functional avidity of dengue virus-specific T-cell clones for variant peptides representing heterologous and previously encountered serotypes. *J Virol*. 2007;81(18):10081-10091.

86. Boon AC, de Mutsert G, van Baarle D, et al. Recognition of homo- and heterosubtypic variants of influenza A viruses by human CD8+ T lymphocytes. *J Immunol*. 2004;172(4):2453-2460.
87. Midgley RS, Bell AI, Yao QY, et al. HLA-A11-restricted epitope polymorphism among Epstein-Barr virus strains in the highly HLA-A11-positive Chinese population: incidence and immunogenicity of variant epitope sequences. *J Virol*. 2003;77(21):11507-11516.
88. Salimi N, Fleri W, Peters B, Sette A. Design and utilization of epitope-based databases and predictive tools. *Immunogenetics*. 2010;62(4):185-196.
89. Rammensee HG, Friede T, Stevanoviíc S. MHC ligands and peptide motifs: first listing. *Immunogenetics*. 1995;41(4):178-228.
90. Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanovic S. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics*. 1999;50(3-4):213-219.
91. Yang IS, Lee J-Y, Lee JS, et al. Influenza sequence and epitope database. *Nucleic Acids Res*. 2009;37(Database issue):D423-430.
92. Kuiken C, Korber B, Shafer RW. HIV sequence databases. *AIDS Rev*. 2003;5(1):52-61.
93. Tong JC, Kong L, Tan TW, Ranganathan S. MPID-T: database for sequence-structure-function information on T-cell receptor/peptide/MHC interactions. *Appl Bioinformatics*. 2006;5(2):111-114.
94. Davies V, Vaughan K, Damle R, Peters B, Sette A. Classification of the universe of immune epitope literature: representation and knowledge gaps. *PLoS ONE*. 2009;4(9):e6948.
95. Vita R, Zarebski L, Greenbaum JA, et al. The immune epitope database 2.0. *Nucleic Acids Res*. 2010;38(Database issue):D854-862.
96. Lafuente E, Reche P. Prediction of MHC-Peptide Binding: A Systematic and Comprehensive Overview. *Curr Pharm Des*. 2009;
97. Tong JC, Ren EC. Immunoinformatics: Current trends and future directions. *Drug Discov Today*. 2009;14(13-14):684-689.
98. Tong JC, Tan TW, Ranganathan S. Methods and protocols for prediction of immunogenic epitopes. *Brief Bioinform*. 2007;8(2):96-108.
99. Lundegaard C, Lund O, Buus S, Nielsen M. Major histocompatibility complex class I binding predictions as a tool in epitope discovery. *Immunology*. 2010;130(3):309-318.
100. Falk K, Rötzschke O, Stevanović S, Jung G, Rammensee HG. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature*. 1991;351(6324):290-296.
101. Lin HH, Ray S, Tongchusak S, Reinherz EL, Brusic V. Evaluation of MHC class I peptide binding prediction servers: applications for vaccine research. *BMC Immunol*. 2008;9:8.
102. Peters B, Bui H-H, Frankild S, et al. A community resource benchmarking predictions of peptide binding to MHC-I molecules. *PLoS Comput Biol*. 2006;2(6):e65.
103. Zhang H, Lundegaard C, Nielsen M. Pan-specific MHC class I predictors: a benchmark of HLA class I pan-specific prediction methods. *Bioinformatics*. 2009;25(1):83-89.
104. Klug F, Miller M, Schmidt H, Stevanovic S. Characterization of MHC Ligands for Peptide Based Tumor Vaccination. *Curr Pharm Des*. 2009;

105. Provenzano M, Panelli MC, Mocellin S, et al. MHC-peptide specificity and T-cell epitope mapping: where immunotherapy starts. *Trends Mol Med*. 2006;12(10):465-472.
106. Li Pira G, Ivaldi F, Moretti P, Manca F. High throughput T epitope mapping and vaccine development. *J Biomed Biotechnol*. 2010;2010:325720.
107. Buchli R, Vangundy RS, Giberson CF, Hildebrand WH. Critical factors in the development of fluorescence polarization-based peptide binding assays: an equilibrium study monitoring specific peptide binding to soluble HLA-A*0201. *J Immunol Methods*. 2006;314(1-2):38-53.
108. Harndahl M, Justesen S, Lamberth K, et al. Peptide binding to HLA class I molecules: homogenous, high-throughput screening, and affinity assays. *J Biomol Screen*. 2009;14(2):173-180.
109. Chen W, Khilko S, Fecondo J, Margulies DH, McCluskey J. Determinant selection of major histocompatibility complex class I-restricted antigenic peptides is explained by class I-peptide affinity and is strongly influenced by nondominant anchor residues. *J Exp Med*. 1994;180(4):1471-1483.
110. Schumacher TN, Heemels MT, Neefjes JJ, et al. Direct binding of peptide to empty MHC class I molecules on intact cells and in vitro. *Cell*. 1990;62(3):563-567.
111. Hosken NA, Bevan MJ. Defective presentation of endogenous antigen by a cell line expressing class I molecules. *Science*. 1990;248(4953):367-370.
112. Kessler JH, Mommaas B, Mutis T, et al. Competition-based cellular peptide binding assays for 13 prevalent HLA class I alleles using fluorescein-labeled synthetic peptides. *Hum Immunol*. 2003;64(2):245-255.
113. van der Burg SH, Visseren MJ, Brandt RM, Kast WM, Melief CJ. Immunogenicity of peptides bound to MHC class I molecules depends on the MHC-peptide complex stability. *J Immunol*. 1996;156(9):3308-3314.
114. Huczko EL, Bodnar WM, Benjamin D, et al. Characteristics of endogenous peptides eluted from the class I MHC molecule HLA-B7 determined by mass spectrometry and computer modeling. *J Immunol*. 1993;151(5):2572-2587.
115. del Guercio M, Sidney J, Hermanson G, et al. Binding of a peptide antigen to multiple HLA alleles allows definition of an A2-like supertype. *J Immunol*. 1995;154(2):685-693.
116. Boisgérault F, Tieng V, Stolzenberg MC, et al. Differences in endogenous peptides presented by HLA-B*2705 and B*2703 allelic variants. Implications for susceptibility to spondylarthropathies. *J Clin Invest*. 1996;98(12):2764-2770.
117. Zeh HJ, Leder GH, Lotze MT, et al. Flow-cytometric determination of peptide-class I complex formation. Identification of p53 peptides that bind to HLA-A2. *Hum Immunol*. 1994;39(2):79-86.
118. Bertoletti A, Costanzo A, Chisari FV, et al. Cytotoxic T lymphocyte response to a wild type hepatitis B virus epitope in patients chronically infected by variant viruses carrying substitutions within the epitope. *J Exp Med*. 1994;180(3):933-943.
119. Bertoni R, Sidney J, Fowler P, et al. Human histocompatibility leukocyte antigen-binding supermotifs predict broadly cross-reactive cytotoxic T lymphocyte responses in patients with acute hepatitis. *J Clin Invest*. 1997;100(3):503-513.

120. Sette A, Sidney J, del Guercio MF, et al. Peptide binding to the most frequent HLA-A class I alleles measured by quantitative molecular binding assays. *Mol Immunol*. 1994;31(11):813-822.
121. Olsen AC, Pedersen LO, Hansen AS, et al. A quantitative assay to measure the interaction between immunogenic peptides and purified class I major histocompatibility complex molecules. *Eur J Immunol*. 1994;24(2):385-392.
122. Ruppert J, Sidney J, Celis E, et al. Prominent role of secondary anchor residues in peptide binding to HLA-A2.1 molecules. *Cell*. 1993;74(5):929-937.
123. Grotenbreg GM, Nicholson MJ, Fowler KD, et al. Empty Class II Major Histocompatibility Complex Created by Peptide Photolysis Establishes the Role of DM in Peptide Association. *J Biol Chem*. 2007;282(29):21425-21436.
124. Bakker AH, Hoppes R, Linnemann C, et al. Conditional MHC class I ligands and peptide exchange technology for the human MHC gene products HLA-A1, -A3, -A11, and -B7. *Proc Natl Acad Sci U S A*. 2008;105(10):3825-3830.
125. Chew SL, Or MY, Chang CX, et al. Stability screening of arrays of major histocompatibility complexes on combinatorially encoded flow cytometry beads. *J Biol Chem*. 2011;286(32):28466-28475.
126. Sylvester-Hvid C, Kristensen N, Blicher T, et al. Establishment of a quantitative ELISA capable of determining peptide - MHC class I interaction. *Tissue Antigens*. 2002;59(4):251-258.
127. Hadrup SR, Toebes M, Rodenko B, et al. High-throughput T-cell epitope discovery through MHC peptide exchange. *Methods Mol Biol*. 2009;524:383-405.
128. Udaka K, Tsomides TJ, Eisen HN. A naturally occurring peptide recognized by alloreactive CD8+ cytotoxic T lymphocytes in association with a class I MHC protein. *Cell*. 1992;69(6):989-998.
129. Cox AL, Skipper J, Chen Y, et al. Identification of a peptide recognized by five melanoma-specific human cytotoxic T cell lines. *Science*. 1994;264(5159):716-719.
130. Schirle M, Keilholz W, Weber B, et al. Identification of tumor-associated MHC class I ligands by a novel T cell-independent approach. *Eur J Immunol*. 2000;30(8):2216-2225.
131. Antwi K, Hanavan PD, Myers CE, et al. Proteomic identification of an MHC-binding peptidome from pancreas and breast cancer cell lines. *Mol Immunol*. 2009;46(15):2931-2937.
132. Hofmann S, Gluckmann M, Kausche S, et al. Rapid and sensitive identification of major histocompatibility complex class I-associated tumor peptides by Nano-LC MALDI MS/MS. *Mol Cell Proteomics*. 2005;4(12):1888-1897.
133. Hickman HD, Luis AD, Buchli R, et al. Toward a definition of self: proteomic evaluation of the class I peptide repertoire. *J Immunol*. 2004;172(5):2944-2952.
134. Altman JD, Moss PA, Goulder PJ, et al. Phenotypic analysis of antigen-specific T lymphocytes. *Science*. 1996;274(5284):94-96.
135. Allan DSJ, Lepin EJM, Braud VM, O'Callaghan CA, McMichael AJ. Tetrameric complexes of HLA-E, HLA-F, and HLA-G. *J Immunol Methods*. 2002;268(1):43-50.
136. Bakker AH, Schumacher TNM. MHC multimer technology: current status and future prospects. *Curr Opin Immunol*. 2005;17(4):428-433.

137. Wooldridge L, Lissina A, Cole DK, et al. Tricks with tetramers: how to get the most from multimeric peptide-MHC. *Immunology*. 2009;126(2):147-164.
138. Hadrup SR, Bakker AH, Shu CJ, et al. Parallel detection of antigen-specific T-cell responses by multidimensional encoding of MHC multimers. *Nat Methods*. 2009;6(7):520-526.
139. Newell EW, Sigal N, Bendall SC, Nolan GP, Davis MM. Cytometry by Time-of-Flight Shows Combinatorial Cytokine Expression and Virus-Specific Cell Niches within a Continuum of CD8(+) T Cell Phenotypes. *Immunity*. 2012;36(1):142-152.
140. Andersen RS, Kvistborg P, Frosig TM, et al. Parallel detection of antigen-specific T cell responses by combinatorial encoding of MHC multimers. *Nat Protoc*. 2012;7(5):891-902.
141. Toebes M, Coccoris M, Bins A, et al. Design and use of conditional MHC class I ligands. *Nat Med*. 2006;12(2):246-251.
142. Rodenko B, Toebes M, Hadrup SR, et al. Generation of peptide-MHC class I complexes through UV-mediated ligand exchange. *Nat Protoc*. 2006;1(3):1120-1132.
143. Rodenko B, Toebes M, Celie PH, et al. Class I major histocompatibility complexes loaded by a periodate trigger. *J Am Chem Soc*. 2009;131(34):12305-12313.
144. Wilson DC, Grotenbreg GM, Liu K, et al. Differential Regulation of Effector- and Central-Memory Responses to Toxoplasma gondii Infection by IL-12 Revealed by Tracking of Tgd057-Specific CD8+ T Cells. *PLoS Pathog*. 2010;6(3):e1000815.
145. Frickel EM, Sahoo N, Hopp J, et al. Parasite stage-specific recognition of endogenous Toxoplasma gondii-derived CD8+ T cell epitopes. *J Infect Dis*. 2008;198(11):1625-1633.
146. Gredmark-Russ S, Cheung EJ, Isaacson MK, Ploegh HL, Grotenbreg GM. The CD8 T-cell response against murine gammaherpesvirus 68 is directed toward a broad repertoire of epitopes from both early and late antigens. *J Virol*. 2008;82(24):12205-12212.
147. Grotenbreg GM, Roan NR, Guillen E, et al. Discovery of CD8+ T cell epitopes in Chlamydia trachomatis infection through use of caged class I MHC tetramers. *Proc Natl Acad Sci U S A*. 2008;105(10):3831-3836.
148. Newell EW, Klein LO, Yu W, Davis MM. Simultaneous detection of many T-cell specificities using combinatorial tetramer staining. *Nat Methods*. 2009;6(7):497-499.
149. Hombrink P, Hadrup SR, Bakker A, et al. High-throughput identification of potential minor histocompatibility antigens by MHC tetramer-based screening: feasibility and limitations. *PLoS One*. 2011;6(8):e22523.
150. Unger WW, Velthuis J, Abreu JR, et al. Discovery of low-affinity preproinsulin epitopes and detection of autoreactive CD8 T-cells using combinatorial MHC multimers. *J Autoimmun*. 2011;37(3):151-159.
151. Lewis CE. Detecting cytokine production at the single-cell level. *Cytokine*. 1991;3(3):184-188.
152. Kern F, Surel IP, Brock C, et al. T-cell epitope mapping by flow cytometry. *Nat Med*. 1998;4(8):975-978.
153. Sylwester AW, Mitchell BL, Edgar JB, et al. Broadly targeted human cytomegalovirus-specific CD4+ and CD8+ T cells dominate the memory compartments of exposed subjects. *J Exp Med*. 2005;202(5):673-685.

154. Jung T, Schauer U, Heusser C, Neumann C, Rieger C. Detection of intracellular cytokines by flow cytometry. *J Immunol Methods*. 1993;159(1-2):197-207.
155. Picker LJ, Singh MK, Zdraveski Z, et al. Direct demonstration of cytokine synthesis heterogeneity among human memory/effector T cells by flow cytometry. *Blood*. 1995;86(4):1408-1419.
156. Prussin C, Metcalfe DD. Detection of intracytoplasmic cytokine using flow cytometry and directly conjugated anti-cytokine antibodies. *J Immunol Methods*. 1995;188(1):117-128.
157. Foster B, Prussin C, Liu F, Whitmire JK, Whitton JL. Detection of intracellular cytokines by flow cytometry. *Curr Protoc Immunol*. 2007;Chapter 6:Unit 6 24.
158. Betts MR, Brenchley JM, Price DA, et al. Sensitive and viable identification of antigen-specific CD8+ T cells by a flow cytometric assay for degranulation. *J Immunol Methods*. 2003;281(1-2):65-78.
159. Wolfl M, Kuball J, Ho WY, et al. Activation-induced expression of CD137 permits detection, isolation, and expansion of the full repertoire of CD8+ T cells responding to antigen without requiring knowledge of epitope specificities. *Blood*. 2007;110(1):201-210.
160. Zajac AJ, Blattman JN, Murali-Krishna K, et al. Viral immune evasion due to persistence of activated T cells without effector function. *J Exp Med*. 1998;188(12):2205-2213.
161. Benson JM, Campbell KA, Guan Z, et al. T-cell activation and receptor downmodulation precede deletion induced by mucosally administered antigen. *J Clin Invest*. 2000;106(8):1031-1038.
162. Mehta BA, Maino VC. Simultaneous detection of DNA synthesis and cytokine production in staphylococcal enterotoxin B activated CD4+ T lymphocytes by flow cytometry. *J Immunol Methods*. 1997;208(1):49-59.
163. Michalek J, Collins RH, Hill BJ, Brenchley JM, Douek DC. Identification and monitoring of graft-versus-host specific T-cell clone in stem cell transplantation. *Lancet*. 2003;361(9364):1183-1185.
164. Houck DW, Loken MR. Simultaneous analysis of cell surface antigens, bromodeoxyuridine incorporation and DNA content. *Cytometry*. 1985;6(6):531-538.
165. McGill J, Legge KL. Cutting edge: contribution of lung-resident T cell proliferation to the overall magnitude of the antigen-specific CD8 T cell response in the lungs following murine influenza virus infection. *J Immunol*. 2009;183(7):4177-4181.
166. Lyons AB, Parish CR. Determination of lymphocyte division by flow cytometry. *J Immunol Methods*. 1994;171(1):131-137.
167. Mintern J, Li M, Davey GM, et al. The use of carboxyfluorescein diacetate succinimidyl ester to determine the site, duration and cell type responsible for antigen presentation in vivo. *Immunol Cell Biol*. 1999;77(6):539-543.
168. Kirak O, Frickel EM, Grotenbreg GM, et al. Transnuclear mice with predefined T cell receptor specificities against Toxoplasma gondii obtained via SCNT. *Science*. 2010;328(5975):243-248.
169. Chattopadhyay PK, Hogerkorp C-M, Roederer M. A chromatic explosion: the development and future of multiparameter flow cytometry. *Immunology*. 2008;125(4):441-449.

170. Precopio ML, Betts MR, Parrino J, et al. Immunization with vaccinia virus induces polyfunctional and phenotypically distinctive CD8(+) T cell responses. *J Exp Med*. 2007;204(6):1405-1416.
171. Wallace PK, Tario JD, Jr., Fisher JL, et al. Tracking antigen-driven responses by flow cytometry: monitoring proliferation by dye dilution. *Cytometry A*. 2008;73(11):1019-1034.
172. Versteegen JM, Logtenberg T, Ballieux RE. Enumeration of IFN-gamma-producing human lymphocytes by spot-ELISA. A method to detect lymphokine-producing lymphocytes at the single-cell level. *J Immunol Methods*. 1988;111(1):25-29.
173. Czerkinsky C, Andersson G, Ekre HP, et al. Reverse ELISPOT assay for clonal analysis of cytokine production. I. Enumeration of gamma-interferon-secreting cells. *J Immunol Methods*. 1988;110(1):29-36.
174. Wulf M, Hoehn P, Trinder P. Identification and validation of T-cell epitopes using the IFN-gamma ELISPOT assay. *Methods Mol Biol*. 2009;524:439-446.
175. Wen J, Duan Z, Jiang L. Identification of a dengue virus-specific HLA-A*0201-restricted CD8+ T cell epitope. *J Med Virol*. 2010;82(4):642-648.
176. Yauch LE, Zellweger RM, Kotturi MF, et al. A protective role for dengue virus-specific CD8+ T cells. *J Immunol*. 2009;182(8):4865-4873.
177. Li CK, Wu H, Yan H, et al. T cell responses to whole SARS coronavirus in humans. *J Immunol*. 2008;181(8):5490-5500.
178. Manz R, Assenmacher M, Pfluger E, Miltenyi S, Radbruch A. Analysis and sorting of live cells according to secreted molecules, relocated to a cell-surface affinity matrix. *Proc Natl Acad Sci U S A*. 1995;92(6):1921-1925.
179. Brosterhus H, Brings S, Leyendeckers H, et al. Enrichment and detection of live antigen-specific CD4(+) and CD8(+) T cells based on cytokine secretion. *Eur J Immunol*. 1999;29(12):4053-4059.
180. Campbell JD. Detection and enrichment of antigen-specific CD4+ and CD8+ T cells based on cytokine secretion. *Methods*. 2003;31(2):150-159.
181. Pittet MJ, Zippelius A, Speiser DE, et al. Ex vivo IFN-gamma secretion by circulating CD8 T lymphocytes: implications of a novel approach for T cell monitoring in infectious and malignant diseases. *J Immunol*. 2001;166(12):7634-7640.
182. Assenmacher M, Lohning M, Scheffold A, et al. Sequential production of IL-2, IFN-gamma and IL-10 by individual staphylococcal enterotoxin B-activated T helper lymphocytes. *Eur J Immunol*. 1998;28(5):1534-1543.
183. Oelke M, Kurokawa T, Hentrich I, et al. Functional characterization of CD8(+) antigen-specific cytotoxic T lymphocytes after enrichment based on cytokine secretion: comparison with the MHC-tetramer technology. *Scand J Immunol*. 2000;52(6):544-549.
184. Rauser G, Einsele H, Sinzger C, et al. Rapid generation of combined CMV-specific CD4+ and CD8+ T-cell lines for adoptive transfer into recipients of allogeneic stem cell transplants. *Blood*. 2004;103(9):3565-3572.
185. Li Pira G, Ivaldi F, Tripodi G, Martinengo M, Manca F. Positive selection and expansion of cytomegalovirus-specific CD4 and CD8 T cells in sealed systems: potential applications for adoptive cellular immunoreconstitution. *J Immunother*. 2008;31(8):762-770.
186. Oelke M, Moehrle U, Chen JL, et al. Generation and purification of CD8+ melan-A-specific cytotoxic T lymphocytes for adoptive transfer in tumor immunotherapy. *Clin Cancer Res*. 2000;6(5):1997-2005.

187. McKinney DM, Skvoretz R, Qin M, Ishioka G, Sette A. Characterization of an in situ IFN-gamma ELISA assay which is able to detect specific peptide responses from freshly isolated splenocytes induced by DNA minigene immunization. *J Immunol Methods*. 2000;237(1-2):105-117.
188. Li Pira G, Ivaldi F, Bottone L, Manca F. High throughput functional microdissection of pathogen-specific T-cell immunity using antigen and lymphocyte arrays. *J Immunol Methods*. 2007;326(1-2):22-32.
189. Li Pira G, Ivaldi F, Dentone C, et al. Evaluation of antigen-specific T-cell responses with a miniaturized and automated method. *Clin Vaccine Immunol*. 2008;15(12):1811-1818.
190. Provenzano M, Mocellin S, Bonginelli P, et al. Ex vivo screening for immunodominant viral epitopes by quantitative real time polymerase chain reaction (qRT-PCR). *J Transl Med*. 2003;1(1):12.
191. Lim JB, Provenzano M, Kwon OH, et al. Identification of HLA-A33-restricted CMV pp65 epitopes as common targets for CD8(+) CMV-specific cytotoxic T lymphocytes. *Exp Hematol*. 2006;34(3):296-307.
192. Chen DS, Soen Y, Stuge TB, et al. Marked differences in human melanoma antigen-specific T cell responsiveness after vaccination using a functional microarray. *PLoS Med*. 2005;2(10):e265.
193. Deviren G, Gupta K, Paulaitis ME, Schneck JP. Detection of antigen-specific T cells on p/MHC microarrays. *J Mol Recognit*. 2007;20(1):32-38.
194. Soen Y, Chen DS, Kraft DL, Davis MM, Brown PO. Detection and characterization of cellular immune responses using peptide-MHC microarrays. *PLoS Biol*. 2003;1(3):E65.
195. Stone JD, Demkowicz WE, Jr., Stern LJ. HLA-restricted epitope identification and detection of functional T cell responses by using MHC-peptide and costimulatory microarrays. *Proc Natl Acad Sci U S A*. 2005;102(10):3744-3749.
196. Kwong GA, Radu CG, Hwang K, et al. Modular nucleic acid assembled p/MHC microarrays for multiplexed sorting of antigen-specific T cells. *J Am Chem Soc*. 2009;131(28):9695-9703.
197. McMichael AJ, Borrow P, Tomaras GD, Goonetilleke N, Haynes BF. The immune response during acute HIV-1 infection: clues for vaccine development. *Nat Rev Immunol*. 2010;10(1):11-23.
198. Walker BD, Burton DR. Toward an AIDS vaccine. *Science*. 2008;320(5877):760-764.
199. Barouch DH. Challenges in the development of an HIV-1 vaccine. *Nature*. 2008;455(7213):613-619.
200. Gaudieri S, Rauch A, Park L, et al. Evidence of viral adaptation to HLA class I-restricted immune pressure in chronic hepatitis C virus infection. *J Virol*. 2006;80(22):11094-11104.
201. Moore CB, John M, James IR, et al. Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. *Science*. 2002;296(5572):1439-1443.
202. de Campos-Lima PO, Gavioli R, Zhang QJ, et al. HLA-A11 epitope loss isolates of Epstein-Barr virus from a highly A11+ population. *Science*. 1993;260(5104):98-100.
203. Tan AT, Koh S, Goh V, Bertoletti A. Understanding the immunopathogenesis of chronic hepatitis B virus: An Asian prospective. *J Gastroenterol Hepatol*. 2008;23(6):833-843.

204. Prugnolle F, Manica A, Charpentier M, et al. Pathogen-driven selection and worldwide HLA class I diversity. *Curr Biol*. 2005;15(11):1022-1027.
205. Rodenhuis-Zybert IA, Wilschut J, Smit JM. Dengue virus life cycle: viral and host factors modulating infectivity. *Cell Mol Life Sci*. 2010;67(16):2773-2786.
206. Halstead SB. Dengue. *Lancet*. 2007;370(9599):1644-1652.
207. Gubler DJ. Dengue and dengue hemorrhagic fever. *Clin Microbiol Rev*. 1998;11(3):480-496.
208. Gubler DJ. Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century. *Trends Microbiol*. 2002;10(2):100-103.
209. Tan GK, Alonso S. Pathogenesis and prevention of dengue virus infection: state-of-the-art. *Curr Opin Infect Dis*. 2009;22(3):302-308.
210. World Health Organization: Impact of Dengue. http://www.who.int/csr/disease/dengue/impact/en/. Updated Accessed 2012.
211. Guzman MG, Halstead SB, Artsob H, et al. Dengue: a continuing global threat. *Nat Rev Microbiol*. 2010;8(12 Suppl):S7-16.
212. Julander JG, Perry ST, Shresta S. Important advances in the field of anti-dengue virus research. *Antivir Chem Chemother*. 2011;21(3):105-116.
213. Bente DA, Rico-Hesse R. Models of dengue virus infection. *Drug Discov Today Dis Models*. 2006;3(1):97-103.
214. Yauch LE, Shresta S. Mouse models of dengue virus infection and disease. *Antiviral Res*. 2008;80(2):87-93.
215. Schul W, Liu W, Xu HY, Flamand M, Vasudevan SG. A dengue fever viremia model in mice shows reduction in viral replication and suppression of the inflammatory response after treatment with antiviral drugs. *J Infect Dis*. 2007;195(5):665-674.
216. Johnson AJ, Roehrig JT. New mouse model for dengue virus vaccine testing. *J Virol*. 1999;73(1):783-786.
217. Shresta S, Sharar KL, Prigozhin DM, Beatty PR, Harris E. Murine model for dengue virus-induced lethal disease with increased vascular permeability. *J Virol*. 2006;80(20):10208-10217.
218. Tan GK, Ng JKW, Trasti SL, et al. A non mouse-adapted dengue virus strain as a new model of severe dengue infection in AG129 mice. In *PLoS Negl Trop Dis*. 2010: e672.
219. Tan GK, Ng JK, Lim AH, et al. Subcutaneous infection with non-mouse adapted Dengue virus D2Y98P strain induces systemic vascular leakage in AG129 mice. *Ann Acad Med Singapore*. 2011;40(12):523-510.
220. Moutaftsi M, Peters B, Pasquetto V, et al. A consensus epitope prediction approach identifies the breadth of murine T(CD8+)-cell responses to vaccinia virus. In *Nat Biotechnol*. 2006: 817-819.
221. Garboczi DN, Hung DT, Wiley DC. HLA-A2-peptide complexes: refolding and crystallization of molecules expressed in Escherichia coli and complexed with single antigenic peptides. *Proc Natl Acad Sci U S A*. 1992;89(8):3429-3433.
222. Grant D, Tan GK, Qing M, et al. A single amino acid in nonstructural protein NS4B confers virulence to dengue virus in AG129 mice through enhancement of viral RNA synthesis. *J Virol*. 2011;85(15):7775-7787.
223. Suwannasaen D, Romphruk A, Leelayuwat C, Lertmemongkolchai G. Bystander T cells in human immune responses to dengue antigens. *BMC Immunol*. 2010;11:47.

224. Budd RC, Cerottini JC, Horvath C, et al. Distinction of virgin and memory T lymphocytes. Stable acquisition of the Pgp-1 glycoprotein concomitant with antigenic stimulation. In *J Immunol*. 1987: 3120-3129.
225. Tripp RA, Hou S, Doherty PC. Temporal loss of the activated L-selectin-low phenotype for virus-specific CD8+ memory T cells. In *J Immunol*. 1995: 5870-5875.
226. Loetscher M, Gerber B, Loetscher P, et al. Chemokine receptor specific for IP10 and mig: structure, function, and expression in activated T-lymphocytes. *J Exp Med*. 1996;184(3):963-969.
227. Deckhut AM, Allan W, McMickle A, et al. Prominent usage of V beta 8.3 T cells in the H-2Db-restricted response to an influenza A virus nucleoprotein epitope. *J Immunol*. 1993;151(5):2658-2666.
228. Kedzierska K, Turner SJ, Doherty PC. Conserved T cell receptor usage in primary and recall responses to an immunodominant influenza virus nucleoprotein epitope. *Proc Natl Acad Sci U S A*. 2004;101(14):4942-4947.
229. Mathew A, Rothman AL. Understanding the contribution of cellular immunity to dengue disease pathogenesis. *Immunol Rev*. 2008;225:300-313.
230. Burke DS, Kliks S. Antibody-dependent enhancement in dengue virus infections. *J Infect Dis*. 2006;193(4):601-603; author reply 603-604.
231. Wahala WM, Silva AM. The human antibody response to dengue virus infection. *Viruses*. 2011;3(12):2374-2395.
232. Mongkolsapaya J, Dejnirattisai W, Xu XN, et al. Original antigenic sin and apoptosis in the pathogenesis of dengue hemorrhagic fever. *Nat Med*. 2003;9(7):921-927.
233. Miller S, Sparacio S, Bartenschlager R. Subcellular localization and membrane topology of the Dengue virus type 2 Non-structural protein 4B. *J Biol Chem*. 2006;281(13):8854-8863.
234. Hanley KA, Manlucu LR, Gilmore LE, et al. A trade-off in replication in mosquito versus mammalian systems conferred by a point mutation in the NS4B protein of dengue virus type 4. *Virology*. 2003;312(1):222-232.
235. National Center for Biotechnology Information: Protein. http://www.ncbi.nlm.nih.gov/protein. Updated Accessed 2012
236. Tajima S, Takasaki T, Kurane I. Restoration of replication-defective dengue type 1 virus bearing mutations in the N-terminal cytoplasmic portion of NS4A by additional mutations in NS4B. *Archives of Virology*. 2011;156(1):63-69.
237. An J, Zhou DS, Zhang JL, et al. Dengue-specific CD8+ T cells have both protective and pathogenic roles in dengue virus infection. *Immunol Lett*. 2004;95(2):167-174.
238. Casalegno-Garduno R, Schmitt A, Yao J, et al. Multimer technologies for detection and adoptive transfer of antigen-specific T cells. *Cancer Immunol Immunother*. 2010;59(2):195-202.
239. Knabel M, Franz TJ, Schiemann M, et al. Reversible MHC multimer staining for functional isolation of T-cell populations and effective adoptive transfer. *Nat Med*. 2002;8(6):631-637.
240. Schmitt A, Tonn T, Busch DH, et al. Adoptive transfer and selective reconstitution of streptamer-selected cytomegalovirus-specific CD8+ T cells leads to virus clearance in patients after allogeneic peripheral blood stem cell transplantation. *Transfusion*. 2011;51(3):591-599.

241. Vincent BG, Young EF, Buntzman AS, et al. Toxin-coupled MHC class I tetramers can specifically ablate autoreactive CD8+ T cells and delay diabetes in nonobese diabetic mice. *J Immunol*. 2010;184(8):4196-4204.
242. Hess PR, Barnes C, Woolard MD, et al. Selective deletion of antigen-specific CD8+ T cells by MHC class I tetramers coupled to the type I ribosome-inactivating protein saporin. *Blood*. 2007;109(8):3300-3307.
243. Davis MM, Altman JD, Newell EW. Interrogating the repertoire: broadening the scope of peptide-MHC multimer analysis. *Nat Rev Immunol*. 2011;11(8):551-558.
244. Constantin CM, Bonney EE, Altman JD, Strickland OL. Major histocompatibility complex (MHC) tetramer technology: an evaluation. *Biol Res Nurs*. 2002;4(2):115-127.
245. Samuel RV, Hanke T. Construction of MHC class I-peptide tetrameric complexes for analysis of T-cell-mediated immune responses. *Methods Mol Med*. 2003;87:279-288.
246. Brackenridge S, Evans EJ, Toebes M, et al. An early HIV mutation within an HLA-B*57-restricted T cell epitope abrogates binding to the killer inhibitory receptor 3DL1. *J Virol*. 2011;85(11):5415-5422.
247. Population Reference Bureau. 2011 World Population Data Sheet. http://www.prb.org/Publications/Datasheets/2011/world-population-data-sheet/data-sheet.aspx. Updated July 2011. Accessed November, 2011.
248. Liu J, Zhang S, Tan S, Zheng B, Gao GF. Revival of the identification of cytotoxic T-lymphocyte epitopes for immunological diagnosis, therapy and vaccine development. *Exp Biol Med (Maywood)*. 2011;236(3):253-267.
249. Dunn PP, Day S, Williams S, Bendukidze N. DNA sequencing as a tissue-typing tool. *Methods Mol Med*. 2004;91:233-246.
250. Cotton LA, Rahman MA, Ng C, et al. HLA class I sequence-based typing using DNA recovered from frozen plasma. *J Immunol Methods*. 2012;
251. Sidney J, del Guercio MF, Southwood S, et al. The HLA-A*0207 peptide binding repertoire is limited to a subset of the A*0201 repertoire. *Hum Immunol*. 1997;58(1):12-20.
252. Hossain MS, Tomiyama H, Inagawa T, et al. Identification and characterization of HLA-A*3303-restricted, HIV type 1 Pol- and Gag-derived cytotoxic T cell epitopes. *AIDS Res Hum Retroviruses*. 2003;19(6):503-510.
253. Heidecker L, Brasseur F, Probst-Kepper M, et al. Cytolytic T lymphocytes raised against a human bladder carcinoma recognize an antigen encoded by gene MAGE-A12. *J Immunol*. 2000;164(11):6041-6045.
254. Immune epitope database and analysis resource. http://www.immuneepitope.org/. Updated November 14, 2011. Accessed November, 2011.
255. Marsh SG, Parham P, Barber LD. *The HLA FactsBook*: Academic Press.2000
256. Falk K, Rotzschke O, Takiguchi M, et al. Peptide motifs of HLA-A1, -A11, -A31, and -A33 molecules. *Immunogenetics*. 1994;40(3):238-241.
257. Gehring AJ, Xue SA, Ho ZZ, et al. Engineering virus-specific T cells that target HBV infected hepatocytes and hepatocellular carcinoma cell lines. *J Hepatol*. 2011;55(1):103-110.
258. Harty JT, Badovinac VP. Shaping and reshaping CD8+ T-cell memory. *Nat Rev Immunol*. 2008;8(2):107-119.

259. Keenan RD, Ainsworth J, Khan N, et al. Purification of cytomegalovirus-specific CD8 T cells from peripheral blood using HLA-peptide tetramers. *Br J Haematol*. 2001;115(2):428-434.
260. Velthuis JH, Unger WW, Abreu JR, et al. Simultaneous detection of circulating autoreactive CD8+ T-cells specific for different islet cell-associated epitopes using combinatorial MHC multimers. *Diabetes*. 2010;59(7):1721-1730.
261. Celie PH, Toebes M, Rodenko B, et al. UV-induced ligand exchange in MHC class I protein crystals. *J Am Chem Soc*. 2009;131(34):12298-12304.
262. Wieboldt R, Ramesh D, Jabri E, et al. Synthesis and characterization of photolabile o-nitrobenzyl derivatives of urea. *J Org Chem*. 2002;67(25):8827-8831.
263. Gandhi MK, Khanna R. Human cytomegalovirus: clinical aspects, immune regulation, and emerging treatments. *Lancet Infect Dis*. 2004;4(12):725-738.
264. Richards JS, Beeson JG. The future for blood-stage vaccines against malaria. *Immunol Cell Biol*. 2009;87(5):377-390.
265. Martin MP, Carrington M. Immunogenetics of viral infections. *Curr Opin Immunol*. 2005;17(5):510-516.
266. Gaudieri S, Rauch A, Park LP, et al. Evidence of viral adaptation to HLA class I-restricted immune pressure in chronic hepatitis C virus infection. *J Virol*. 2006;80(22):11094-11104.
267. Chang CX, Dai L, Tan ZW, et al. Sources of diversity in T cell epitope discovery. *Front Biosci*. 2011;17:3014-3035.
268. Parham P, Ohta T. Population biology of antigen presentation by MHC class I molecules. *Science*. 1996;272(5258):67-74.
269. Broen K, Greupink-Draaisma A, Woestenenk R, et al. Concurrent detection of circulating minor histocompatibility antigen-specific CD8+ T cells in SCT recipients by combinatorial encoding MHC multimers. *PLoS One*. 2011;6(6):e21266.
270. Oh HL, Chia A, Chang CX, et al. Engineering T cells specific for a dominant severe acute respiratory syndrome coronavirus CD8 T cell epitope. *J Virol*. 2011;85(20):10464-10471.
271. Zandvliet ML, Falkenburg JH, Jedema I, et al. Detailed analysis of IFNg response upon activation permits efficient isolation of cytomegalovirus-specific CD8+ T cells for adoptive immunotherapy. *Journal of Immunotherapy*. 2009;32(5):513-523.
272. Pircher H, Moskophidis D, Rohrer U, et al. Viral escape by selection of cytotoxic T cell-resistant virus variants in vivo. *Nature*. 1990;346(6285):629-633.
273. Fernandez CS, Stratov I, De Rose R, et al. Rapid viral escape at an immunodominant simian-human immunodeficiency virus cytotoxic T-lymphocyte epitope exacts a dramatic fitness cost. *J Virol*. 2005;79(9):5721-5731.
274. Neumann-Haefelin C, McKiernan S, Ward S, et al. Dominant influence of an HLA-B27 restricted CD8+ T cell response in mediating HCV clearance and evolution. *Hepatology*. 2006;43(3):563-572.
275. Welsh RM, Selin LK, Szomolanyi-Tsuda E. Immunological memory to viral infections. *Annu Rev Immunol*. 2004;22:711-743.

276. Mothe B, Llano A, Ibarrondo J, et al. CTL responses of high functional avidity and broad variant cross-reactivity are associated with HIV control. *PLoS One*. 2012;7(1):e29717.
277. Billerbeck E, Kang YH, Walker L, et al. Analysis of CD161 expression on human CD8+ T cells defines a distinct functional subset with tissue-homing properties. *Proc Natl Acad Sci U S A*. 2010;107(7):3006-3011.
278. Klenerman P, Zinkernagel RM. Original antigenic sin impairs cytotoxic T lymphocyte responses to viruses bearing variant epitopes. *Nature*. 1998;394(6692):482-485.
279. Larsen MV, Lelic A, Parsons R, et al. Identification of CD8+ T cell epitopes in the West Nile virus polyprotein by reverse-immunology using NetCTL. *PLoS One*. 2010;5(9):e12697.
280. Lund O, Nascimento EJ, Maciel M, Jr., et al. Human leukocyte antigen (HLA) class I restricted epitope discovery in yellow fewer and dengue viruses: importance of HLA binding strength. *PLoS One*. 2011;6(10):e26494.
281. Hoof I, Peters B, Sidney J, et al. NetMHCpan, a method for MHC class I binding prediction beyond humans. *Immunogenetics*. 2009;61(1):1-13.
282. Nielsen M, Lundegaard C, Blicher T, et al. NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. *PLoS One*. 2007;2(8):e796.
283. Lundegaard C, Hoof I, Lund O, Nielsen M. State of the art and challenges in sequence based T-cell epitope prediction. *Immunome Res*. 2010;6 Suppl 2:S3.
284. Zhang L, Udaka K, Mamitsuka H, Zhu S. Toward more accurate pan-specific MHC-peptide binding prediction: a review of current methods and tools. *Brief Bioinform*. 2011;
285. Lehmann PV, Zhang W. Unique Strengths of ELISPOT for T Cell Diagnostics. *Methods Mol Biol*. 2012;792:3-23.
286. Anthony DD, Lehmann PV. T-cell epitope mapping using the ELISPOT approach. *Methods*. 2003;29(3):260-269.
287. Bendall SC, Simonds EF, Qiu P, et al. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science*. 2011;332(6030):687-696.
288. Rosenberg SA, Yang JC, Sherry RM, et al. Durable complete responses in heavily pretreated patients with metastatic melanoma using T-cell transfer immunotherapy. *Clin Cancer Res*. 2011;17(13):4550-4557.
289. Morgan RA, Dudley ME, Wunderlich JR, et al. Cancer regression in patients after transfer of genetically engineered lymphocytes. *Science*. 2006;314(5796):126-129.
290. Rosenberg SA. Cell transfer immunotherapy for metastatic solid cancer--what clinicians need to know. *Nat Rev Clin Oncol*. 2011;8(10):577-585.
291. Butler MO, Friedlander P, Milstein MI, et al. Establishment of antitumor memory in humans using in vitro-educated CD8+ T cells. *Sci Transl Med*. 2011;3(80):80ra34.
292. Gajewski TF. Monitoring specific T-cell responses to melanoma vaccines: ELISPOT, tetramers, and beyond. *Clin Diagn Lab Immunol*. 2000;7(2):141-144.
293. Pittet MJ, Speiser DE, Valmori D, et al. Ex vivo analysis of tumor antigen specific CD8+ T cell responses using MHC/peptide tetramers in cancer patients. *Int Immunopharmacol*. 2001;1(7):1235-1247.

294. Bettinotti MP, Panelli MC, Ruppe E, et al. Clinical and immunological evaluation of patients with metastatic melanoma undergoing immunization with the HLA-Cw*0702-associated epitope MAGE-A12:170-178. *Int J Cancer*. 2003;105(2):210-216.
295. Betts MR, Exley B, Price DA, et al. Characterization of functional and phenotypic changes in anti-Gag vaccine-induced T cell responses and their role in protection after HIV-1 infection. *Proc Natl Acad Sci U S A*. 2005;102(12):4512-4517.
296. He XS, Holmes TH, Mahmood K, et al. Phenotypic changes in influenza-specific CD8+ T cells after immunization of children and adults with influenza vaccines. *J Infect Dis*. 2008;197(6):803-811.
297. Co MD, Kilpatrick ED, Rothman AL. Dynamics of the CD8 T-cell response following yellow fever virus 17D immunization. *Immunology*. 2009;128(1 Suppl):e718-727.
298. Molldrem JJ, Lee PP, Wang C, et al. Evidence that specific T lymphocytes may participate in the elimination of chronic myelogenous leukemia. *Nat Med*. 2000;6(9):1018-1023.
299. Lee PP, Yee C, Savage PA, et al. Characterization of circulating T cells specific for tumor-associated antigens in melanoma patients. *Nat Med*. 1999;5(6):677-685.
300. Liu HG, Fan ZP, Chen WW, et al. A mutant HBs antigen (HBsAg)183-191 epitope elicits specific cytotoxic T lymphocytes in acute hepatitis B patients. *Clin Exp Immunol*. 2008;151(3):441-447.
301. Urbani S, Amadei B, Fisicaro P, et al. Heterologous T cell immunity in severe hepatitis C virus infection. *J Exp Med*. 2005;201(5):675-680.
302. Cobbold M, Khan N, Pourgheysari B, et al. Adoptive transfer of cytomegalovirus-specific CTL to stem cell transplant patients after selection by HLA-peptide tetramers. *J Exp Med*. 2005;202(3):379-386.
303. Neudorfer J, Schmidt B, Huster KM, et al. Reversible HLA multimers (Streptamers) for the isolation of human cytotoxic T lymphocytes functionally active against tumor- and virus-derived antigens. *J Immunol Methods*. 2007;320(1-2):119-131.
304. Nagata Y, Hanagiri T, Takenoyama M, et al. Identification of the HLA-Cw*0702-restricted tumor-associated antigen recognized by a CTL clone from a lung cancer patient. *Clin Cancer Res*. 2005;11(14):5265-5272.
305. Keane NM, Roberts SG, Almeida CA, et al. High-avidity, high-IFNgamma-producing CD8 T-cell responses following immune selection during HIV-1 infection. *Immunol Cell Biol*. 2012;90(2):224-234.
306. Immune epitope database and analysis resource. http://www.immuneepitope.org/. Updated Accessed June 2010.
307. Ichiki Y, Takenoyama M, Mizukami M, et al. Simultaneous cellular and humoral immune response against mutated p53 in a patient with lung cancer. *J Immunol*. 2004;172(8):4844-4850.