

Anti-CMV CD8+ T-Cell Epitope Specificities and T-Cell Receptor Repertoires in African Study Subjects



Amna Malik

Wolfson College, University of Oxford

Supervised by:

Prof. Philip J. R. Goulder

Dr. Philippa C. Matthews

This thesis is submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Paediatrics

Michaelmas Term 2017

Anti-CMV CD8+ T-Cell Epitope Specificities and T-Cell Receptor Repertoires in African Study Subjects

Amna Malik — Wolfson College — DPhil in Paediatrics – Michaelmas 2017

Abstract

This work focuses on chronic viral infection in populations in sub-Saharan Africa. Although some of these viruses are endemic, there is a paucity of data characterising the immune responses in the relevant populations affected. The primary focus of this thesis is cytomegalovirus (CMV). Almost 90% of children in Africa are infected by the age of 12 months. A strong virus-specific immune response controls but does not clear CMV infection, and CMV typically remains latent; however, reactivation is possible, particularly in the context of HIV infection and other states of immunosuppression. The high-frequency, immunodominant CD8+ T-cell responses, the CMV epitopes targeted, and the T-cell receptor repertoires underlying these responses have not previously been well defined in African populations. Furthermore, it is not known how the T-cell repertoire changes over the course of infection in response to CMV. To gain a better understanding of these aspects of CMV infection and immunity, I defined the immunodominant CMV-specific CD8+ T-cell responses within a cohort of sub-Saharan African study subjects (Chapter 3). Using next-generation sequencing of the CMV-specific CD8+ T-cell receptor (TCR) repertoire, focusing on a single immunodominant HLA-B*44:03-restricted response, I studied TCR repertoire bias (Chapter 4) and the changes in the TCR repertoire within children and adults followed longitudinally over 10 years (Chapter 5). Two other viruses, human Parvovirus 4 (PARV4) and Hepatitis B virus (HBV) are also endemic in these populations, may, like CMV, also be of particular significance in the setting of HIV co-infection, but are under-represented in the literature for Africa. I present data that I contributed to projects setting out to improve our understanding of the local epidemiology of these infections and to characterise relevant CD8+ T-cell responses (Chapter 6). Together, this work contributes towards a better understanding of immune responses that control chronic infections prevalent in African populations.

Acknowledgements

I would like to start by thanking my two supervisors who made this thesis possible, providing me with constant support, guidance and directing me with their exceptional knowledge and expertise. Thank you, Philip, for giving me the opportunity to do this D.Phil. I am extremely grateful for your continuous support, ideas, for giving me the opportunities to attend conferences and creating opportunities for collaborations. I wouldn't have come this far without you! Philippa, thank you for being there for me in my most difficult times and for believing in me. You are an inspiration to me. I have learnt a great deal from you, not just because you are an amazing scientist but also because you are a truly amazing human being. I will forever be indebted to you.

I would also like to thank all the people in the Medawar, in particular, Anna, Emily, Cathy, Chan, Chloe, Chris, Jodie, Kate, Masa, and Matt Jones. Special thanks to Siobhan for troubleshooting with me at late hours of night.

Thank you to all my dear friends in Oxford and beyond! Thank you Mehak for always putting things in perspective. Chloe, thank you for being there for me at difficult times and being just a phone call away.

Mohsan, my dear husband, you have been my rock through this D.Phil. I would not have completed this without you. Thank you for inspiring and motivating me everyday! Thank you for everything you do for me.

And finally, enormous thanks to my amazing family for always believing in me and being there for me through thick and thin. Thank you, Ami, Papa, Aqib, Jamal, Ami-ji, Abu-ji, Mama-ji, Papa-ji, and Ahsan for your immense love and support. My dear mum — I am here because of you. You are my reason for everything I do.

Contributions to this thesis

This work would not have been possible without the willingness of the study participants and the excellent efforts of nurses, doctors and lab technicians at the sites in South Africa and UK.

I am very grateful to the following individuals who have contributed their expertise to the work presented in this thesis:

Chapter 3:

- Dr. Emily Adland contributed to the conception of the study, participated in ELISpot assays;
- Dr. Henrik Kloverpris performed some of the tetramer stainings;
- Prof. Soren Buus' lab generated tetramers.

Chapter 4:

- Dr. Henrik Kloverpris performed cell sorting on Durban (Gateway cohort) samples;
- Dr. Meriem Attaf performed deep sequencing and participated in data analysis.

Chapter 5:

- Dr. Julia Roider performed the cell sorting on Durban (PEHSS cohort) samples;
- Dr. Meriem Attaf performed deep sequencing;

- Ahsan Alvi provided continued guidance and expertise in data analysis.

Chapter 6:

- Dr. Colin Sharp performed the PARV4 IgG assays and PARV4 PCRs;
- Dr. Apostolos Beloukas participated in HBsAg assays;
- Dr. Sheila Lumley led the literature review performed for the Hepi-topes project.

I would like to thank Rebecca Alan for proof-reading this thesis.

This D.Phil. was funded by the scholarship provided by the Department of Paediatrics.

Publications

First author publications

- **Malik A**, Adland E, Laker L, Kløverpris H, Fardoos R, Roider J, et al. Immunodominant Cytomegalovirus-specific CD8+ T-cell responses in sub-Saharan African populations. PLoS ONE. 2017 (*Under review*)
- **Malik A**, Attaf M, Roider J, Ndung'u T, Buus B, Kløverpris H, et al. TCR Bbias within the immunodominant HLA-B*44:03-restricted, CMV-specific CD8+ T-cell response. The Journal of Immunology. 2017 (*Late draft manuscript ready*)

Contributing author publications

- Matthews PC, **Malik A**, Simmons R, Sharp CP, Simmonds P, Klennerman P. PARV4: an emerging tetraparvovirus. PLoS Pathogen. 2014;10(5)
- Matthews PC, Sharp CP, **Malik A**, Gregory WF, Adland E, Jooste P, et al. Human parvovirus 4 infection among mothers and children in South Africa. Emerging Infectious Diseases. 2015; 21(4):713-5
- Ndhlovu Z, Kanya P, Mewalal N, Kløverpris H, Nkosi, Pretorius K, Laher F, Ogunshola F, Chopera D, Shekhar K, Ghebremichael M, Ismael N, **Malik A**, et al. Magnitude and kinetics of CD8+ T cell activation during hyperacute HIV infection impacts viral set point. Immunity. 2015;43(3):591-604

- Matthews PC, Beloukas A, **Malik A**, Carlson JM, Jooste P, Ogwu A, et al. Prevalence and Characteristics of Hepatitis B Virus (HBV) Coinfection among HIV-Positive Women in South Africa and Botswana. PLoS ONE. 2015;10(7)
- Matthews PC, Carlson JM, Beloukas A, **Malik A**, Jooste P, Ogwu A, et al. HLA-A is a Predictor of Hepatitis B e Antigen Status in HIV-Positive African Adults. The Journal of Infectious Diseases. 2016;213(8):1248-1252
- Lumley S, Noble H, Hadley MJ, Callow L, **Malik A**, Chua YY, et al. Hepitopes: A live interactive database of HLA class I epitopes in hepatitis B virus. Wellcome Open Research. 2016;1:9
- Sharp CP, Gregory WF, Hattingh L, **Malik A**, Adland E, Carlson JM, et al. PARV4 prevalence, phylogeny, immunology and coinfection with HIV, HBV and HCV in a multicentre African cohort. Wellcome Open Research. 2017;2:26

Abbreviations

α — alpha

β — beta

$\gamma\delta$ — gamma delta

A

aa — amino acid

Ag — antigen

APCs — antigen presenting cells

AIDS — acquired immune deficiency syndrome

ART — antiretroviral therapy

ATP — adenosine triphosphate

Asn — Asparagine

Ala — Alanine

B

BSA — bovine serum albumin

C

CDR3 —complementary determining region 3

CA — capsid

CDC — centers for disease control

CMV — (Human) cytomegalovirus

CTD — C-terminal domain

CTL — cytotoxic T-lymphocytes

C — cysteine

D

D — Diversity

DC — dendritic cell

DNA — deoxyribonucleotide triphosphate

dNTP — deoxynucleotide triphosphate

E

EBV — Epstein Barr virus

E. coli — Escherichia coli

ELISA - enzyme-linked immunosorbent assay

ELISpot — enzyme-linked immunospot assay

ER — endoplasmic reticulum

F

FACS — fluorescence activated cell sort

FCS — foetal calf serum

G

Gly — glycine

H

HIV — human immunodeficiency virus

HLA — human leukocyte antigen

HBV — Hepatitis B

I

ICS — Intracellular cytokine assay

IDUs — intravenous drug users

IFN γ — Interferon gamma

IE1 — Immediate early protein 1

IE2 — Immediate early protein 2

IgG — Immunoglobulin G

IL-2 — interleukin-2

IQR — interquartile range

J

J — joining

K

kb — kilobase

KIR — killer cell immunoglobulin-like receptor

L

Leu — Leucine

M

MHC — major histocompatibility complex

MSM — men who have sex with men

MCMV- Murine cytomegalovirus

N

NGS — next generation sequencing

NK — natural killer cells

NS — non-structural

NC — nucleocapsid

NFB — nuclear factor kappa B

O

OLPs — overlapping peptides

ORF —open reading frame

P

PARV4 — Human parvovirus 4
PBMCs — peripheral blood mononuclear cells
PBS — phosphate buffered saline
PCR — polymerase chain reaction
PHA — phytohaemagglutinin
PP65 — phosphoprotein 65

R

RT — real-time / reverse transcriptase
RNA — ribonucleic acid
RNAPII — RNA polymerase II
RPMI — roswell park memorial Institute medium

S

SD — standard deviation
SEM — standard error of the mean
SFC — spot-forming cells
SFU — spot-forming units
SIV — simian immunodeficiency viruses
ss — single stranded

T

TAP — transporter for antigen processing
TCR — T cell receptors
T_{CM} — central memory T-cells
TdT — terminal deoxynucleotidyl transferase
T_{EM} — effector memory T-cells
TLR — Toll-like receptors
TNF- α — Tumor necrosis factor alpha
Treg — regulatory T cell
TRAD — TCR alpha diversity gene
TRBD — TCR beta diversity gene
TRAJ — TCR alpha joining gene
TRBJ — TCR beta joining gene
TRAV — TCR alpha variable gene
TRBV — TCR beta variable gene

V

V — variable
VZV — Varicella zoster virus
VP — viral capsid

Contents

Chapter 1: Introduction	1
1.1 Chronic viral infections	1
1.2 CMV	2
1.2.1 History	2
1.2.2 Classification	3
1.2.3 Genome	4
1.2.4 Epidemiology	5
1.2.5 Transmission and risk factors	7
1.2.6 CMV latency and reactivation	7
1.2.7 Pathogenesis	9
1.2.8 CMV life cycle	11
1.2.9 CMV cell tropism	12
1.3 Innate immune responses to CMV	13
1.4 Adaptive immune response to CMV	16
1.4.1 Antigen presentation and MHC molecules	16
1.4.2 Structure and assembly of TCR	20
1.4.3 T-cell development	22
1.4.4 V(D)J recombination	25
1.4.5 TCR receptor diversity	28

1.4.6	CD8+ T-cells	29
1.4.6.1	CD8+ T-cell differentiation	29
1.4.6.2	Effector mechanisms of CD8+ T-cells.....	31
1.4.7	CD8+ T-cell responses to CMV	33
1.4.8	Proportion/size of CD8+ T-cells responses to CMV	34
1.4.9	Course of CMV infection	34
1.4.10	Phenotype of CMV-specific CD8+ T-cells.....	36
1.4.11	Factors involved in memory inflation	37
1.4.11.1	Antigen exposure.....	37
1.4.11.2	Naive and memory T-cells	38
1.4.11.3	Co-stimulatory molecules and inflammatory cytokines	39
1.4.12	Oligoclonality of CMV-specific repertoire	40
1.4.13	Other host defences to CMV.....	41
1.4.13.1	CD4+ T-cell responses to CMV	41
1.4.13.2	Delta T-cell responses to CMV	41
1.4.13.3	Humoral immune responses to CMV	42
1.4.14	Immune evasion techniques	43
1.4.14.1	MHC class I pathway.....	43
1.4.14.2	MHC class II pathway	43
1.4.14.3	NK cells	44
1.4.14.4	Homologues	45
1.5	TCR repertoire of antigen-specific CD8+ T-cell responses	46
1.5.1	Diversity of TCR repertoire and efficacy of immune response .	46
1.5.2	Diversity of TCR repertoire and magnitude of CD8+ T-cell re- sponse.....	47
1.5.3	Affinity of TCR repertoire influences immune protection	48
1.5.4	Cross reactivity	48

1.5.5	Size of naive T-cell repertoire and antigen-specific CD8+ T-cell repertoire	49
1.5.6	Genetic influence on naive repertoire	50
1.6	Immune ontogeny in children	51
1.7	Hypotheses and Aims	56
 Chapter 2: Materials and methods		57
2.1	Study subjects	57
2.2	Blood processing	59
2.2.1	PBMC extraction	59
2.2.2	Cyropreservation of cells	60
2.2.3	DNA extraction	60
2.3	HLA typing	61
2.4	CD4+ T-cell counts and viral load determinations	61
2.5	Interferon-gamma (IFN-γ) enzyme-linked immunosorbent spot (ELISpot assay)	61
2.5.1	ELISpot procedure	61
2.5.2	Peptide dilution and screening for CMV-specific responses by megamatrix	62
2.5.3	ELISpot peptide titration assay	63
2.6	Intracellular cytokine staining (ICS) assay	64
2.7	Tetramer synthesis	64
2.8	Cell staining	65
2.9	Cell sorting	65
2.10	RNA extraction	66
2.11	TCR sequencing	66
2.11.1	cDNA synthesis	66

2.11.2	Primer design	68
2.11.3	PCR amplification of TCR- α and TCR- β chains	70
2.11.4	Agarose gel electrophoresis	73
2.11.5	DNA gel extraction.....	73
2.12	Illumina high-throughput sequencing	73
2.13	TCR sequence analysis	73
2.14	Statistical analyses	74
Chapter 3:	Immunodominant cytomegalovirus-specific CD8+ T-cell re-	
	sponses in sub-Saharan African populations	75
3.1	Introduction	75
3.2	Materials and methods	77
3.2.1	Study subjects	77
3.2.2	Epitope Prediction using <i>HLA-Restrictor</i> — a tool for predic- tion of HLA restriction elements and optimal epitopes within peptides	79
3.2.3	Statistical analysis	79
3.3	Results.....	79
3.3.1	Immunodominant CMV-specific CD8+ T-cell responses within pp65, IE1 and IE2	79
3.3.2	Identification of additional novel CD8+ T-cell epitopes within pp65, IE1 and IE2	82
3.3.3	Closely related HLA I molecules shape distinct CMV-specific CD8+ T-cell hierarchies	85
3.3.4	Definition of a high frequency HLA-B*44:03-restricted IE2- specific response	89

3.3.5	Relative contribution of HLA-A, -B and -C alleles to immune recognition of CMV	91
3.4	Discussion	94
Chapter 4: TCR bias within the immunodominant HLA-B*44:03-restricted, CMV-specific CD8+ T-cell response		98
4.1	Introduction	98
4.2	Materials and methods	100
4.2.1	Study subjects	100
4.2.2	Statistical analysis	101
4.2.3	Data analysis	102
4.3	Results.....	102
4.3.1	NW8-specific T-cells display limited TCR diversity	102
4.3.2	The NW8-specific CD8+ T-cell repertoire is characterised by restricted V and J segment usage.....	104
4.3.3	TCR sharing in NW8-specific T-cell responses.....	107
4.3.4	High-frequency TCR- α chains are highly shared amongst HLA-B*44:03-infected individuals	107
4.3.5	Public NW8-specific TCRs are produced by convergent recombination	109
4.3.6	Public CDR3 motifs are strictly conserved amongst HLA-B*44:03-positive individuals.....	110
4.4	Discussion	113
Chapter 5: Studying changes in CMV-specific TCR repertoires in mother-child pairs over a span of 10+ years		116
5.1	Introduction	116
5.2	Materials and methods	119

5.2.1	Study subjects	119
5.2.2	TCR sequencing	119
5.3	Results	119
5.3.1	Analysis of NW8-specific CD8+ T-cell responses	119
5.3.2	Characteristics of NW8-repertoires in children and adults	121
5.3.3	TCR publicity	128
5.3.4	Changes in NW8-specific TCR repertoire diversity	130
5.3.5	The relationship between CD8+ T-cell response magnitude and TCR repertoire diversity in children and mothers	131
5.3.6	Overlap of TCR- β repertoires for related and unrelated mother- child pairs	133
5.4	Discussion	139

**Chapter 6: PARV4 and Hepatitis-B prevalence, CD8+ T-cell responses
and co-infection with HIV and CMV in African children and adults** **144**

6.1	Introduction	144
6.1.1	Parvovirus 4 (PARV4)	145
6.1.2	Hepatitis B virus (HBV)	147
6.2	Materials and methods	152
6.2.1	Patient cohorts	152
6.2.2	PARV4 IgG assay	153
6.2.3	HBV testing	153
6.2.4	PARV4 IFN- γ ELISpot assays	154
6.2.5	PARV4 DNA	154
6.2.6	HIV status	154
6.2.7	HIV-viral load and CD4 T-cell count	155
6.2.8	Systematic literature review for “Hepitopes” database	155

6.2.9	Statistical analysis	156
6.3	Results.....	156
6.3.1	Determination of the level of PARV4, HBV, CMV and HIV co- infection in South Africa	156
6.3.2	Impact of PARV4 and HBV on HIV disease progression	157
6.3.3	High breadth and magnitude of PARV4-specific CD8+ T-cell responses in children and adults	160
6.3.4	Impact of PARV4 infection on CMV- and HIV-specific CD8+ T-cell responses	162
6.3.5	Impact of HBV on HIV-specific CD8+ T-cell responses	165
6.3.6	HBV epitopes	165
6.4	Discussion	169
 Chapter 7: Conclusions		 172
 Appendix		 178
 Bibliography		 182

List of Figures

1.1	Schematic representation of structural components and genome organisation of CMV.	5
1.2	Seroprevalence of CMV worldwide in adult population.	6
1.3	Microbial ligand specificities of human Toll-like receptors and their location.	15
1.4	Schematic presentation of MHC class I molecule.	19
1.5	Schematic representation of MHC I antigen processing.	20
1.6	TCR protein and gene structure.	22
1.7	Hematopoietic stem cells in the bone marrow gives rise to lymphoid and myeloid progenitors.	24
1.8	Overview of V(D)J recombination of TCR- α and - β chains.	27
1.9	The expression of surface markers associated with differentiation is presented for CD8+ T-cells.	31
1.10	Frequency of CMV ORF recognition by CD8+ T-cells.	35
1.11	Course of CMV infection.	36
1.12	Evolution of the immune system.	52
1.13	Age-dependent changes in TLR-induced immune regulatory function.	53
2.1	Example of megamatrix setup.	63
2.2	Representative example of the gating strategy.	66
2.3	cDNA synthesis.	67

3.1	Immunodominant CMV-specific CD8+ T-cell responses within pp65, IE1 and IE2 measured by IFN- γ ELISpot assay for CMV+ subjects. ...	81
3.2	Validation of novel CD8+ T-cell epitopes in CMV using HLA-Class I tetramers.	84
3.3	Validation of additional novel CD8+ T-cell epitopes in CMV using HLA-Class I tetramers and intracellular cytokine assay.	87
3.4	Phenotypic frequencies of the class I HLA molecules in the African study cohort and a representative white population.	88
3.5	Phenotypic frequencies of closely related HLA I molecules and their impact on CMV specific CD8+ T-cell hierarchies.	90
3.6	Individuals making CD8+ T-cell responses against CMV IE2 epitope HLA-B*44:03-NW8.	92
3.7	Number of peptides recognised in association with individual HLA allele expression.	93
4.1	Distribution of TCR- α chain clonotypes in HLA-B*44:03 NW8-specific CD8+ T-cells.	105
4.2	Distribution of TCR- β chain clonotypes in HLA-B*44:03 NW8-specific CD8+ T-cells.	106
4.3	V and J gene usage in TCR- α and - β chain repertoires of HLA-B*44:03/NW8-specific CD8+ T-cells.	108
4.4	TCR sharing amongst NW8-specific CD8+ T-cells from CMV-infected individuals.	109
4.5	Clonal dominance and TCR publicity.	110
4.6	Number of nucleotide sequences and TCR publicity.	111
4.7	Amino acid motif conservation in public clonotypes from HLA-B*44:03 NW8-specific CD8+ T-cells.	112

5.1	Distribution of TCR- α and TCR- β chains in mother-child pairs over the course of 10+ years.	124
5.2	V and J gene usage in TCR- α and - β chain repertoires of HLA-B*44:03-NW8-specific CD8+ T-cells.	126
5.3	TCR- α and - β chain pairings.	127
5.4	Clonal dominance and TCR publicity.	129
5.5	Number of nucleotide sequences and TCR publicity.	130
5.6	Changes in diversity of TCR- α and - β chain clonotypes in children and their mothers over the course of 10+ years.	132
5.7	The relationship between CD8+ T-cell response magnitude and TCR repertoire diversity in children and mothers.	134
5.8	The CD8+ T-cell response magnitude and composition of TCR- α and - β chain clonotypes.	138
5.9	TCR sharing amongst NW8-specific CD8+ T-cells from related versus unrelated mother-child pairs.	139
6.1	Schematic diagram of PARV4 genome and schematic views of HLA-B*57:03-restricted epitope variants in PARV4 genome.	150
6.2	Relationship of PARV4 IgG status and HBsAg status with age and HIV status.	158
6.3	Venn diagram of CMV, HIV, PARV4 and HBV prevalence and co-infection in multicentre sub-Saharan African cohort.	159
6.4	IFN- γ CD8+ T-cell responses to PARV4 peptides determined by ELISpot assay.	161
6.5	Validation of a novel HLA-B*5703-restricted epitope in PARV4 NS protein using peptide titration assay.	164
6.6	Impact of PARV4 IgG serostatus on CMV and PARV4-specific CD8+ T-cell responses.	166

6.7 Sample screenshot of the “Hepitope” database. 168

List of Tables

1.1	Table of herpesviruses that affect human and their clinical manifestations.....	3
1.2	Clinical features of CMV infection.....	10
2.1	PCR cycling condition for first step of cDNA synthesis	68
2.2	cDNA synthesis master mix reaction.....	68
2.3	PCR cycling condition for second step of cDNA synthesis	69
2.4	Step-out PCR master mix reaction	70
2.5	Nested PCR master mix reaction.....	71
2.6	PCR primers for amplifying TCR- α and TCR- β chains	71
2.7	PCR cycling conditions for step-out PCR	72
2.8	PCR cycling conditions for Nested PCR	72
3.1	Cohort Characteristics	78
3.2	HLA associations with the most targeted 15-mer OLP's in CMV pp65, IE1 and IE2.....	83
3.3	HLA associations with the remaining 15-mer OLP's in CMV pp65, IE1 and IE2 that were targeted by $\geq 4\%$ of the study cohort.....	86
4.1	Subject Characteristics.....	103
5.1	Subject Characteristics.....	120

6.1	Prevalence and characteristics of CMV, PARV4, and HBV infection. . .	148
6.2	Prevalence of PARV4 reported in previously published studies of African populations.	149
6.3	Serological markers associated with HBV infection.	151
6.4	HLA associations with the immunogenic 15mer OLPs in NS.....	163
A1	HLA frequencies of cohort used in chapter 3.....	178
A2	Overlapping peptide sequences.....	181

Chapter 1

Introduction

1.1 Chronic viral infections

Chronic viral infections result when the primary infection is not cleared by the adaptive immune response. Diseases caused by viruses such as HIV, hepatitis B and C viruses and human herpesviruses include AIDS, chronic hepatitis, and cancer. I studied the impact of three chronic viral infections: (i) CMV, (ii) PARV4 and, (iii) Hepatitis B in sub-Saharan African populations. These opportunistic infections are especially of a great concern in sub-Saharan African populations where HIV infection is also endemic. The immunodeficiency caused by chronic HIV infection influences the consequences of co-infection with CMV, HBV and PARV4, increasing the risk of long-term morbidity and mortality. Despite these concerns, the burden and impact of CMV, HBV and PARV4 in sub-Saharan Africa have not been well characterised. I therefore set out to characterise the prevalence and characteristics of CMV, PARV4 and HBV infection in a large cohort of African individuals, with CMV being the main focus of this thesis.

1.2 CMV

1.2.1 History

Human cytomegalovirus (CMV) has been identified as an important human virus for over a century. In 1881, Ribbert first identified inclusion-bearing, large cells in sections of a kidney and parotid glands of children (Ribbert 1904). He was unable to interpret them until a report by Jesionek and Kiolemenoglou (1904), where they described similar cells as “protozoan cells” in the lungs, kidneys and liver of an 8-month luetic fetus. The eccentrically placed nuclei of these large cells contained a “central nuclear body” surrounded by two well-defined zones, an inner dark and an outer clear halo, which they named “owl's eye inclusion bodies”. Glahn and Pappenheimer (1925) reported similar structures in the lungs, liver, and kidneys of a 2 months old infant and suggested that these abnormal cells were produced by viruses. This was the first indication that cells with intranuclear inclusions might be related to a group of viruses. By 1923, several cases of large cells with similar structures (containing the typical intranuclear inclusions) isolated from a variety of organs were described. Wyatt et al. (1950) suggested the name, generalized cytomegalic inclusion disease (CID), although the viral aetiology was still unknown. The first confirmation that CID was caused by a virus came from Minder (1953). Minder observed 199 nm particles by electron microscopy, suggestive of a virus, in pancreatic cells (with intranuclear inclusion).

It was not until studies were carried out on human embryonic cell cultures by Enders et al. (1949) that virologists were able to culture viruses from the clinical cases of CID. Smith in 1956, Rowe et al. in 1956, and Weller et al. in 1957 independently isolated human CMV strains. Weller in 1960 named the virus “cytomegalovirus” and subsequently isolated CMV from the urine of infants with generalised disease

(Weller 1970).

1.2.2 Classification

CMV or human herpesvirus 5 (HHV-5) is a member of the β -herpesvirus sub-family of the Herpesviridae family. The other genus within β -herpesvirus group that infects humans is Roseolovirus (which includes HHV-6 and HHV-7). Herpesviridae comprise a large family of double-stranded DNA viruses that persist latently in the host. There are 8 members of the herpesvirus family that routinely infect humans (Table 1.1). These viruses are ancient and have co-evolved with their host for over a millennium (McGeoch et al. 1995). All herpesviruses share characteristics of a DNA genome encoding 100-200 genes and a virion structure that encapsulates the genome within an enveloped icosahedral capsid. This classification was originally based on species specificity and slow growth in cell culture. Advances in availability of nucleotide data has now led to confirmation by genetic sequence homologies to other herpesvirus genomes (Knipe and Howley 2007).

Name	Synonym	Clinical manifestations
HHV-1	Herpes simplex virus-1 (HSV-1)	Oral herpes
HHV-2	Herpes simplex virus-2 (HSV-2)	Genital herpes
HHV-3	Varicella zoster virus (VZV)	Chickenpox and shingles
HHV-4	Epstein–Barr virus (EBV)	Infectious mononucleosis
HHV-5	Cytomegalovirus (CMV)	Described in Table 1.2
HHV-6		Sixth disease (Roseola infantum)
HHV-7		Sixth disease (Roseola infantum)
HHV-8	Kaposi's sarcoma-associated herpesvirus	Kaposi's sarcoma associated with AIDS

Table 1.1: Table of herpesviruses that affect human and their clinical manifestations.

1.2.3 Genome

CMV has the largest genome of any known human virus, at 236 kb. The genome is a linear, double-stranded DNA molecule consisting of two covalently linked unique regions [Unique Long, (UL) and Unique Short (US)], which are flanked on one end by terminal repeated sequences (TRL and TRS) and on the other end by internal repeats (IRL and IRS) (Fig 1.1) (Dolan et al. 2004). The CMV genome encodes approximately 165 genes. The virion structure ranges from 200 to 300 nanometer in size and consists of a double-stranded linear DNA core in an icosahedral nucleocapsid, enveloped by a proteinaceous matrix (tegument) (Chen et al. 1999). These components are surrounded by host cell-derived lipid bilayer containing a number of virus encoded surface glycoproteins which are incorporated during the maturation process and are required for cell attachment and penetration (Crough and Khanna 2009).

The most abundant tegument protein is the lower matrix phosphoprotein 65 (pp65), also termed unique long 83 (UL83) (Varnum et al. 2004). The tegument also consists of other major proteins (pp71, pp150 and UL48 gene products), proteins that are present in small amounts and some cellular and viral RNA. Tegument proteins play important roles in all stages of the viral life cycle, including, viral entry, gene expression, immune evasion, assembly, and egress (Tomtishen 2012).

Productive infection leads to gene expression in three overlapping stages, based on the time of synthesis after infection: immediate-early (IE), delayed early (E), and late (L). IE genes are expressed during the first six hours post infection. Delayed early genes are expressed within 24 hours of infection and late genes are expressed after 24 hours of infection (Stinski 1978; Crough and Khanna 2009). IE gene expression acts as a master regulator to initiate the lytic cycle of the virus.

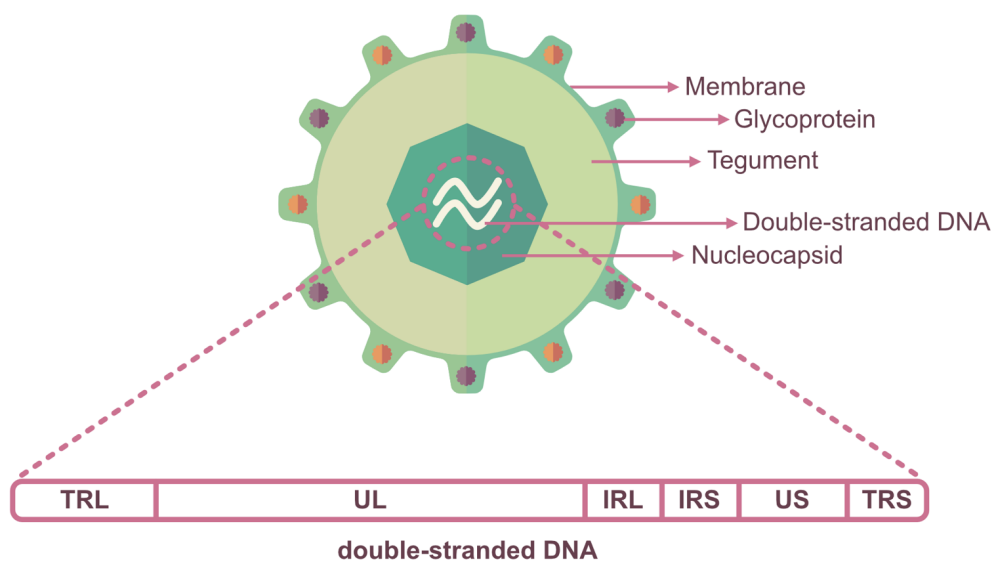


Figure 1.1: Schematic representation of structural components and genome organisation of CMV. The model of CMV is showing various components of the virus. The lettering within the individual regions of the genome represent the following features: terminal repeat long (TRL), unique long (UL), unique short (US), internal repeat long (IRL), internal repeat short (IRS), terminal repeat short (TRS), and internal repeat (IR). *Figure adapted from Crough and Khanna 2009*

1.2.4 Epidemiology

Several epidemiological studies have confirmed that human CMV is ubiquitous among the world's population, with varied seroprevalence, based on factors including breast-feeding, hygiene, socio-economic status, living standards and age. CMV is endemic in sub-Saharan African populations with almost two-thirds of infants infected by 3 months of age and 85% infected by the time they are a year old (Miles et al. 2007). By adolescence, CMV infection is virtually universal in sub-Saharan Africa (Cannon et al. 2010). The prevalence of CMV infection is also very high in developed countries (Cannon et al. 2010; Adland et al. 2015a) and increases with age (Staras et al. 2006; Dollard et al. 2011). In the USA, Australia and Europe, CMV seroprevalence among adults is estimated at between 36 and 77% (Adland et al. 2015a). Fig 1.2 shows the prevalence of CMV around the world.

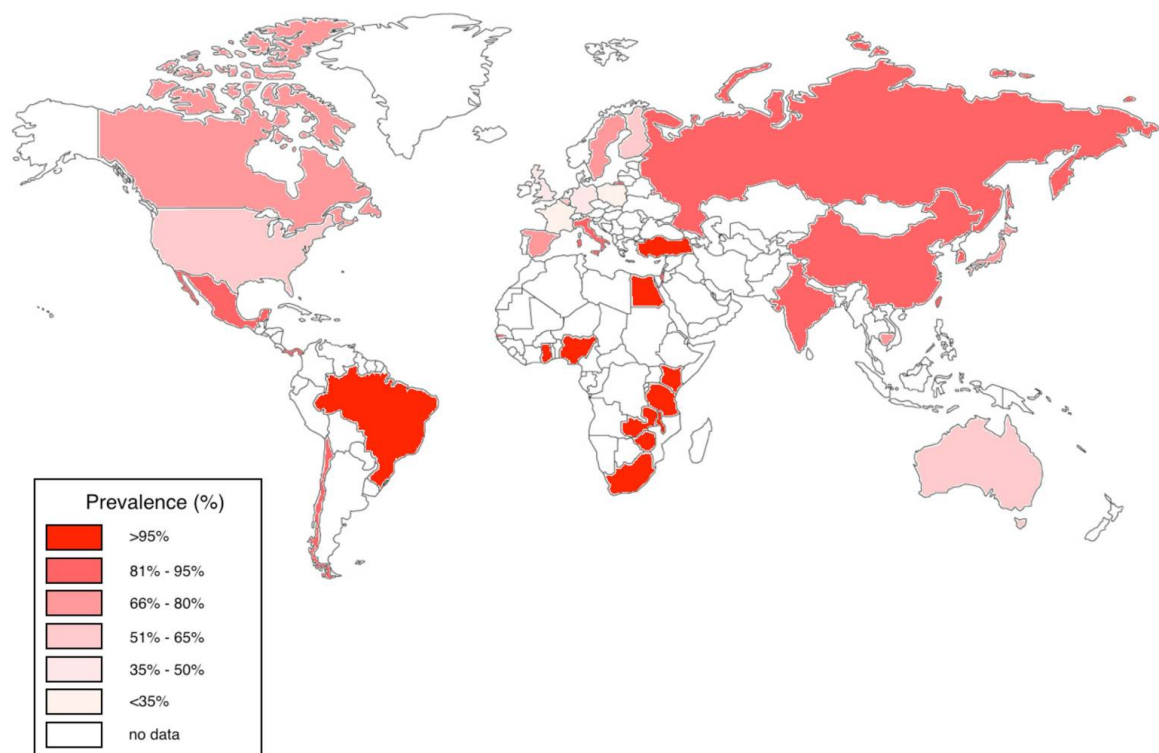


Figure 1.2: Seroprevalence of CMV worldwide in adult population. The map represents studies of adults aged 16-50 years published between 2005 and 2015 from Australia, Belgium, Brazil, Canada, Cambodia, Chile, China, Finland, France, Gambia, Germany, Ghana, India, Israel, Italy, Japan, Kenya, Mexico, Nigeria, Panama, South Africa, Spain, Sweden, Taiwan, Tanzania, Turkey, UK, USA, Zambia, and Zimbabwe. *Figure from Adland et al. 2015a*

1.2.5 Transmission and risk factors

CMV can be transmitted via saliva, sexual contact, placental transfer, breast-feeding, blood transfusion, solid-organ transplantation (SOT) and hematopoietic stem cell transplantation (Adland et al. 2015a). Breastfeeding is a common route for CMV transmission, particularly in populations with high CMV seroprevalence and high rates of breastfeeding. In African studies, co-infection with HIV, poor nutrition, low body weight and poor living conditions (overcrowding) are associated with increased CMV seroprevalence (Tembo et al. 2015; Adland et al. 2015a). The studies carried out in the western countries show that rate of CMV seropositivity is much higher among adults with HIV, suggesting that risk factors associated for acquisition of HIV and CMV are shared (Compston et al. 2009; Gianella et al. 2015). Alternatively, HIV infection may increase susceptibility to acquiring CMV infection, or vice versa.

1.2.6 CMV latency and reactivation

CMV, like other herpesviruses, is able to establish a lifelong latency within the host after primary infection, with periodic reactivation. The development of CMV latency in children takes longer than adults (Prendergast et al. 2012). Primary CMV infection in young children leads to prolonged viral excretion continuing up to five years (Pass et al. 1983; Tu et al. 2004). In contrast, adults with primary CMV infection stop continual viral shedding by 3-6 months after acquisition (Revello et al. 1998), and have only infrequent recurrences of shedding thereafter (Tu et al. 2004).

The immune response mounted during primary infection is able to limit lytic viral replication, however latency is always established and the virus is never cleared. During latency, viral gene expression is minimised to avoid immune surveillance.

Viral latency can be defined as the ability of the viral genome to maintain itself in the absence of production of infectious virions, but with the ability of the viral genome to reactivate under certain conditions (Sinclair 2008).

Although all the sites of latency have not been fully established, the most common site for CMV to develop latency is in cells of myeloid lineage, CD34+ myeloid progenitor cells and CD14+ monocytes. Studies have shown that viral DNA can be found in myeloid cells from peripheral blood of healthy seropositive individuals without any detectable infectious virus (Mendelson et al. 1996).

Three mechanisms have been proposed by which latency in CMV infection develops. The first option involves entry of virus, following attachment, directly into a latent state without expressing de novo viral genes. The second possibility is that following its entry, the virus initiates productive infection that is prematurely interrupted, which may subsequently lead to latency. Thirdly, following entry, the virus expresses a subset of viral genes unrelated to productive infection but which are necessary for successful establishment of latency (Reeves et al. 2005; Cheung et al. 2006; Crough and Khanna 2009; Goodrum et al. 2011).

There is evidence suggesting that transcriptional repression of the viral major immediate early promoter (MIEP), which would normally drive lytic cycle, through post-translational modification of histones around the MIEP, is the key event in latency (Sinclair and Sissons 2006; Sinclair 2010). A number of latency-associated genes are detectable, including UL138 (first viral gene sequenced and proven to be functionally required for CMV latency) and UL82AS (inhibits the expression of the UL82 product pp71, which activates viral IE transcription) (Bego et al. 2005; Goodrum et al. 2011). These latency associated unique genes are expressed in

the absence of productive virus replication (Cheung et al. 2006). Another possible latency-associated transcript has been identified which maps to the viral UL111.5A gene; this gene encodes a putative viral homologue of the immunomodulatory cytokine IL-10 (Rosa and Diamond 2012; Jenkins et al. 2008).

Differentiation of latently infected CD34+ cells or monocytes into macrophages or dendritic cells can cause de-repression of the MIEP and triggers the initiation of a cascade of viral genes including, immediate early, early and late gene products. This leads to viral DNA replication and reactivation of de novo virus production and lytic infection. CMV reactivation can be caused by a number of factors including, immunosuppression (for example, haematopoietic bone marrow or solid organ transplant recipients or patients with AIDS), stress, infection or inflammation (Crough and Khanna 2009; Prösch et al. 2000; Kutza et al. 1998), although the exact mechanism of reactivation is unclear. Stress catecholamines, epinephrine and norepinephrine and proinflammatory prostaglandins contribute to CMV reactivation by stimulating the IE enhancer/promoter via cyclic AMP-dependent signalling pathways (Prösch et al. 2000; Kline et al. 1998). TNF- α is the key mediator of CMV reactivation. It interacts with the TNF- α receptor on latently infected cells which is a strong stimulus for the transcription factor NF- κ B, which in turn plays an important role in stimulating the IE enhancer/promoter (Fietze et al. 1994). The reactivation of CMV from latency is a key step in pathogenesis from human CMV infection.

1.2.7 Pathogenesis

CMV is typically asymptomatic in healthy individuals, but can cause a self-limiting “glandular fever” syndrome similar to another herpesvirus, EBV (Crough and Khanna 2009). Table 1.2 shows the clinical manifestation of CMV in different hosts. CMV related disease is a major problem in immunocompromised individuals. CMV is

Hosts	Clinical manifestations
Healthy individual	Usually asymptomatic; infrequently mononucleosis with fever; myalgia; adenopathy; splenomegaly
Congenital infection	Jaundice; hepatosplenomegaly; petechiae; microcephaly;; hypotonia; seizures; lethargy
SOT recipient	Febrile illness with leukopenia and malaise; pneumonitis; enterocolitis; esophagitis or gastritis; hepatitis; retinitis
Hematopoietic SCT recipient	Pneumonitis; enterocolitis; esophagitis or gastritis; less commonly retinitis; encephalitis; hepatitis
HIV/AIDS patient	Retinitis; enterocolitis; esophagitis or gastritis; pneumonitis; hepatitis

Table 1.2: Clinical features of CMV infection. *Source: Crough and Khanna 2009*

a serious opportunistic infection in individuals with AIDS due to impaired immunity and in addition CMV is associated with increased progression to AIDS (Detels et al. 1994; Erice et al. 2003; Deayton et al. 2004; Fielding et al. 2011; Adland et al. 2015a).

CMV is a significant cause of chronic rejection of solid organ transplants (Broers et al. 2000; Castro-Malaspina et al. 2002) and the most common infection in solid organ transplant recipients. Due to the prolonged period of immunosuppression required for organ transplant, transplant recipients are at a great risk of acquiring CMV infection or reactivation. The risk of CMV related disease is most common in a group of CMV seronegative individuals receiving an organ from a seropositive donor (Crough and Khanna 2009).

Although CMV is generally not considered as an onco-accessory, it has been implicated in malignant cancers (Pourghesari et al. 2010; Michaelis et al. 2009). CMV has been shown to be associated with glioblastoma multiforme (GBM)/astrocytoma

grade IV, which is a malignant and lethal brain cancer, and it was implied that CMV might play an oncomodulatory role in GBM (Ranganathan et al. 2012).

CMV is thought to induce systemic immune activation. During primary CMV infection, there is an increase in type 1 cytokines in serum (van de Berg et al. 2010) and the resultant systemic immune activation is associated with several vascular diseases, including atherosclerosis. Recently, a study on aging showed that individuals with CMV IgG antibody titres in the highest quartile had a 1.43-fold greater all-cause mortality, with a higher risk of cardiovascular mortality compared with lower quartiles (Roberts et al. 2010).

CMV seropositivity has been shown to have a negative influence on immune parameters and is associated with 'immune risk phenotype' (IRP) in the elderly. The parameters comprising the IRP include clonal expansion of CD8⁺ T-cells, inverted CD4⁺/CD8⁺ T-cell ratio (<1), an increased proportion of highly differentiated CD8⁺ cells, and elevated levels of pro-inflammatory cytokines (Akbar and Fletcher 2005). The Swedish longitudinal-ageing studies have shown that CMV is associated with 'IRP' (Olsson et al. 2000). The IRP is a predictor of increased mortality in individuals over the age of 80. CMV infection contributes to the significant expansion of CD8⁺ T-cells in elderly. The co-existence of CMV in elderly individuals and its association with IRP suggests that CMV infection may be a potential accelerator in immune system ageing.

1.2.8 CMV life cycle

The life cycle of CMV initiates with a series of distinct steps, including attachment and penetration to human cells via direct fusion or endocytic pathway, gene transcription and DNA replication followed by virion assembly and egress (Compton

2004).

Virally encoded glycoproteins in the lipid bilayer attach to their receptor (epidermal growth factor receptor) on target cells, which triggers fusion and leads to internalisation of both the nucleocapsid and the tegument proteins into the cell (Varnum et al. 2004). The nucleocapsids are translocated to the nucleus and viral DNA is released, which initiates the expression of immediate early genes (IE1 and IE2). The viral gene expression during productive infection occurs in an ordered cascade, starting with immediate-early (IE), followed by delayed-early (E), and late (L) genes. Initial transcription involving the expression of IE genes is not dependent on the translation of other gene products, however delayed-early (E) genes require the expression of functional products of IE gene (transcriptional regulatory proteins). Early genes encode proteins that regulate viral DNA synthesis, while late genes encode proteins for structural components of the virions and permit the assembly and egress of newly formed progeny viral particles (Kalejta 2008). Replicated viral DNA molecules are encapsulated as capsids, which are then transported from nucleus to the cytoplasm, where they undergo a second envelopment in the endoplasmic reticulum (ER)-Golgi intermediate compartment. This is followed by a complex two-stage final envelopment and egress process of newly formed virions (Crough and Khanna 2009; Compton 2004). In addition to infectious virions, dense bodies are also produced from CMV-infected cell. Dense bodies are enveloped tegument proteins that lack capsids, therefore do not contain viral genome (Kalejta 2008).

1.2.9 CMV cell tropism

Primary CMV infection induces a robust innate and adaptive immune response, after which the virus establishes a life-long latency in the host. CMV infects a wide

variety of cells, predominately endothelial cells, epithelial cells (including retinal cells), smooth muscle cells, fibroblasts, dendritic cells and hematopoietic cells. Primary CMV infection in healthy individuals usually initiates with replication in mucosal epithelium and subsequently the virus disseminates to cells of myeloid lineage including CD34+ progenitor bone marrow cells and monocytes where it establishes life-long latency (Söderberg-Nauclér et al. 1997; Sinclair 2008). Infection of these cells along with endothelial cells facilitates systemic spread within the host. The differentiation of these virus-infected CD34+ cells or monocytes to macrophages or dendritic cells can initiate productive infection. Whether a cell permits active viral replication is dependent directly on state of the cell differentiation. Viral replication only takes place in differentiated cells. CMV is able to enter undifferentiated cells but the cells are non-permissive for viral replication, however cells that are not permissive for viral replication do play an important role in the dissemination of the virus in the body (Sinclair and Sissons 2006).

1.3 Innate immune responses to CMV

The innate immune response is important in the defence against CMV and also to prime adaptive immune responses. Antigen presenting cells (APCs) detect virus particles or virus-associated dense bodies via toll-like receptors (TLRs) and activate signal transduction pathways. TLRs are expressed on APCs, including macrophages, dendritic cells (DCs), and B lymphocytes. To date, 10 functional TLRs have been identified in humans. All TLRs consist of extracellular leucine-rich repeats and a cytoplasmic Toll/Interleukin-1 Receptor (TIR) domain. Some TLRs (TLR1, -2, -4, -5, -6, and -10) are expressed at the cell surface and generally recognise lipids, whereas those expressed intracellularly (TLR3, -7, -8, and -9) recognise nucleic acids (Mogensen 2009). CMV virions are detected by TLR2

(Compton 2004). Fig 1.3 shows their cellular localization and origin. Activated TLRs induce the secretion of a range of inflammatory cytokines, and the upregulation of costimulatory molecules such as CD80 and CD86, which are important for the activation of adaptive immunity.

The components of the innate immune system include neutrophils, natural killer cells (NK cells), monocytes, macrophages and DCs. NK cells are an integral part of the innate immune response. Similar to CD8⁺ T-cells, NK cells are able to induce apoptosis by secretion of cytotoxic cytokines such as TNFs or by delivery of Fas ligands to infected cells (Topham and Hewitt 2009). According to the 'missing self' hypothesis (Ljunggren and Karre 1990), NK cells selectively recognise and kill infected cells that abrogate expression of surface major histocompatibility complex (MHC) molecules in order to evade recognition by CD8⁺ T-cells. This recognition of infected cells is mediated by a complex balance of activating and inhibitory receptors on NK cells (Ljunggren and Karre 1990). Recognition of MHC class I molecules by inhibitory receptors on NK cells, suppresses the cytotoxicity activity of NK cells and in contrast recognition by activating receptors induces NK cell mediated cytotoxicity activity.

The role of NK cells in control of human CMV is unclear but, in principle, NK cells should be able to kill CMV infected cells that have downregulated expression of MHC molecules (Crough and Khanna 2009). NK cell activating KIR (killer-cell immunoglobulin-like receptors) genotypes (aKIR2DS2 + aKIR2DS4) of donors have been shown to control CMV reactivation in hematopoietic cell transplant recipients (Zaia et al. 2009). Furthermore, in a study of 43 bone marrow transplant patients with CMV reactivation, the levels of non-specific NK cell cytotoxicity correlated with recovery from CMV infection (Quinnan et al. 1982). Increased NK

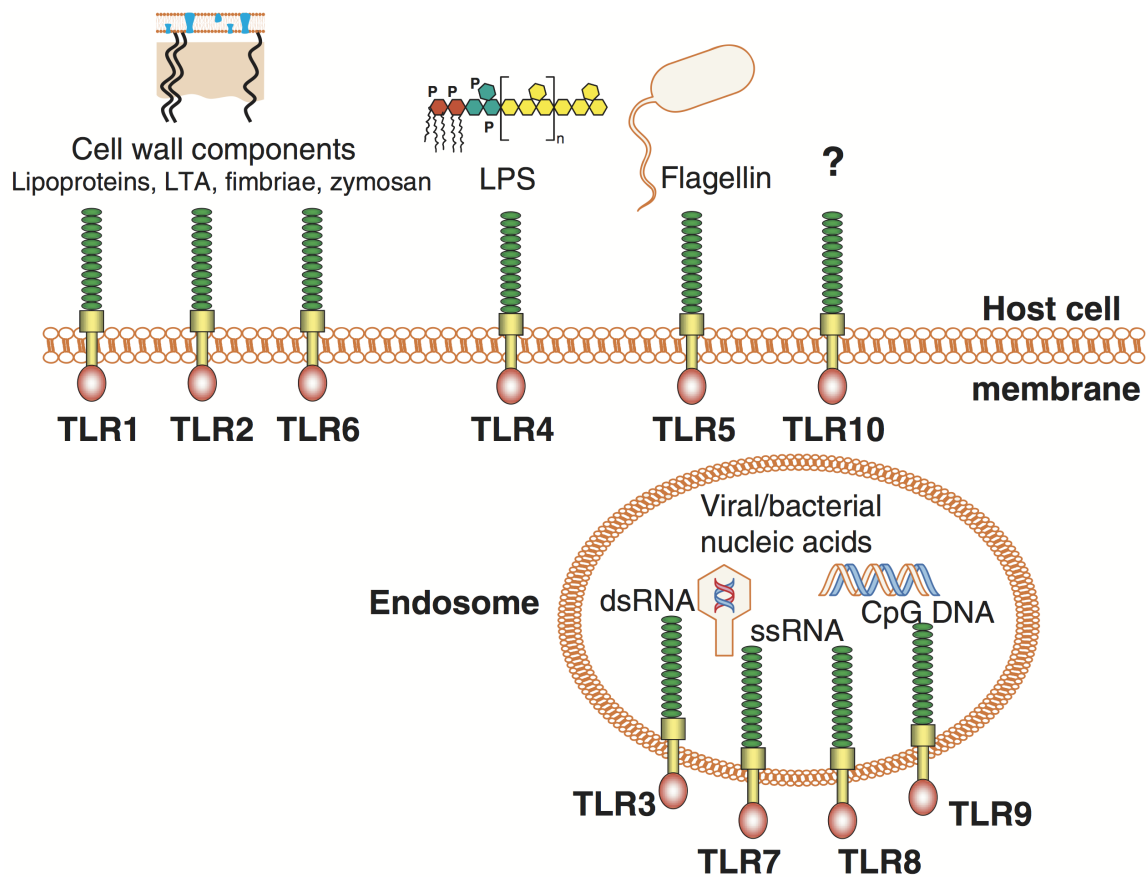


Figure 1.3: Microbial ligand specificities of human Toll-like receptors and their location. TLRs 1, 2, 4, 5 and 6 are expressed on the host cell surface and they recognize extracellular microbial structures. TLR1, TLR2 and TLR6 mostly detect microbial cell wall components, such as lipoproteins, lipoteichoic acid (LTA), fimbriae, or yeast zymosan. TLR4 recognises lipopolysaccharide (LPS) and TLR5 recognises bacterial flagellin, whereas ligand for TLR10 is yet to be identified. TLRs 3, 7, 8 and 9 are expressed intracellularly on endocytic vesicles and they specialise in detecting viral or bacterial nucleic acids. TLR7 and TLR8 recognise single-stranded viral RNA, TLR3 recognizes double-stranded viral RNA and TLR9 detects microbial CpG DNA. (LPS, lipopolysaccharide; LTA, lipoteichoic acid). *Figure from Krauss et al. 2010*

cell activity has also been reported in primary and recurrent CMV infection in renal transplant patients (Venema et al. 1994). Taken together, these studies, along with the fact that CMV employs multiple immune evasion mechanisms (see Section 1.4.14.4) to avoid NK cell activation, indicate an important role of NK cells in controlling CMV infection in humans.

Studies in animal models show a clear role for NK cells in control of murine CMV. (Bukowski et al. 1983). Adoptive transfer of NK cells in MCMV infected mice was shown to control MCMV infection (Bukowski et al. 1985).

1.4 Adaptive immune response to CMV

The adaptive immune system has evolved to deal with pathogens that evade the innate immune responses. The adaptive response displays high specificity by virtue of antigen-specific receptors expressed on the surfaces of T and B lymphocytes. T-cells can be divided into two groups, $\alpha\beta$ -receptor expressing T-cells and $\gamma\delta$ -receptor expressing T-cells. T-cells with the $\gamma\delta$ -TCR have been characterised to a lesser extent.

The $\alpha\beta$ -receptor expressing T-cells can be further sub-divided into two groups based on their cell surface expression, helper T-cells and cytotoxic T-cells, bearing either CD4 or CD8 molecules on their surface, respectively.

1.4.1 Antigen presentation and MHC molecules

TCRs recognise fragments of foreign antigen presented on MHC by antigen presenting cells (Parkin and Cohen 2001). Antigen presentation by MHC class I and class II molecules occurs via two different pathways, but in each case resulting in

the presentation of peptides with a core length of approximately 9 amino acids to CD8⁺ and CD4⁺ T-cells. The open peptide-binding groove of MHC II means that class II presented peptides are longer (Kisselev et al. 1999). Antigenic peptides produced endogenously during an infection with an intracellular pathogen (viral or tumour proteins) are complexed with MHC class I molecules through intracellular processing pathways. Alternatively, exogenous antigens taken up by endocytosis by the specialized professional APCs are complexed with MHC class II molecules via a different pathway and are presented to CD4⁺ T-cells. Professional APCs include dendritic cells, B cells and macrophages. These cells can also present exogenous peptides via the MHC class I pathway, in a process called “cross presentation” (Heath and Carbone 2001). Similarly, endogenous peptides can be presented using the MHC class II molecules when they are degraded through autophagy (Nair-Gupta and Blander 2013; Heath and Carbone 2001). Antigens are presented to the T-cells within the peptide binding groove of an MHC molecule (Fig 1.4).

MHC class I molecules are found on the cell surface of all nucleated cells in the body, and present antigens to CD8⁺ T-cells. In humans, the MHC locus is termed the Human Leukocyte Antigen (HLA) locus (Doherty and Zinkernagel 1975). A series of genes encode the MHC class I: HLA-A, HLA-B and HLA-C genes and the less diverse and less variable HLA-E, -F, -G genes. These genes are highly polymorphic, reflecting the central role of CD8⁺ T-cells in the recognition of pathogens and amelioration of disease (Doherty and Zinkernagel 1975). Differences between HLA molecules have a critical influence on survival, hence the strong selection pressure operating on the HLA gene loci over the course of human evolution.

MHC class I molecules comprise four domains, $\alpha 1$, $\alpha 2$ and $\alpha 3$, with $\beta 2M$ contribut-

ing a fourth domain (Fig 1.4) (Jones 1997). The peptide-binding groove is formed by $\alpha 1$ and $\alpha 2$ domains consisting of a sheet of eight β -pleated strands with two flanking walls constituted by an α -helix and typically binds peptides of 8-11 amino acids long. The transmembrane $\alpha 3$ domain together with $\beta 2$ -microglobulin supports the peptide binding site. MHC class I molecules possess pockets within the peptide-binding groove, which preferentially bind particular amino acid residues, usually at position 2 and at the C terminus of the peptide. The peptide-binding groove forms six A-F pockets, although not every pocket is necessarily occupied, depending on different HLA types. Due to different pockets in MHC I molecules, they display differential preference to amino acids binding (Falk et al. 1991).

The polymorphic heavy chain (α chain) of MHC class-I molecule is stabilized by the chaperone calnexin, prior to association with the $\beta 2$ -microglobulin. Endoplasmic reticulum resident chaperone proteins (calreticulin, Erp57, protein disulfide isomerase (PDI) and tapasin) facilitate the folding of MHC class-I molecules. Tapasin is critical for linking the class-I/chaperone complex to TAP (transporter associated with antigen presentation). Peptides generated in the cytosol by proteasomes are translocated into the ER by TAP (Abele and Tampé 1999). These translocated peptides may require further trimming in the ER before binding to the newly synthesized MHC class I molecules, which is mediated by ER aminopeptidase (ERAP)(Reits et al. 2003; Rock et al. 1994). MHC molecules dissociate from TAP after peptide loading. The peptide-MHC class I complex is then transported to the plasma membrane via the Golgi apparatus (Spiliotis et al. 2000) so that it can be presented to CD8+ T-cells, triggering CD8+ T-cell responses (Fig 1.5). In some cases, MHC class I molecules never bind peptides and have to be degraded by the ER-associated protein degradation (ERAD) system and sometimes peptides are not able to associate with MHC class I and they have to be returned to the cytosol

for degradation.

It should be noted that 30 to 70% of peptides presented by MHC class I molecules are derived from defective ribosomal translation products (DRiPs), which are degraded by the multi-subunit proteasome complex in the cytosol due to defective transcription and translation (Schubert et al. 2000). This process allows the MHC class I pathway to sample proteins immediately after synthesis, thereby allowing the viral peptides to be sampled and presented very quickly in an infected cell.

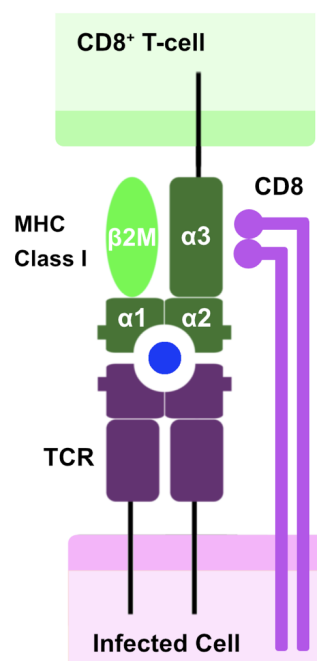


Figure 1.4: Schematic presentation of MHC class I molecule. MHC class I molecules are heterodimers made up of α (heavy) and β 2-microglobulin (β 2m; light) chains. The α chain consists of three polypeptide domains (α 1, α 2, α 3) and they are noncovalently linked with β 2-microglobulin. The α 3 domain is plasma membrane-spanning and interacts with the CD8 co-receptor of T-cells. The α 1 and α 2 domains fold to make up a peptide groove, in which the antigen peptide is presented. The α 3-CD8 interaction holds the MHC I molecule in place while the T-cell receptor on the CD8⁺ T-cells binds the peptide presented in the binding cleft.

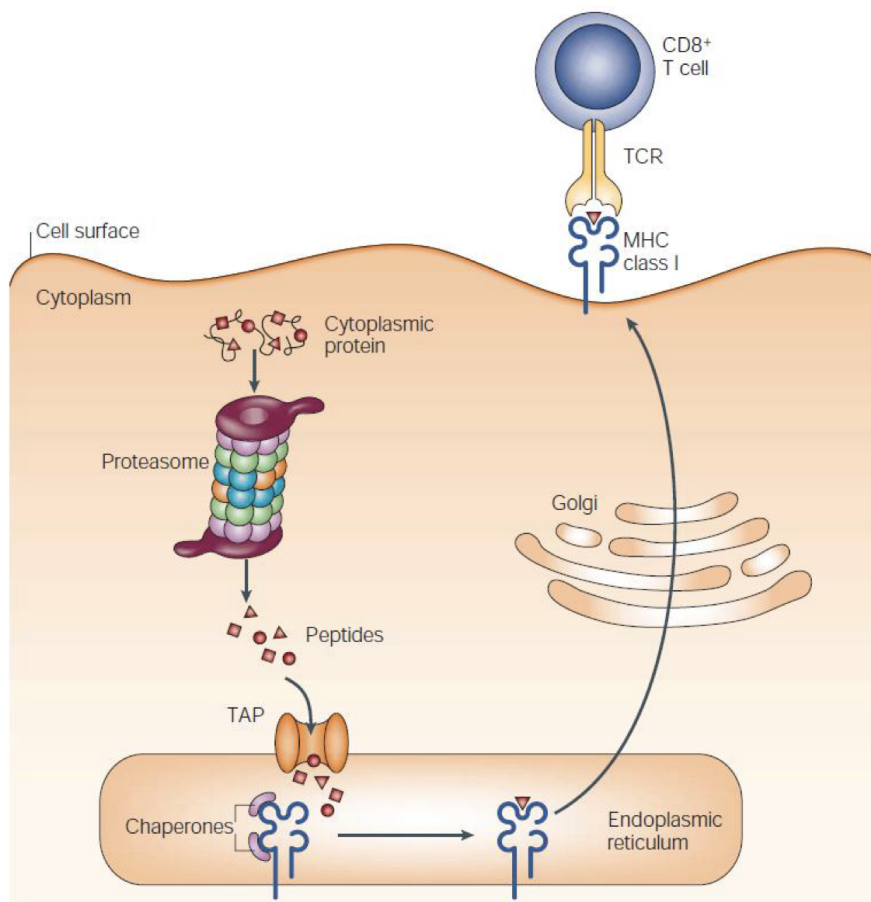


Figure 1.5: Schematic representation of MHC I antigen processing. Proteins of intracellular pathogen are digested by the proteasome into peptides. The transporter for antigen processing (TAP) then translocates them into the the lumen of the endoplasmic reticulum (ER). Peptide is loaded onto MHC class I molecule present in the ER, with the assistance of chaperone molecule. Subsequently, the MHC-peptide complex is translocated via the Golgi apparatus to the cell surface and presented to the CD8⁺ T-cells. *Figure from Yewdell et al. 2003*

1.4.2 Structure and assembly of TCR

The TCR is a heterodimer of one α and one β -chain, or one γ and one δ chain, which are disulphide-linked. Most T-cells express α/β -TCRs (approximately 95% of human T-cells express α/β -TCRs and 5% express γ/δ -TCRs). Each TCR chain is composed of a large extracellular domain, made up of a constant and a variable domain, followed by a membrane-spanning region and a short cytosolic tail (Fig 1.6A) (Kappler et al. 1983; Chothia et al. 1988; Raulet 1989). The β -chain com-

prises of variable (V), diversity (D), joining (J), and constant (C) gene segments whilst the α -chain comprises of variable (V), joining (J), and constant (C) gene segments.

A complete TCR RNA transcript is divided into a leader sequence, a framework region (FR) and a complementarity-determining region (CDR). Typically, four FRs and three CDRs are juxtaposed in the sequence of FR1, CDR1, FR2, CDR2, FR3, CDR3 and FR4 to form a TCR (Lefranc et al. 2003). The primary site of antigen contact is thought to be the CDR regions and more specifically the CDR3 regions (Fig 1.6B). The α and β chains pair soon after their biosynthesis to yield the $\alpha\beta$ -TCR heterodimer.

The production of antigen-specific receptors in T-cells is the result of a process of random gene rearrangement (Hozumi and Tonegawa 1976) and splicing together of multiple DNA segments that code for the antigen-binding areas of the receptors (CDR). Gene rearrangement is facilitated by recombination-activating genes (RAG1 and RAG2 genes). TCR chains are assembled somatically during T-cell development (discussed later) before exposure to antigen, which enables a generation of a highly diverse receptor repertoire.

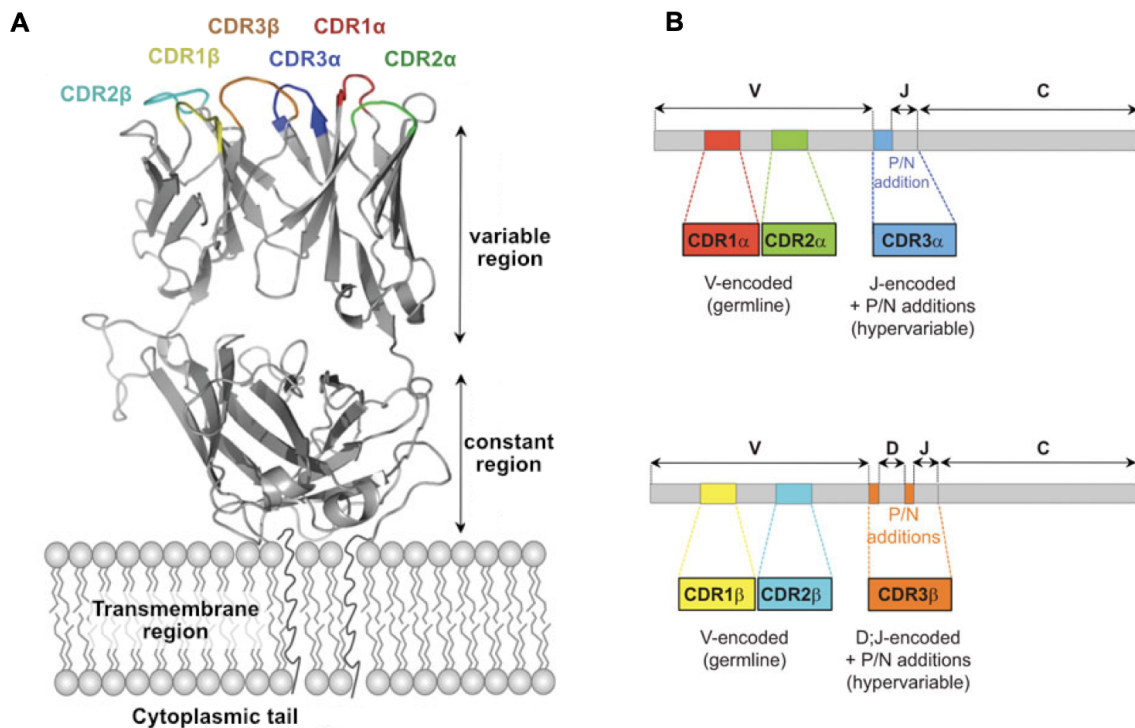


Figure 1.6: TCR protein and gene structure. **(A)** The TCR is a disulfide-bound heterodimer composed either of α and β or γ and δ , here showing alpha and beta chain TCR. Each chain has a variable region and a constant region (designated V_α and C_α , V_β and C_β etc.). Diversity within TCR molecules is mainly concentrated at the short hairpin loops, complementarity determining regions (CDRs, colored) of the α and β chains which also form the antigen binding cleft. **(B)** TCR chains are assembled somatically during T-cell development by joining of V, (D) and J gene segments. CDR1 and CDR2 are entirely encoded in the germline V genes, whereas CDR3 lies at the junction between the rearranged V and J segments (TCR- α) and V, D and J segments (TCR- β). The CDR3 junctional sites are assembled by random addition and deletion of template and non-template nucleotides (blue for TCR- α and orange for TCR- β). CDR, complementarity-determining region; TCR, T-cell receptor. *Figure from Attaf et al. 2015*

1.4.3 T-cell development

The T-cells originate from lymphoid progenitor cells in the bone marrow and migrate to the thymus, through a process known as haematopoiesis (Fig 1.7). A series of V(D)J recombination events take place in the thymus and T-cells develop their specific markers, including TCR, CD3, CD4 or CD8 (Germain 2002; Pui et al.

1999).

Committed T-cell precursors lack TCR, CD4 and CD8 and are termed double negatives (DN; no CD4 or CD8). DN thymocytes differentiate in four sequential stages, which are identified by their surface expression of CD44 (adhesion molecule) and CD25 (interleukin 2 α -receptor): DN1, CD44+CD25-; DN2, CD44+CD25+; DN3, CD44-CD25+; and DN4, CD44-CD25- (Godfrey et al. 1993). Double negative T-cells give rise to either of $\alpha\beta$ T-cells or $\gamma\delta$ T-cells, which have different types of T-cell receptors (van Oers et al. 1995). Cells that proceed along the $\alpha\beta$ -TCR pathway, can develop into two distinct functional subsets, CD4 and CD8 T-cells.

CD44+ CD25- cells at DN3 stage, along the $\alpha\beta$ -TCR pathway, undergo a process termed β -selection. The β -chain of the TCR is rearranged and upon successful rearrangement, it is paired with a preTCR α -chain, allowing the assembly of a complete preTCR (Huang et al. 2005). At the cell surface, the pre-TCR- $\alpha\beta$ pair is associated with CD3 molecules to provide a complete CD3:TCR receptor complex that initiates signal transduction which halts any further β -chain rearrangement (van Oers et al. 1995). This signalling is also critical for T-cell proliferation and the expression of CD4 or CD8 molecules (double positive thymocytes). After successful β -chain selection and subsequent cell proliferation, the pre TCR- α locus is replaced with a newly rearranged TCR- α chain, which yields a complete $\alpha\beta$ -TCR (Huang et al. 2005).

These $\alpha\beta$ -TCR+ CD4+ CD8+ double positive (DP) thymocytes then undergo thymic selection which usually permits only one in a hundred thymocytes to enter the periphery (Goldrath and Bevan 1999). The factors that control this selection process include the components of the antigen processing machinery (Attaf et al. 2015).

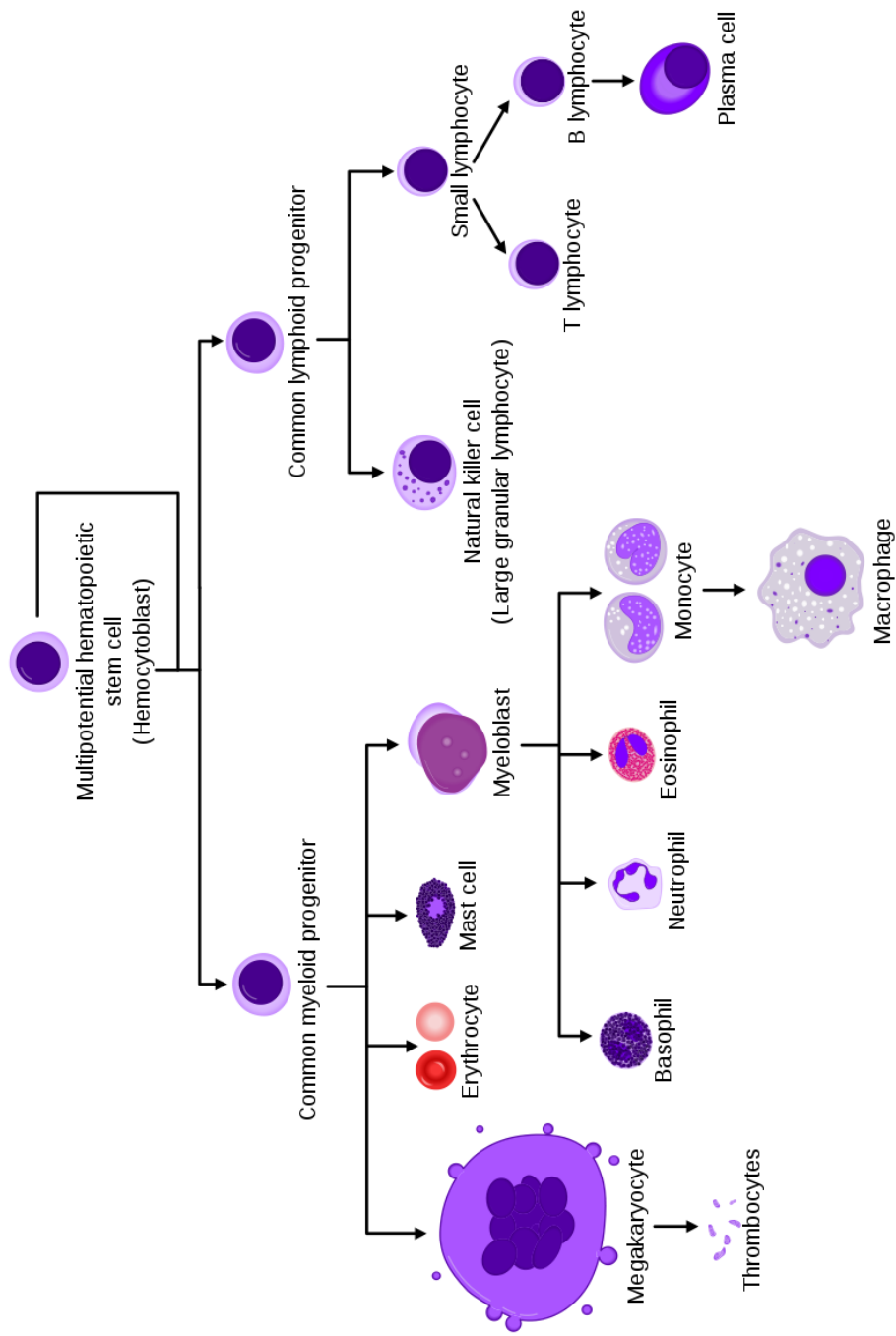


Figure 1.7: Hematopoietic stem cells in the bone marrow gives rise to lymphoid and myeloid progenitors. The pluripotent hematopoietic stem cells divide to produce two specialized types of stem cells, a common lymphoid progenitor and a common myeloid progenitor. The common lymphoid progenitor gives rise to the T and B lymphocytes or natural killer cells. The common myeloid progenitor gives rise to different types of leukocytes (white blood cells), erythrocytes (red blood cells that carry oxygen), and the megakaryocytes that produce platelets that are important in blood clotting. *Figure from Mikael Häggström, used with permission*

MHC polymorphism (Messaoudi et al. 2002; Dyllal et al. 2000), ER-associated aminopeptidase 1 (ERAP1) polymorphism (Reeves et al. 2013), thymoproteasome (Nitta et al. 2010) are all components of antigen processing machinery and determine the peptide diversity generated in the thymus. Peptide diversity in thymus regulates the diversity of post-selection thymocyte population (Bevan 1997).

Thymocytes are selected for maturation and migration into peripheral lymphoid tissues, through interaction with thymic epithelial cells expressing a wide variety of self-peptide-MHC (Robey and Fowlkes 1994; von Boehmer et al. 1989). The majority (90%) of the generated thymocytes express TCR that bind with self-peptide-MHC so weakly that the signal required to sustain viability is not generated, hence causing their 'death by neglect'. A small proportion of thymocytes bind self-peptide-MHC so strongly as to pose an autoimmune risk if they were permitted to leave the thymus. These thymocytes undergo rapid apoptotic death (negative selection). Thymocytes that express TCR that recognise the self-peptide-MHC molecule with appropriate intensity (intensity between those resulting in neglect or negative selection) are selected (positive selection). DP cells that pass through thymic selection, ultimately differentiate into either CD4⁺ or CD8⁺ T-cells and leave to enter the periphery (Singer et al. 2008). This system ensures a production T-cells with a wide array of TCRs that are capable of recognition and activation in response to diverse antigens.

1.4.4 V(D)J recombination

TCR genes, unlike the majority of other genes, are not germline encoded, instead the loci consist of discontinuous sections of DNA that have been recombined together to form a complete coding sequence. The genes encoding the $\alpha\beta$ -TCR are generated by somatic recombination from several non-contiguous gene segments

(Fig 1.8).

Recombination of gene segments can be initiated by recognition of conserved recombination signal sequences (RSS) at the gene boundaries. Gene-segment joining is carried out by the recombinase complex, the product of two recombinase activating genes, RAG-1 and RAG-2, which work co-operatively to mediate recombination between RSS (Oettinger et al. 1990; Fugmann et al. 2000; Gellert 2002). Gene segments of the α and β -chains are flanked by RSS. Each RSS is composed of a conserved nonamer and heptamer nucleotide sequence, separated by a space of either 12 or 23 nucleotides (Tonegawa 1983). Recombination only occurs efficiently between different length spacers and this promotes correct order of assembly, for example, recombination occurs only between 12-RSS and a 23-RSS, known as the “12/23 rule” (Hiom and Gellert 1998). The distribution of RSS in TCR- α chain directs recombination between the J-V gene segment. The distribution of RSS in the TCR- β chain first directs recombination between 3' D-23RSS to J-12RSS and then V-23RSS with 5' D-12RSS gene, avoiding V-23RSS to J-12RSS recombination. The mechanism of avoiding V-23RSS to J-12RSS recombination is referred to as the “beyond 12/23 rule” (Olaru et al. 2005) and this ensures the inclusion of a D- β gene segment in the assembled TCR- β chain. Once the two RSSs are cleaved together the intervening segment is cleaved out and the coding ends are ligated by ubiquitously expressed non-homologous DNA end-joining (NHEJ) proteins (Malu et al. 2012) which also results in addition or deletion of templated nucleotides or addition of non-templated nucleotides.

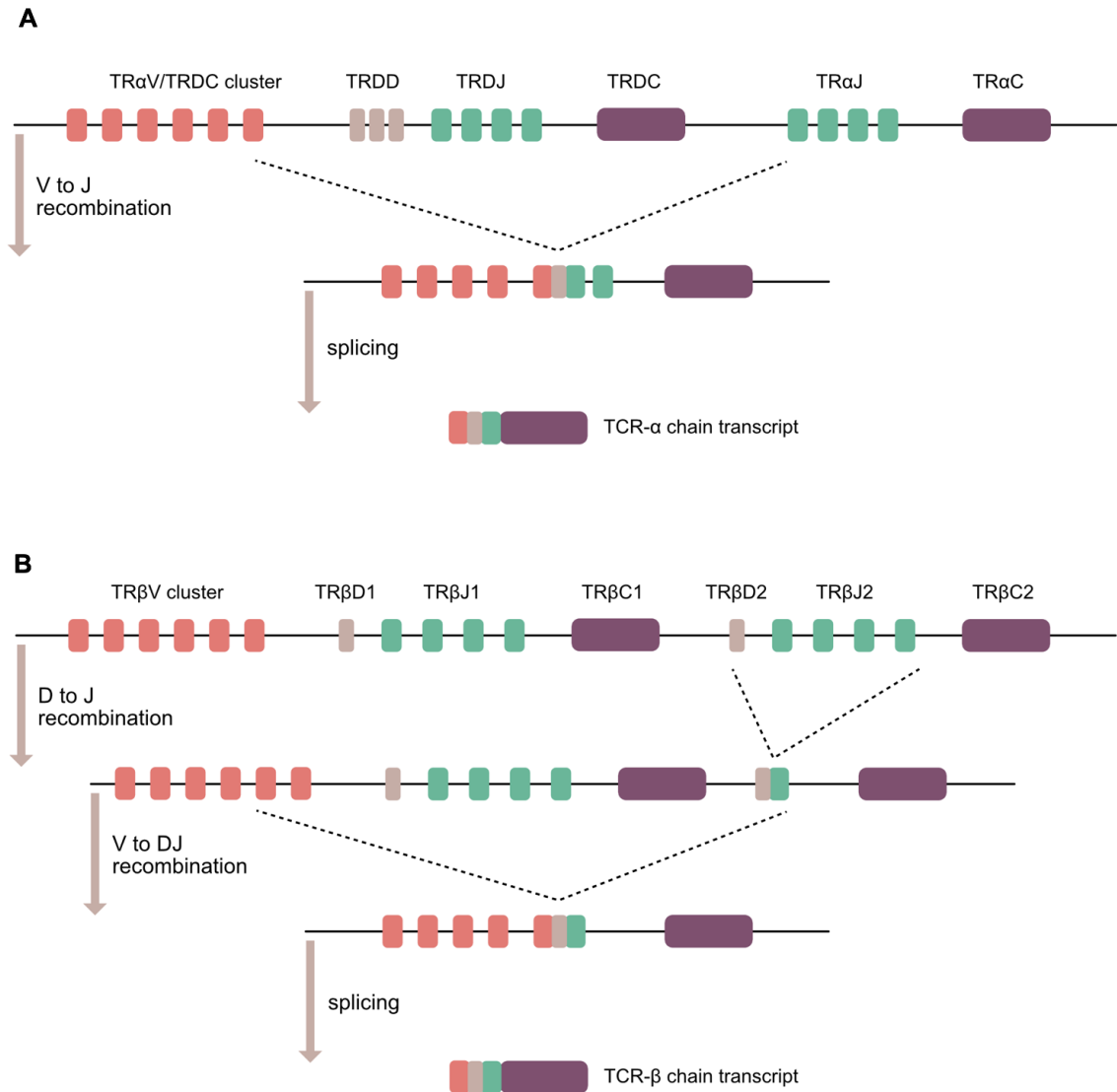


Figure 1.8: Overview of V(D)J recombination of TCR- α and - β chains. (A) VJ recombination at the tra locus. The tra locus also has TCR- δ , or trd locus embedded in it. It comprises a 5' V gene segment cluster (46 TRAV segments) followed by a central J cluster (51 TRAJ segments) and a single C gene segment (TRAC). TCR- δ D and J segments (TRDD and TRDJ, respectively) present are also shown. V to J recombination brings together one of many TRAV segments to one of many TRAJ segments, splicing out the intervening sequences, producing a TCR- α transcript in which V, J and C segments are directly adjacent. **(B)** VDJ recombination at the trb locus. The trb locus is composed of a 5' V cluster (48 TRBV gene segments) followed by two 3' TRBD-TRBJ-TRBC clusters. VDJ recombination is a two-step, ordered process. Firstly, D to J recombination occurs, where TRBD1 is brought to join one of the six TRBJ1 segments or TRBD2 to one of the seven TRBJ2 segments. V to DJ recombination subsequently brings the rearranged DJ join to one of many TRBV segments. The intervening sequences are then spliced out which generates a TCR- β transcript with V, D, J and C segments directly adjacent to each other. *Figure adapted from Attaf et al. 2015*

1.4.5 TCR receptor diversity

V(D)J recombination gives rise to diversification of the CDR region in TCRs (Fig 1.8) (Market and Papavasiliou 2003), which is mainly generated by virtue of there being multiple V, D and J genes. The theoretical diversity of human TCR repertoire is estimated at 10^{18} clonotypes (Sewell 2012); most of these specificities are never used and it is estimated that human TCR repertoire is composed of 25 million clonotypes (Arstila et al. 1999).

The α -locus comprises a 5' V gene segment cluster (46 TRAV segments) followed by a central J cluster (51 TRAJ segments) and a single C gene segment (TRAC) (Lefranc et al. 2003). V to J recombination in the α -chain brings together one of the 46 TRAV segments to one of the 51 TRAJ segments, splicing out the intervening segments, producing a TCR- α transcript in which V, J and C segments are adjacent.

The β -locus is composed of a 5' V gene segment cluster (48 TRBV segments) followed by two 3' TRBD-TRBJ-TRBC clusters. The human TCR β -locus contains 48 V, 2 D, and 13 J segments (Lefranc et al. 2003). VDJ recombination in β -chain is an ordered two-step process. D to J recombination in the β -chain occurs first whereby TRBD1 is juxtaposed to one of the six TRBJ1 segments or TRBD2 to one of the seven TRBJ2 segments. V to DJ recombination then brings the rearranged DJ to join one of the 48 TRBV segments, splicing out any intervening segments to produce a TCR- β transcript in which, V, D, J and C segments are directly adjacent (Attaf et al. 2015).

Although V(D)J recombination is a major factor in TCR diversity, diversity of TCR is significantly increased by random addition and deletion of nucleotides. There are

two types of nucleotide additions, template and non-templated. Non-templated nucleotides (N) are added by the enzyme terminal deoxynucleotidyl transferase (TdT) (Desiderio et al. 1984; Repasky et al. 2004). TdT is able to add deoxynucleotides without a template, with a preference for guanine residues that results in N regions being generally GC-rich (Gilfillan et al. 1995). These insertions are relatively specific to V(D)J recombination as the expression of TdT is usually limited to early lymphoid cells where V(D)J recombination is active (Davis and Bjorkman 1988). Addition of template nucleotides is also found in the coding joint. The ends of the coding DNA are converted to hairpins during recombination. These hairpins can be nicked by a few bases, due to exonucleolytic activity of DNA repair enzymes, which can lead to formation of 'P-nucleotides' at the recombining edges. These P nucleotide insertions (P for palindromic) can be incorporated in the junction (Gellert 2002). Very little is known about the nucleotide deletion in the coding junction.

1.4.6 CD8+ T-cells

1.4.6.1 CD8+ T-cell differentiation

CD8+ T-cells that have not yet encountered their cognate antigen and have not been activated are called antigen naive CD8+ T-cells. These T-cells leave the thymus and circulate through blood and the lymphatic system providing immune surveillance, and reside in secondary lymphoid organs (lymph node and spleen) (Spits 2002; Boyman et al. 2009). APCs acquire antigens in non-lymphoid tissue throughout the body and migrate into secondary lymphoid organs where they present antigens to CD8+ T-cells. When a naive CD8+ T-cell recognises its specific antigen, presented by a TCR in context of MHC molecule on an APC, it ceases migrating and undergoes extensive clonal expansion (Kaech et al. 2002; Spits 2002). Interaction of the TCR with the cognate antigen triggers intracellu-

lar signalling events that leads to activation and proliferation of these naive CD8+ T-cells (CD45RA+, CD27+, CD28+, CCR7+) into effector CD8+ T-cells. Short-lived effector CD8+ T-cells migrate to a site of infection and carry out their effector functions, such as cell-mediated cytotoxicity and production of cytokines such as IFN-gamma. The majority of the effector CD8+ T-cells will die after the antigen is cleared and the remaining will become memory CD8+ T-cells which serve to protect the body against the same antigen in the future. These long-lived memory CD8+ T-cells are able to mount a faster response than their naive counterparts and therefore are able to clear the pathogen faster. Memory CD8+ T-cells consist of two distinct subpopulations, central memory CD8+ T-cells (T_{CM}) and effector memory CD8+ T-cells (T_{EM}). T_{CM} and T_{EM} can be distinguished based on their homing capacity and effector functions (Sallusto et al. 2004; Willinger et al. 2005).

CD8+ T_{CM} are CD45RA-/CCR7+ memory cells that constitutively express two lymph node-homing receptors (CCR7 and CD62L) and express co-stimulatory molecules CD27 and CD28. Lymph node-homing receptors are required for migration of T-cells to areas of secondary lymphoid organs (Sallusto et al. 1999). CD8+ T_{CM} are able to differentiate into T_{EM} following TCR triggering. CD8+ T_{EM} are CD45RA-/CCR7- CD27+ CD28-/+, characterized by their rapid effector function and loss of CCR7 expression, are heterogeneous for CD62L expression, and display characteristic sets of chemokine receptors and adhesion molecules that allow them to gain access to inflamed peripheral tissues (Sallusto et al. 1999). In humans, there is a third subset of T-cell memory compartment, effector memory RA (T_{EMRA}), which includes cells that re-express CD45RA+ but lack expression of CCR7 and are CD27- CD28-. T_{EMRA} are highly differentiated cells that have undergone extensive cell division. They display potent effector function but have little proliferative capacity and short telomeres (Appay et al. 2008). Markers that distinguish be-

tween naive, effector, central memory, effector memory and T_{EMRA} cells are shown in Fig 1.9 (Appay et al. 2008).

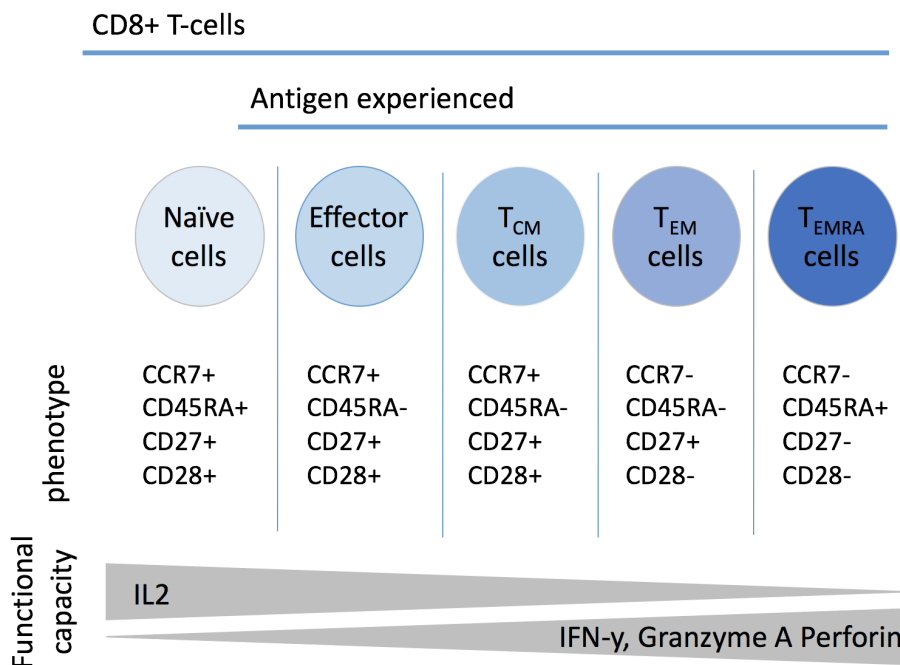


Figure 1.9: The expression of surface markers associated with differentiation is presented for CD8+ T-cells. Five distinct subsets of circulating CD8 T-cells are defined based on expression of CD27, CD28, CCR7, and CD45RA. As the CD8+ T-cells become antigen experienced, they express different cell surface markers and have different functional capacity. CM: Central memory; EM: Effector memory. *Figure adapted from Appay et al. 2008.*

1.4.6.2 Effector mechanisms of CD8+ T-cells

The CD8+ T-cells, often called cytotoxic T lymphocytes (CTL), are very important in immune defence against intracellular pathogens, including viruses and bacteria, and are also vital in the destruction of tumour cells. Naive CD8+ T-cells are activated as a result of two distinct signals: (i) binding of TCR to peptide/MHC complex on APC and (ii) signalling of a variety of co-stimulatory molecules expressed on T-cells (e.g. CD28 to its ligand CD80 or CD86 on APC) (Andersen et al. 2006). Upon activation, naive CD8+ T-cells undergo clonal expansion in the presence of

Interleukin 2 (IL-2) and differentiate into effector cells (Broere et al. 2011). The activated CD8⁺ T-cells can kill infected or malignant cells by direct cell-to-cell contact or by production of cytokine. The CD8⁺ T-cells primarily produce TNF- α and IFN- γ , which have anti-tumour and anti-viral effects. TNF- α is able to engage its receptor on the target cells and initiate caspase cascade, leading to targeted cell death. IFN- γ , on the other hand, induces transcriptional activation of the MHC class I antigen presentation pathway and Fas in target cells. This leads to upregulation of antigen peptides by MHC class I and increased Fas-mediated target-cell lysis (Andersen et al. 2006).

Although CD8⁺ T-cells are capable of producing cytokines, their main function appears to be largely driven towards elimination of infected cells by cytotoxic means. This is most commonly accomplished by delivery of cytotoxic granule toxins into the cytosol of the infected cell by a CD8⁺ T-cell. Upon contact with target cell, the CD8⁺ T-cells form an immunological synapse and cytotoxic secretory granules are released into the synaptic cleft. The cytotoxic secretory granules include pore forming proteins, perforins, and other structurally related serine proteases, Granzyme A and B. These toxic proteins diffuse into the cell membrane of infected cells in a Ca²⁺ dependent manner. These proteins are highly cytotoxic and induce apoptosis in target cells through caspase-dependent and -independent routes (Trapani and Smyth 2002). This method of killing by CD8⁺ T-cells is highly controlled, ensuring killing of only the target cells in a rapid and efficient manner.

The other mechanism of cell-to-cell killing is mediated by Fas ligands, which are expressed on the surface of CD8⁺ T-cells, and which bind to Fas receptors (Fas CD95). This binding of Fas ligands to Fas receptors in target cells triggers apoptosis of target cells via the classical caspase cascade (Waring and Müllbacher 1999).

Rapid expansion of CD8⁺ T-cells and the ability of a single CD8⁺ T-cell to destroy many virally infected cells, while sparing “innocent” bystanders, make them very efficient antigen-specific effector cells (Broere et al. 2011).

1.4.7 CD8⁺ T-cell responses to CMV

CD8⁺ T-cells are thought to be the principal mediator of protection against CMV. The importance of CD8⁺ T-cell responses to CMV comes from observations in animal models in which the elimination of CD8⁺ T-cells correlated with increased levels of reactivation and dissemination of viral infection (Reddehase et al. 1985; Barry et al. 2007).

Studies in humans have also established that deficiency of CMV-specific CD8⁺ T-cell responses correlate with CMV replication and CMV disease, suggesting an important role for the CD8⁺ T-cell in control of CMV infection (Sester et al. 2002). IFN- γ -producing CD8⁺ T-cells in patients with AIDS appear to protect against CMV associated retinitis (Jacobson et al. 2004). Clinical data from bone marrow transplant patients showed that development of CMV-specific CD8⁺ T-cell responses protect them from CMV associated pneumonia (Reusser et al. 2015) and other CMV-related disease (Li et al. 1994). In another study of bone marrow transplant recipients, the presence of CD8⁺ T-cells correlated with recovery from CMV infection (Quinnan et al. 1982). Furthermore, two pivotal studies in bone marrow transplants patients by (Riddell et al. 1992) and (Walter et al. 1995) showed that transferring donor derived CMV-specific CD8⁺ T-cells into recipients effectively restored antigen specific CD8⁺ T-cell immunity and prevented CMV viraemia and CMV disease. A similar important role of CD8⁺ T-cell immunity has been shown in solid organ transplant recipient, where presence of CD8⁺ T-cells correlates with

viral control and mediates protection from CMV disease (Reusser et al. 1999).

1.4.8 Proportion/size of CD8+ T-cells responses to CMV

The magnitude of CD8+ T-cell responses to CMV is unusually high, constituting 10% of the CD8+ T-cell memory compartment (Sylwester et al. 2005). The proportion of CD8+ T-cell increases with age and CMV-specific CD8+ T-cell responses are higher in the elderly, reaching up to 40% (Khan et al. 2004). It is not understood what makes CMV so immunogenic and what impact this has on the host's response to other pathogens. Recent data using overlapping 15-mer peptides from open reading frame (ORF) covering the whole proteome revealed that CMV-specific CD8+ T-cell responses are diverse. Out of the 213 CMV ORFs assessed, 107 (50%) elicited at least one CD8+ T-cell response (Sylwester et al. 2005). This and other studies have shown that CMV-specific CD8+ T-cell responses are highly diverse, targeting a variety of proteins expressed at different stages of viral replication (IE, E, late), with diverse function such as structural, immune evasive and regulatory (Elkington et al. 2003; Sylwester et al. 2005). There is an interesting hierarchy of immunogenicity in CMV encoded proteins: UL48, UL83 (pp65), and UL123 (IE1) have been shown to be the three most immunogenic CMV proteins in a study by Sylwester et al. et al. (2005), recognised by more than 50% of the subjects studied in the cohort. Fig 1.10 shows the schematic representation of the hierarchy of the most frequently recognised CMV antigens.

1.4.9 Course of CMV infection

During an acute infection with CMV, naive T-cells recognise CMV epitopes, and undergo proliferative expansion (Fig 1.11). Typically, during viral infections, the establishment of a memory T-cell population in response to an antigen is thought

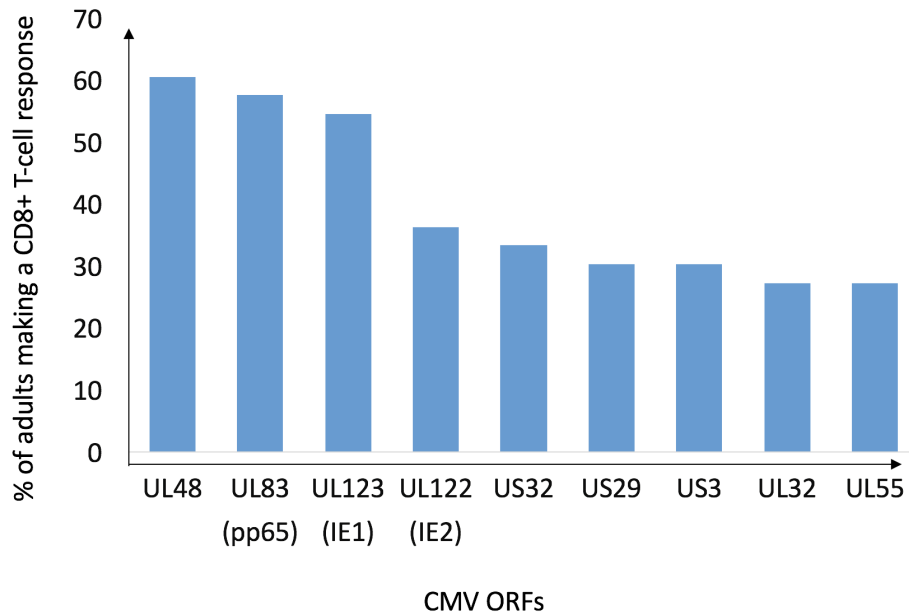


Figure 1.10: Frequency of CMV ORF recognition by CD8+ T-cells. The graph shows the percentage of CMV seropositive individuals (n=33) making a CD8+ T-cell response to the 10 most frequently recognised CMV ORF. *Source: Sylwester et al. 2005*

to involve expansion followed by contraction into a stable pool of memory cells, after the infection is cleared. However, in CMV infection, the CD8+ T-cells specific for certain epitopes do not contract but are maintained at high frequencies (O’Hara et al. 2012). This increase in CMV-specific CD8+ T-cell response is called “memory inflation”. “Memory inflation” is a term derived from studies of murine cytomegalovirus (MCMV) that detected unusually high numbers of MCMV-specific CD8+ T-cells and maintenance of these CD8+ T-cells at high frequencies, in immune competent mice (Karrer et al. 2003). It is now a well described phenomenon in CMV positive humans (Klenerman and Oxenius 2016). Fig 1.11 shows the development of memory CD8+ T-cell responses during the course of CMV infection. The course of CMV infection involves clonal expansion of CMV-specific CD8+ T-cells followed by contraction and subsequent maintenance of CMV-specific memory CD8+ T-cells. Despite the persistent nature of CMV infection, inflationary T-cells lack features of exhaustion and they maintain their functionality. The fact that

CMV can induce high CD8+ T-cell responses can be exploited therapeutically. Relatively high transmission of CMV together with the high inflationary CD8+ T-cell response has led to consideration of the development of vaccines using CMV as a vector (Karrer et al. 2004; Hansen et al. 2009; O'Hara et al. 2012).

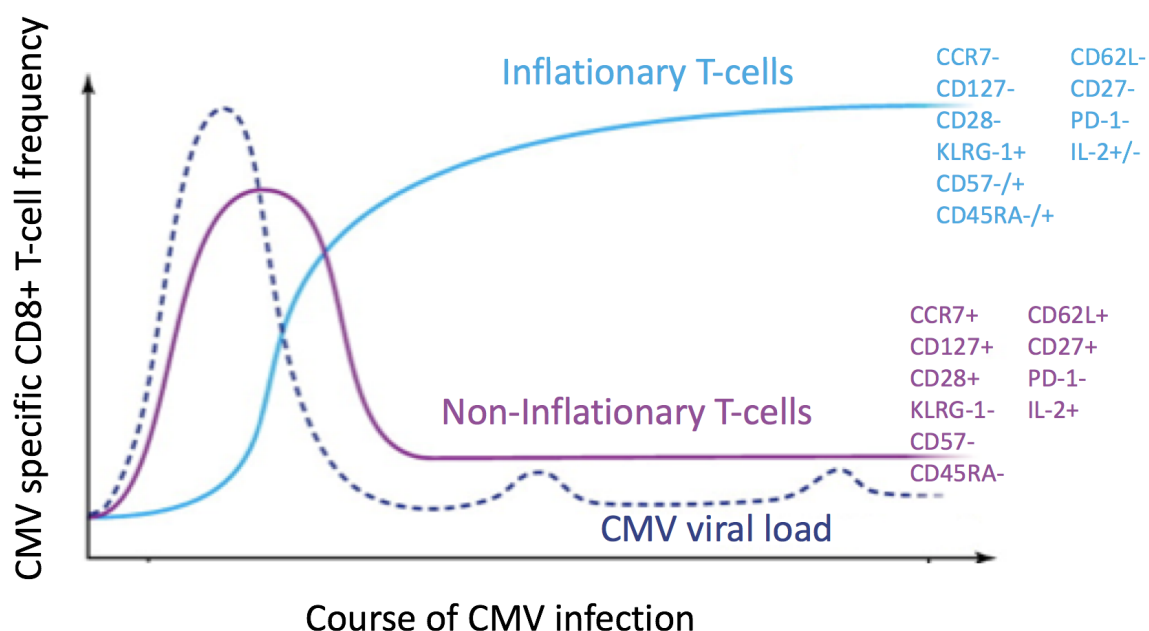


Figure 1.11: Course of CMV infection. During CMV infection, non-inflationary T-cells undergo expansion, contraction and formation of stable memory pools. By contrast, inflationary T-cell pools are characterized by lack of contraction and gradually accumulate in frequency over time aided by low levels of viral reactivation. Characteristic phenotypic features of inflationary versus non-inflationary T-cells are shown. *Figure adapted from O'Hara et al. 2012*

1.4.10 Phenotype of CMV-specific CD8+ T-cells

Typically, during an acute CMV infection, CMV-specific CD8+ T-cells exhibit a classical central memory response, displaying a CD45RA- CD45RO+ CD27+ CD28+ CCR7+ phenotype (Fig 1.11). Following resolution of CMV viraemia, CMV-specific CD8+ T-cells are maintained (inflationary T-cells) and they display an effector memory phenotype (CD45RA-/+ CD27-, CD28-, CCR7-). Consistent with this pheno-

type, they are present at higher quantities in blood and tissues (such as the liver and lungs), compared with lymph nodes. These cells are characterised by gradual downregulation in expression of co receptors CD27 and CD28 and upregulation in expression of CD57 and KLRG1 (killer cell lectin-like receptor subfamily G member 1) and effector molecules (such as perforin and granzyme). Together with this, effector memory CMV-specific CD8⁺ T-cells are also deficient in expression of CD62L (lymph node homing receptor) and CD127 (the IL-7-receptor α -chain).

Another defining feature of these inflationary CMV-specific CD8⁺ T-cells is the reversion of expression of CD45RA (which is associated with naive T-cells) therefore being known as “revertant” memory cells (T_{EMRA}) (Appay et al. 2002; Moss and Khan 2004; Rosa and Diamond 2012). CMV-specific CD8⁺ T-cells do not show features of exhaustion, and retain their proliferative and cytotoxic capacity, including the ability to secrete cytokines (Hertoghs et al. 2010; O’Hara et al. 2012; Klenerman and Oxenius 2016). Furthermore, they display low PD-1 expression (programme death receptor) which is usually found on differentiated cells and serves as a negative regulator of T-cell activation (Hertoghs et al. 2010; O’Hara et al. 2012). A common transcriptional signature found in human CMV-specific T-cells seems to be driven by a small number of transcriptional regulators including T-bet, EOMES and BLIMP1 (Hertoghs et al. 2010; O’Hara et al. 2012).

1.4.11 Factors involved in memory inflation

1.4.11.1 Antigen exposure

The underlying mechanisms of memory T-cell inflation are not yet fully elucidated. The major driving force of this phenomenon must relate to repetitive antigen exposure (Klenerman and Hill 2005). This notion that CMV-specific memory T-cell infla-

tion is caused by repetitive antigen exposure has been investigated using murine models. It was demonstrated that adoptive transfer of MCMV-driven inflationary T-cells into naive recipients resulted in their failure to divide and survive, whereas the same cells survived in chronically infected mice, indicating that antigen stimulation is required for memory inflation (Snyder et al. 2008; Kim et al. 2015). Furthermore, the dose of viral inoculum seems to influence the magnitude of memory inflation in CMV infection. Recently, a study conducted using low dose of MCMV infection severely hampered the accumulation of inflationary memory T-cells (Redeker et al. 2014).

1.4.11.2 Naive and memory T-cells

It has been reported that inflationary T-cells are functional, short-lived cells with a half-life of approximately 40-50 days (O'Hara et al. 2012). Naive T-cells contribute to maintenance of the inflationary T-cell population during chronic CMV infection by replacing the decaying short-lived pool, however they are not crucial. The memory cells that are primed early in infection are the main contributors to maintenance of the total population (Snyder et al. 2008; Loewendorf et al. 2011). Adoptive transfer of MCMV-specific memory CD8⁺ T-cells into latently infected mice leads to their differentiation into T_{EM}. Differentiated cells proliferate and migrate to peripheral tissues. This suggests that new T_{EM} cells are derived from the restimulation and differentiation of CMV-specific T_{CM} cells after viral reactivation events (Snyder et al. 2008; Klenerman and Oxenius 2016).

Despite the fact that CMV can acquire potent evasion mechanisms that interfere with MHC class I-restricted antigen presentation, strong CMV-specific CD8⁺T-cell are still elicited. This suggests that CMV-specific CD8⁺ T-cells are primed by cross-presentation of CMV antigens by professional APCs. However, cross-presenting

antigen presenting cells alone cannot elicit a CD8⁺ T-cell response. The role of cross-priming in memory inflation has been analysed using *Batf3*^(-/-) mice (Torti et al. 2011). These mice have impaired cross-presentation owing to the lack of CD8 α ⁺ and CD103⁺ DCs. Even in the absence of cross-presenting DCs, inflation of two immunodominant MCMV-specific CD8⁺ T-cell populations was largely normal, suggesting that inflation during latency was mainly dependent on direct antigen presentation.

1.4.11.3 Co-stimulatory molecules and inflammatory cytokines

In addition to antigen presentation, co-stimulatory molecules and inflammatory cytokines are involved in CMV-specific inflationary CD8⁺ T-cell responses. The main co-stimulatory receptor for T-cells is CD28, however, terminally differentiated inflationary CD8⁺ T-cells lose their ability to express CD28. Instead, other co-stimulatory molecules CD137 (4-1BB) and OX40 are up-regulated on inflationary T-cells upon TCR triggering, which are important for CMV-specific memory inflation. Co-stimulatory molecules mediated signalling contributes to the restimulation of CMV-specific CD8⁺ T-cells and thus contributes to the size of inflationary CMV-specific CD8⁺ T-cell pool (Humphreys et al. 2007; Waller et al. 2007; Klenerman and Oxenius 2016).

CMV-specific inflationary T-cell responses depend on IL-2, a key cytokine for T-cell proliferation. Chimeric mice harbouring both IL-2R-competent and IL-2R-deficient T-cells were used to assess the CMV-specific inflationary T-cell response. Absence of high-affinity IL-2 receptors in mice led to a reduction of the CMV-specific inflationary response (Bachmann et al. 2007). Interestingly, it has been shown that autocrine IL-2, but not CD4⁺ T-cell or DC-derived IL-2, is important for optimal secondary expansion of CMV-specific memory CD8⁺ T-cell populations (Feau

et al. 2011). Furthermore, the presence of IL-2 is not only important for priming of CMV-specific memory T-cells, but also for dictating the proliferative capacity of these cells. This is suggested by a study showing that, in the absence of IL-2, these T-cells exhibit an anergic phenotype (Bachmann et al. 2007).

1.4.12 Oligoclonality of CMV-specific repertoire

Inflationary CMV-specific CD8⁺ T-cells often show restricted TCR usage, targeting immunodominant CMV epitopes (O'Hara et al. 2012). The expansion of CMV-specific CD8⁺ T-cells is usually oligoclonal although in some cases it can be monoclonal (Khan et al. 2002). Analysis of the clonal composition of the memory CD8⁺ T-cell repertoire specific to HLA-A2- and HLA-B7-restricted immunodominant pp65 epitopes has shown a high degree of clonal focusing, with selection of TCRs containing conserved amino acid motifs in their TCR hypervariable regions that are shared between individuals, known as public TCRs (Day et al. 2007). Dominant CMV-specific clonotypes selected into the long-term memory pool are distinct from subdominant clonotypes observed during acute infection, with higher functional avidity of dominant clonotypes (Day et al. 2007).

The marked clonal expansion of CMV-specific CD8⁺ T-cells, with subsequent reduction in naive cells and a highly differentiated phenotype (CD8⁺ CD28⁻ CD57⁺), has been implicated in immunosenescence (Looney et al. 1999). Immunosenescence is characterised by a reduction in levels of naive cells, the accumulation of clonally expanded CD28⁻ memory T-cells, and a decline in immune responsiveness (Crough and Khanna 2009).

1.4.13 Other host defences to CMV

1.4.13.1 CD4+ T-cell responses to CMV

Although the role of CD8+ T-cells in control of CMV infection is clearly evident and predominant, CD4+ T-cells also play a critical part. As with the CD8+ T-cells, the frequency of CMV-specific CD4+ T-cell responses is also unusually high, constituting 10% of the CD4+ T-cell memory compartments in adult peripheral blood (Sylwester et al. 2005). The delayed appearance of CD4+ T-cells is associated with CMV reactivation and with symptomatic infection in immunocompromised individuals with primary infection (Antoine et al. 2012). The HIV-infected individuals with absolute CD4 counts below 50 cells/mm³ are at high risk of CMV retinitis. Evidence of CMV pneumonitis is a frequent finding at autopsy in AIDS cases (Wallace and Hannah 1987). In children, the low frequency of CMV-specific CD4+ T-cells is associated with delayed clearance of CMV from the urine until approximately 5yrs of age (Chen et al. 2004). CMV-specific CD4+ T-cells play an important protective role in renal transplant, where low levels of CD4+ T-cells correlate significantly with CMV associated disease (Gamadia et al. 2003; Sester et al. 2005). Furthermore, adoptive transfer of predominantly CD4+ T-cells was shown to result in reduction of CMV viraemia in five out of the seven recipients (Einsele et al. 2002), suggesting a role of CD4+ T-cells in the control of CMV infection.

1.4.13.2 Delta T-cell responses to CMV

As mentioned earlier, an alternative to the $\alpha\beta$ -chain T-cell is the $\gamma\delta$ T-cell, which is composed of the rearranged products of the gamma and delta chain genes. The $\gamma\delta$ T-cells comprise a small fraction (0.5-6%) of peripheral blood in humans, however it represents a greater proportion of lymphoid cells in areas of the body exposed to the external milieu, such as the intestinal mucosa in humans (Déchanet et al.

1999). $\gamma\delta$ T-cells have been shown to play an important protective role in murine herpes simplex virus type 1 infection (Sciammas et al. 1997). Additionally, a protective role of circulating $\gamma\delta$ T-cells has been suggested in renal transplant patients, as marked expansion of these cells is associated with active CMV infection (Dechanet et al. 1999). In these patients, the delayed $\gamma\delta$ T-cell expansion was associated with prolonged and elevated antigenemias and increased severity of CMV disease thus suggesting a protective role of circulating $\gamma\delta$ T-cells in a CMV infected person whose α/β T-cell response has been weakened by immunosuppression (Dechanet et al. 1999).

1.4.13.3 Humoral immune responses to CMV

CMV triggers all arms of the immune system, including humoral immunity. Humoral immune responses are important in preventing viral dissemination and clinical manifestations of CMV-related disease (Jonjić et al. 1994; Gerna et al. 2008). Antibodies specific to structural tegument proteins (pp65 and pp150), envelope glycoproteins and IE proteins are elicited during primary CMV infection (Rosa and Diamond 2012). CMV glycoproteins gB and gH are the main target for neutralizing antibodies. The glycoproteins gB and gH are involved in cell attachment and penetration, and fusion of the viral envelope with the host cell membrane, respectively (Britt et al. 1990; Crough and Khanna 2009). The importance of neutralising antibodies in humans was displayed when the transfer of antibodies from a CMV-seropositive mother to a new-born infant protected against CMV infection caused by a seropositive blood transfusion (Yeager et al. 1981). Furthermore, antibody response of low avidity and poor neutralizing activity is associated with a higher chance of CMV transmission from the mother to foetus (Boppana and Britt 1995). However, infants born to CMV-seropositive mothers are not protected from CMV infection in the first months of life from transplacentally acquired maternal IgG an-

tibodies.

1.4.14 Immune evasion techniques

1.4.14.1 MHC class I pathway

An effective mechanism for clearing CMV infection involves CD8+ T-cell detection of viral antigens presented on infected cells. A major evasion mechanism is focused on decreasing cell-surface expression of MHC class I molecules. CMV is able to remain undetected by interfering with proteasome thus preventing antigen presentation. PP65 has a kinase activity and can phosphorylate the IE1 protein which in turn selectively blocks the processing and presentation of IE-derived antigenic peptides via the MHC class I pathway and hence prevents an IE1-specific CD8+ T-cell response (Gilbert et al. 1996).

CMV encodes five immunomodulatory proteins, glycoprotein (gp) US2, gpUS3, gpUS6, gpUS10 and gpUS11, that can alter MHC class I expression. In the presence of these viral proteins, there is an inhibition in generation of antigenic epitope, blockage of TAP, retention of MHC class I molecules in the endoplasmic reticulum, and recycling of nascent class I heavy chains back to the cytosol where they are degraded by proteasomes. As a result, MHC class I protein expression on the infected cell surface is downregulated (Tomasec et al. 2000; Tortorella et al. 2000; Mocarski 2002; Chevalier and Johnson 2003).

1.4.14.2 MHC class II pathway

As observed with MHC class I regulation, CMV has also evolved multiple mechanisms by which to interfere with antigen presentation via the MHC class II pathway to evade recognition by CD4+ T-cells. CMV encodes a variety of proteins that

can downregulate MHC II cell surface expression, including US2 which targets the MHC class II DR- α and DM- α molecules resulting in proteasome-dependent degradation of MHC class II molecules in a number of different cell types (Tomazin et al. 1999). Another CMV encoded protein gpUS3 binds newly synthesized class II heterodimers in the ER and reduce their association with the invariant chain. Loss of invariant chain causes class II complexes to be inefficiently sorted to the MHC class II loading compartment (Hegde et al. 2002; Chevalier and Johnson 2003). CMV pp65 protein also downregulates MHC class II expression by trafficking MHC class II molecules to lysosomes causing their destruction (Odeberg et al. 2003; Miller-Kittrell and Sparer 2009) .

1.4.14.3 NK cells

As discussed earlier, NK cells play an important role in control of CMV infection. NK cells selectively recognise and kill targets that lack expression of MHC class I molecules (Ljunggren and Karre 1990). The downregulation of MHC class I in CMV-infected cells should therefore result in increased susceptibility to attack by NK cells. However, CMV-infected cells evade NK cell recognition via the production of virus-encoded homologues of MHC class I molecules that act as decoy proteins (Tortorella et al. 2000; Crough and Khanna 2009). Other mechanisms used by CMV to evade NK cell recognition include:(i) CMV encoded UL142-mediated inhibition of NK cell mediated lysis, (ii) CMV encoded pp65 protein-mediated inhibition of the NK cell mediated cytotoxicity by an interaction with receptor NKp30 and (iii) CMV UL141-mediated blocking of the surface expression of CD155, a ligand for activating NK receptors (Arnon et al. 2005; Wills et al. 2005; Tomasec et al. 2010)

1.4.14.4 Homologues

CMV can encode a wide range of other homologues to subvert the host immune response and facilitate persistence. These homologues have distinct functions and can mimic the behaviour of host proteins. Examples of these include:

1. CMV UL111a, which encodes a homologue of the cytokine IL-10, a powerful anti-inflammatory agent and a potent immunosuppressor (Kotenko et al. 2000),
2. CMV UL146 encodes a potent IL-8-like chemokine, viral CXC-1, which induces the chemotaxis of human peripheral blood neutrophils, thereby providing for efficient viral dissemination (Penfold et al. 1999),
3. CMV UL36 and UL37 which encode various anti-apoptotic gene products, vMIA and vICA, and inhibit Fas-mediated apoptosis (Goldmacher et al. 1999; Skaletskaya et al. 2001),
4. CMV UL40 gene encodes a sequence homologous to signal peptides, which upregulates the expression of a non-classical MHC class I molecule, HLA-E. Surface expressed HLA-E suppresses NK cell recognition of virus-infected cells by binding to inhibitory NKG2A receptors on NK cells (Tomasec et al. 2000).

1.5 TCR repertoire of antigen-specific CD8+ T-cell responses

1.5.1 Diversity of TCR repertoire and efficacy of immune response

The first evidence for a link between TCR diversity and protection from infection was reported by Messaoudi et al. (2002), using a mouse model of Herpes simplex virus (HSV-1) infection. The B6 mice, with H-2K^b background, were more susceptible to HSV infection whereas bm8 mice, with H-2K^{bm8} background, were resistant (Messaoudi et al. 2002). C57BL/6 mice expressing H-2K^b (B6; mhc haplotype H-2^b) or H-2K^{bm8} (bm8; mhc H-2^{bm8}) differ from each other only by four amino acids. Both haplotypes can both present the immunodominant HSV epitope (SSIEFARL), with similar binding affinity, thus, ruling out differential binding of HSV epitope to H-2K^b and H-2K as the cause of the differences observed in CD8+ T-cell-mediated HSV-1 resistance. Increased resistance in bm8 mice was shown to be linked to greater TCR diversity and selection of higher avidity gB-specific CD8+ T-cell (Messaoudi et al. 2002).

Several other studies have presented evidence that a diverse TCR repertoire correlates with immune protection against viral escape (Meyer-Olson et al. 2006; Price et al. 2004). Large-scale analysis in a simian immunodeficiency virus (SIV)-Rhesus macaque model using TL8 (TTPESANL) and CM9 (CTPYDINQM) epitope specific TCR sequences (Price et al. 2004) showed that the CD8+ T-cells which recognise CM9 epitope exhibit a diverse TCR repertoire with a diverse array of CDR3- β sequences; whereas the TL8 specific CD8+ T-cells display a highly conserved TCR- β CDR3 motif. The narrow TCR repertoire of CD8+ T-cells responding to the TL8

epitope was associated with rapid selection of TL8 escape variants, whereas, the more diverse clonotypic TCR repertoire of CM9-specific CD8+ T-cells was not associated with viral escape. A likely explanation for this difference is that the more diverse TCR repertoire contains TCRs capable of cross-reacting with epitope variants, thus preventing the emergence of escape variants (Turner et al. 2009).

1.5.2 Diversity of TCR repertoire and magnitude of CD8+ T-cell response

It is unclear whether a relationship exists between TCR diversity and the magnitude of CD8+ T-cells. With respect to the HLA A11-restricted CD8+ T-cell responses against Epstein-Barr virus (EBV) nuclear antigen-4, the subdominant response (to epitope EBNA4 399-408) was populated with a narrow range of TCRs with conserved $V\beta$ usage and identical length amino acid motifs in the CDR3 regions, whereas CD8+ T-cells specific for the dominant epitope (EBNA4 416-424) utilized a diverse TCR repertoire, based on different combinations of TCR- α/β V and J segments and CDR3 regions (Campos-Lima et al. 1997). On the other hand, some studies have observed a correlation between immunodominant epitopes and restricted TCR repertoire. In murine influenza A virus infection, CD8+ T-cell responses to a sub-dominant epitope (D^bPA_{224}) displayed a relatively higher TCR diversity than the immunodominant (D^bNP_{366}) epitope (La Gruta et al. 2006). In HIV-infected humans expressing HLA-B*42:01, the immunodominant Gag-specific TL9 response was more likely to be populated by public TCRs than the subdominant responses (Kloverpris et al. 2015).

1.5.3 Affinity of TCR repertoire influences immune protection

Selection of a diverse TCR repertoire conferring high-affinity TCRs determines viraemic control of HIV infection and limits the emergence of escape mutations (Varela-Rohena et al. 2008; Ladell et al. 2013). A study showed that a high affinity TCR specific for the immunodominant HLA-A*02-restricted HIV gag epitope SL9 (SLYNTVATL) could recognise all common variants of the epitope and produce a greater range of soluble factors and more IL-2 than the wild-type TCR (Varela-Rohena et al. 2008). Studies of immune-mediated control in HLA-B*27-positive individuals, in whom the immunodominant response is towards the Gag KK10 epitope (Gag 263-272), have shown a similar impact of the TCR avidity on outcome. Escape from the HLA-B*27-restricted Gag KK10-specific public TCRs, that were high avidity, via the L268M Gag mutation, typically results in the emergence of a L268M-variant-specific response. However the TCR clonotypes underlying this L268M-specific response are lower avidity and less effective in control of viraemia (Iglesias et al. 2011).

1.5.4 Cross reactivity

To be effective the TCR repertoire needs to include a sufficient number of TCRs to recognise a wide range of foreign antigens, but simultaneously needs to be highly selective so that it can distinguish between self and non-self-antigens (Nikolich-Žugich 2008). A TCR has the ability to recognise different peptides bound to MHC molecules which gives rise to functional diversity, known as cross-reactivity (Mason 1998; Jones et al. 2008). Cross-reactivity can potentially operate to ensure T-cell reactivity to epitope variants that arise as a result of virus escape. An example of this was described above in Mamu-A*01 Rhesus monkeys infected with SIV (Charini et al. 2001). In this study, a strong CD8+ T-cell response was observed

to the immunodominant wild type epitope Gag-p11C (CTPYDINQM) together with a strong response to its variant. The T-cell recognition of the wild type and variant epitope was attributed to the discrete subsets of p11C-specific T-cells that cross-react with cells presenting variant peptides, suggesting that cross reactive CD8+ T-cells are important for recognition of immune escape variants.

1.5.5 Size of naive T-cell repertoire and antigen-specific CD8+ T-cell repertoire

The magnitude of the antigen-specific T-cell repertoire in mature T-cells correlates with the size of corresponding naive T-cell repertoire (Marc and Moon 2012; Moon and Jenkins 2015). An example of this comes from Lymphocytic Choriomeningitis Virus (LCMV) mouse model (Kotturi et al. 2008), in which CD8+ T-cells responded to at least 28 different epitopes, however about one third of the total response was directed towards three of the dominant epitopes. These three dominant peptides displayed similar binding affinity to MHC molecules as the other subdominant epitopes, suggesting that dominance in the CD8+ T-cell response was not entirely due to the peptide-MHC binding affinity. Instead the magnitude of CD8+ T-cell responses to each of 28 identified epitopes was found to correlate with naive CD8+ T-cell precursor frequency. The naive precursor population specific for the dominant peptides was about 10 times larger than those specific for the less dominant peptides (Kotturi et al. 2008). This correlation has also been seen in responses to other viruses, including an experimental model of influenza virus infection in humanized HLA-transgenic mice (Tan et al. 2011) and humans infected with hepatitis C virus (Schmidt et al. 2011).

There is evidence that naive T-cell precursor frequency can also profoundly in-

fluence the pathway along which CD8⁺ memory T-cells develop. Adoptive transfer of naive TCR-transgenic T-cells in a LCMV mouse model correlates with increased differentiation of T_{EM} cells to T_{CM} (Marzo et al. 2005). These mice usually show a linear differentiation pathway from naive to T_{EM} and then to T_{CM} cells.

Together with the size of the naive T-cell repertoire, the extent of naive epitope specific precursor T-cell recruitment from the available pool and/or ongoing proliferation through the course of the primary host response also determines the magnitude of antigen driven T-cell responses (La Gruta et al. 2010). The failure to recruit the complete naive precursor pool and an inability to sustain proliferative responses were the reason for influenza D^bPB1-F2₆₂- and K^bNS2₁₁₄ epitope-specific subdominant responses in virus-infected B6 mice (La Gruta et al. 2010).

1.5.6 Genetic influence on naive repertoire

Since generation of the TCR repertoire involves negative selection of autoreactive T-cell populations, the genetic background of the host can also have a strong influence on the composition of the TCR repertoire. MHC molecules are hugely polymorphic, and each of them preferentially selects some and deletes other TCR specificities. Such deletions are referred to as a “hole” in the repertoire (Nikolich-Zugich 2008).

A mouse study demonstrated that the spectrum of MHC alleles expressed can modify the magnitude of the CD8⁺ T-cell response to an individual epitope (Belz et al. 2000). There was a diminished CD8⁺ T-cell response to H-2D^b-restricted influenza virus epitope in F1 MHC I-heterozygous mice compared to the homozygous mice. This reduction in response was characterized by the deletion of a predominant clonotype in the TCR repertoire of the F1 mice (Belz et al. 2000).

An example of MHC induced alteration in TCR repertoire has been demonstrated in human CD8⁺ T-cell responses to EBV. In EBV seropositive individuals expressing HLA-B*08 and HLA-B*44:02, there was a deletion of CD8⁺ T-cells expressing a dominant public TCR β -chain specific for a HLA-B*08-restricted EBV epitope (FLRGRAYGL). The CD8⁺ T-cells using this conserved TCR β -chain were found to be cross-reactive with the alloantigen HLA-B*44:02. This is an example of self-tolerance creating a potential “hole” in the TCR repertoire for foreign antigen with self-homology (Burrows et al. 1994).

1.6 Immune ontogeny in children

Understanding the ontogeny of the human immune system is fundamental to developing interventions aimed at providing long-term protection against infectious pathogens causing infection in early life (Marchant and Kollmann 2015). The mortality and morbidity of a range of human infectious diseases are both strongly affected by the age at which infection arises. Viruses like CMV and HSV can cause severe infection if these arise in utero or in the neonatal period, whereas they are often asymptomatic if they arise later in childhood. Prolonged replication of HBV and CMV viruses is also commonly observed in early childhood (Marchant and Goldman 2005). There exists a so-called “honeymoon period” of infectious diseases in which levels of disease decreases from infancy to a nadir in early to mid-childhood (3-5yo), increasing again to high levels in teenage years and adulthood (Ahmed et al. 2007; Muenchhoff and Goulder 2014; Marchant and Kollmann 2015) (Fig 1.12). What causes this “honeymoon period” is not well understood due to the paucity of data collected during the childhood years.

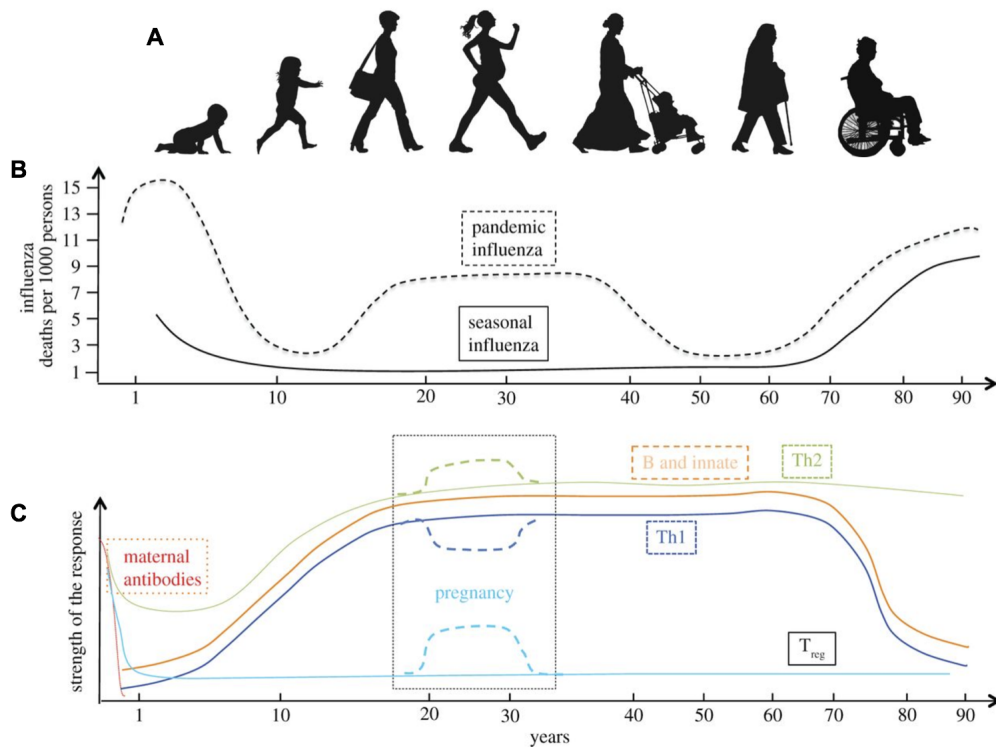


Figure 1.12: Evolution of the immune system. (A) Graphic representation of ages of woman. (B) Excess deaths from seasonal or pandemic influenza over the lifetime of an individual represented as number of deaths per 1000 persons. (C) Changes in immune response to influenza over the lifetime of an individual. *Figure from Simon et al. 2015*

The innate immune response of newborns differs to that of adults (Fig 1.13). TLR stimulation studies reveal quantitative and qualitative differences between neonatal and adult cells (Prendergast et al. 2012). In response to TLR activation, neonatal plasmacytoid DCs produce lower amounts of type I interferons (IFN), in comparison to those of adults (Kollmann et al. 2009). Type I interferons are important for the prevention of viral spread and the induction of antiviral adaptive immune responses (Stetson and Medzhitov 2006). Neonatal innate immune cells also produce lower levels of IFN γ and IL-12, cytokines that support Th1 cell differentiation, and higher levels of IL-6, IL-23 and IL-1 β , which support Th17 cell differentiation (Kollmann et al. 2009). Overall, neonates express an innate Th2 and Th17 cell polarization, weak Th1-polarization, and low innate antiviral type 1 interferon response (Pren-

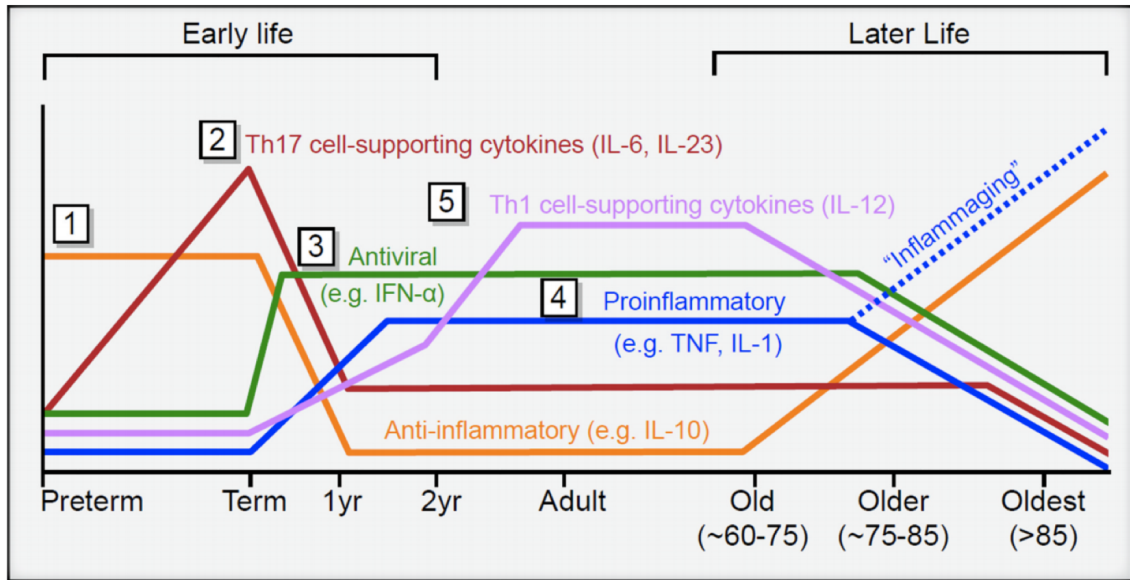


Figure 1.13: Age-dependent changes in TLR-induced immune regulatory function. Figure from Kollmann et al. 2012

dergast et al. 2012). Impaired Th1 responses and lower amount of Type I IFNs in newborns correspond to increased risk of infection with intracellular pathogens such as *Listeria monocytogenes* (David and Levy 2014), *Mycobacterium tuberculosis* (Donald et al. 2010) and herpes simplex virus (Corey and Wald 2009). Consistent with reduced pro-inflammatory/T helper 1 (Th1) cell-polarizing function, impaired responses to vaccines given around birth have also been noted (Siegrist 2001; Levy 2007). Overall, neonates express distinct innate immune defences which enable protection against extracellular bacterial pathogens rather than intracellular bacterial pathogens and viruses, and this is presumably to provide a survival advantage in early life against rapidly fatal bacterial infections (Zhang et al. 2017).

The weak innate Th1-polarizing cytokine production observed in neonates strengthens relatively slowly through childhood, not reaching adult levels until past adolescence. Type I IFN production, however, reaches adult levels in infancy, and

the production of Th17 cell-supporting cytokines, IL-6 and IL-23, and of the anti-inflammatory cytokine IL-10, also reduce rapidly in infancy (Corbett et al. 2010; Kollmann et al. 2012; Prendergast et al. 2012). There is a marked increase in antiviral capacity after the neonatal period, owing to increased type I IFN production, and a gradual increase in responses to intracellular pathogens, although in early childhood these responses remain below adult levels (Prendergast et al. 2012). The adaptive value of this pattern of immune changes through childhood is illustrated by the modest paediatric immune response to common childhood viruses such as VZV that cause chicken pox. The average age of infection in countries such as UK, in temperate zones, is 4 years, at which time most children experience minimal disease. In contrast, adults becoming infected with VZV may suffer significant immunopathology. In countries such as Sri Lanka where 50% of VZV infections occur in adults, chickenpox is the commonest cause of hospital admission (Welgama et al. 2003).

Due to limited exposure to antigens in utero, required to prime adaptive immune responses, infants are strongly dependent on transplacentally acquired maternal antibody and innate immunity for protection against infections (David and Levy 2014). As would be expected, low numbers of memory-effector T and B cells are detected after birth into early infancy and the T-cell and B cell lymphocyte compartments display age-dependent maturation (Adkins 2007). However, as far as CMV is concerned, the CMV-specific CD8⁺ T-cell populations in infancy are similar to those of adults in terms of size, phenotype and function (Appay et al. 2002; Komatsu et al. 2006; Miles et al. 2007, 2008). In chronically infected immunocompetent children, aged 1.5-4yrs, CMV-specific CD8⁺ T-cell responses were similar in magnitude to those observed in infected adults. However, the CD4⁺ T-cell response in children at this age, when CMV can still be detected in the urine, was less than 10% of the

magnitude of responses observed in infected adults.

Studies of the TCR repertoire in childhood have been very limited and those that have been undertaken pre-date the next generation sequencing era. From birth onwards, the TCR repertoire gradually changes, particularly in terms of its diversity (Wedderburn et al. 2001). The neonatal TCR repertoire is highly diverse (Garderet et al. 1998). Over the course of early childhood into adulthood and later in life, the clonal composition of the TCR repertoire changes, becoming less diverse and more oligoclonal (Monteiro et al. 1995), reflecting the experience of the immune system. The changes in TCR repertoire are thought to be associated with clonal expansions of CD8⁺ T-cells that are a result of immune responses to infections with viruses, such as the herpesviruses, EBV (Callan et al. 1996) and CMV (Wang et al. 1995). Clonally expanded CD8⁺ T-cell populations are predominantly within the CD8⁺ CD28⁻ subpopulation and generally oligoclonal, arising from a very limited number of unique TCR clonotypes (Morley et al. 1995; Wedderburn et al. 2001). Clonal T-cell expansions are generally not detectable in neonatal blood with the exception of oligoclonal CD8⁺ T-cell expansions observed during congenital CMV infection (Shen et al. 1998; Marchant et al. 2003; Huygens et al. 2015). Oligoclonal TCR expression in children have previously been associated either with HIV infection (Niehues et al. 2000), autoimmunity (Wedderburn et al. 1999) or immunodeficiency (Bousso et al. 2000).

1.7 Hypotheses and Aims

The aims and hypotheses are as follow:

- i.) **Hypothesis:** Immunodominant CD8+ T-cell epitopes in CMV in African populations will be novel and differ from those published to date from studies of Caucasian subjects.

Aim: To define immunodominant CMV-specific CD8+ T-cell responses in a study cohort comprising Southern African individuals, focusing on three of the most immunogenic CMV proteins pp65, IE1 and IE2.

- ii.) **Hypothesis:** High frequency CD8+ T-cells responses to immunodominant CMV epitopes arise as a result of the use of public TCRs.

Aim: Focusing on a single, novel, immunodominant HLA-B*44:03-restricted CMV-specific response, to analyse the TCR sequences of individuals showing a response to those immunodominant epitope and to determine whether TCR repertoire bias is evident in these individuals.

- iii.) **Hypothesis:** The TCR repertoire populating adult CMV-specific responses remains stable over a decade of chronic infection, whilst that of children alters over the first decade of life.

Aim: To track changes in the TCR repertoire in HLA-B*44:03-positive mother-child pairs over a 10-year period

- iv.) **Hypothesis:** PARV4 and HBV are highly prevalent in sub-Saharan African populations, eliciting potent immune responses like CMV.

Aims: To study the prevalence of PARV4 and HBV in sub-Saharan African populations and to further the understanding of CD8+ T-cell immune responses directed at them.

Chapter 2

Materials and methods

Chapter-specific methods are presented in the methods section of the relevant chapters.

2.1 Study subjects

Subjects from the following cohorts were studied, and approved by the University of Oxford Research Ethics Committees:

- **Thames Valley cohort** (Payne et al. 2010): Chronically HIV-infected ART-naive adults recruited from the Royal Berkshire Hospital, Reading, UK; Northampton General Hospital, Northampton, UK; Churchill Hospital, Oxford, UK; and Wycombe Hospital, High Wycombe, UK. Samples from this study were used to study CMV-specific CD8⁺ T-cell responses in chapter 3.
- **Kimberley HIV-positive cohort** (Adland et al. 2015b; Matthews et al. 2011): HIV-positive children and their mother attending paediatric HIV clinics in Kimberley, South Africa. Samples from this cohort were used to study CMV-specific CD8⁺ T-cell responses in chapter 3 and to study PARV4/hepatitis B

prevalence. Ethics approval was given by the University of the Free State Ethics Committee, Bloemfontein.

- **Gateway Cohort** (Payne et al. 2014): Chronically HIV-infected women recruited in Durban, South Africa. Samples from this cohort were used to study CMV-specific CD8+ T-cell responses in chapter 3 and for TCR repertoire studies in chapter 4. Ethics approval was given by the University of KwaZulu-Natal Biomedical Research Ethics Committee.
- **Paediatric early HAART and strategic interruption study (PHESS) cohort** (Prendergast et al. 2008): HIV-infected infants born to HIV-positive mothers at St Mary's and Prince Mshiyeni Hospitals in KwaZulu-Natal, South Africa, in 2003-2005. Pregnant mothers and infants were recruited and followed up. Samples from this cohort were used in TCR repertoire studies in chapter 5. This study was approved by the Biomedical Research Ethics Committee of the University of KwaZulu-Natal.
- **Gaborone cohort** (Shapiro et al. 2010): Chronically HIV-positive women recruited from the antenatal clinics in Gaborone, Botswana. Samples from this cohort were used to study PARV4/hepatitis B prevalence in chapter 6. The study was approved by the Office of Human Research Administration, Harvard School of Public Health, Boston, USA; and the Health Research Development Committee, Botswana Ministry of Health, Gaborone, Botswana.
- **Sinikithemba cohort** (Brumme et al. 2010; Kiepiela et al. 2007): Chronically HIV-positive women recruited at antenatal clinics in Durban, South Africa. Samples from this cohort were used to study PARV4/hepatitis B prevalence in chapter 6. Ethics approval was given by the Biomedical Research Ethics Committee of the University of KwaZulu-Natal.

- **Masibambisane cohort:** HIV-negative women recruited from antenatal clinics in Durban. Samples from this cohort were used to study PARV4/hepatitis B prevalence in chapter 6. Ethics approval was given by the Biomedical Research Ethics Committee of the University of KwaZulu-Natal.
- **Kimberley Respiratory (KReC) cohort:** HIV-negative children, aged between 9 and 48 months, recruited in the paediatric department at Kimberley Hospital following respiratory tract infection. Samples from this cohort were used to study PARV4/hepatitis B prevalence in chapter 6. The study was approved by the University of the Free State Ethics Committee, Bloemfontein.

All adult subjects, or the parent/guardian for children, provided written informed consent for participation.

2.2 Blood processing

2.2.1 PBMC extraction

Fresh whole blood samples were processed within 24 hours of collection in heparin containers. Prior to the peripheral blood mononuclear cells (PBMC) separation, 3 ml blood was taken and preserved for subsequent DNA extraction. Whole blood was pooled together in a 50 ml falcon tube and spun at 2000 rpm for 10 min to separate plasma. The plasma layer was removed and aliquoted in 1.5 ml cryovials and stored at -80°C for further use. The remaining sample was diluted with PBS and layered onto Lymphoprep (Axis-Shield, Norway) at a blood: PBS: Lymphoprep ratio of 1:1:1 and spun for 24 min at 2000 rpm using the slowest brake setting. The buffy coat layer containing PBMC was collected with a Pasteur pipette, transferred to a new 50 ml falcon tube and washed three times in RPMI-1640. After the final wash, cells were re-suspended in R10 (RPMI-1640, 100 μ /ml penicillin, 100 μ g/ml

streptomycin, 2mM L- Glutamine, 10% heat-inactivated filtered foetal calf serum), and manually counted using Trypan blue staining under a light microscope.

2.2.2 Cyropreservation of cells

$5-10 \times 10^6$ cells were re-suspended in cold heat-inactivated filtered foetal calf serum (0.45 μ m, Sigma) containing 10% DMSO (Sigma, USA) in a drop-wise manner, transferred to pre-chilled cryovials and placed into a pre-chilled Mr Frosty container (Nalgene, USA) at -80°C for at least 80 min before transferring into liquid nitrogen for long-term storage. To thaw, a cryovial with frozen cells was placed in a 37°C water bath to loosen frozen cells, quickly transferred to pre-warmed R10 and spun at 2000 rpm for 5 min. Cells were further washed twice with pre-warmed R10.

2.2.3 DNA extraction

Genomic DNA was extracted from 3 ml of fresh whole blood using QIAmp DNA Blood reagents (Qiagen, Germany). The blood was mixed well with 9 ml red blood cell lysis solution and incubated for 10 min at room temperature. Following the incubation, the mixture was pelleted at 2000 rpm spin for 10 min and re-suspended in 3 ml of white blood cell lysis solution for more than 48 hours at room temperature. Protein precipitation solution of 1 ml was added to the sample and mixed well by vortexing for 20 seconds. The supernatant was then extracted by spinning down the sample at 2000 rpm for 10 min and transferred into 3 ml of isopropanol and inverted at least 50 times until threads of DNA were visible. The sample was then pelleted at 2000 rpm for 3 min and washed with 70% ethanol. The DNA pellet was left to air-dry at room temperature overnight and re-hydrated in RNase/DNase free water (Thermo Fisher, USA). The concentration and quality of DNA were measured on Thermo Scientific NanoDrop 1000 spectrophotometer. DNA was kept at -20°C

for long-term storage or 4 °C for immediate use.

2.3 HLA typing

Four-digit high resolution Sequence Based Typing of HLA-A, -B, and -C was performed from genomic DNA in the CLIA/ASHI accredited laboratory of William Hildebrand, PhD, (ABHI) at the University of Oklahoma Health Sciences Centre using a locus-specific PCR amplification strategy and a heterozygous DNA sequencing methodology for exon 2 and 3 of the class I PCR amplicon. Relevant ambiguities (Cano et al. 2007) were resolved by homozygous sequencing.

2.4 CD4+ T-cell counts and viral load determinations

CD4+ T-cell counts were determined by flow cytometry using standard clinical protocol. HIV plasma viral load was measured using the Roche Amplicor version 1.5 assay with COBAS Amplicor. Data of CD4+ T-cell count and viral load were provided by each centre providing clinical care.

2.5 Interferon-gamma (IFN- γ) enzyme-linked immunosorbent spot (ELISpot assay)

2.5.1 ELISpot procedure

ELISpot assays (Altfeld et al. 2000) were performed using a panel of overlapping CMV peptides (see below). ELISpot assays were performed using sterile 96-well plates with a polyvinylidene fluoride membrane (Millipore, USA). The plate was pre-coated with anti-human IFN- γ monoclonal antibody (Mabtech, Sweden) in

PBS at 1:2000 and incubated at 4°C overnight. Prior to use, the plate was washed with blocking buffer (1% heat-inactivated filtered foetal calf serum in PBS) 6 times and 50 μ l of R10 was added to each well. Peptides were added to a final concentration of 40 μ g/ml per well. Each plate contained two positive controls with 250 μ g/ml of CMV lysate (Virusys) and 250 μ g/ml of phytohaemagglutinin (PHA); and four negative controls with no peptide/CMV lysate. $0.5-1 \times 10^5$ PBMC per well were added and the plate was incubated for 14-16 hours in a humidified incubator at 37°C 5% CO₂. After 6 washes with PBS, biotinylated anti-IFN- γ -antibody (Mabtech, Sweden) in PBS at 1:2000 was added and incubated at room temperature for 90 min in the dark. The plate was washed 6 times with PBS, incubated with streptavidin-alkaline phosphatase conjugate (Mabtech, Sweden) in PBS at 1:2000 for 45 min at room temperature in the dark, washed as before and developed using substrate colour solution (Bio-Rad laboratories, UK) and read using ELISpot reader (ELISpot v.4.0, Autoimmune Diagnostika, Germany) and manually checked. Background (mean of the four negative controls + 3 x standard deviation) was subtracted from values of all wells) and final result was expressed as number of spot forming cells (SFC) per million PBMC.

2.5.2 Peptide dilution and screening for CMV-specific responses by megamatrix

Peptides were synthesised by Schafer-N (Denmark) in powder form. In vitro CMV-specific CD8+ T-cell responses were determined using a panel of 401 overlapping 15mer OLPs spanning the CMV pp65, IE1 and IE2 proteins synthesised. Following dilution to 400 μ g/ml, peptide pools were produced with each pool containing 100 μ l of 11-12 separate peptides, with each peptide present in two different pools. The peptide pools were generated as followed: 120 peptides in IE1 making 22

A

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12
C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
						-	-	-	-	+	+

B

	Pool A1	Pool A2	Pool A3	Pool A4	Pool A5	Pool A6	Pool A7	Pool A8	Pool A9	Pool A10	Pool A11	Pool A12
Pool B1	1	2	3	4	5	6	7	8	9	10	11	12
Pool B2	13	14	15	16	17	18	19	20	21	22	23	24
Pool B3	25	26	27	28	29	30	31	32	33	34	35	36
Pool B4	37	38	39	40	41	42	43	44	45	46	47	48
Pool B5	49	50	51	52	53	54	55	56	57	58	59	60
Pool B6	61	62	63	64	65	66	67	68	69	70	71	72
Pool B7	73	74	75	76	77	78	79	80	81	82	83	84

Figure 2.1: Example of megamatrix setup. (A) ELISpot 96-well plate layout with CMV-specific peptide pools shown as A1-A12, B1-B12, C1-C2, etc. “+” indicates positive control wells, “-” indicates negative controls. Blue shading shows positive responses (pool A1, A3 and B5). **(B)** Overlapping 15mer peptides constituting pool table is shown (only Pools A1-A12 and B1-B7 are shown for convenience). Blue shading indicates the four positive responses in (A). Responses to pools A1/B5 and A3/B5 identify reactivity to OLP 49 and 51, respectively. OLP 49 and 51 would then be used for subsequent confirmatory ELISpot assay.

pools; 143 peptides in IE2 making 24 pools and 138 peptides in pp65 making 23 pools. A peptide was identified as reactive by the positive responses in the two corresponding pools. An example is shown in Fig 2.1. OLPs identified by the megamatrix approach were then tested individually by ELISpot assay.

2.5.3 ELISpot peptide titration assay

The peptide titration assay was performed using the methods described above, with the exception that addition of OLP peptides occurred in progressively diluted quantities. Six 1 in 10 dilutions ranging from 100-0.001 $\mu\text{g/ml}$ were used. The test was carried out 3 times for each sample on separate plates at all peptide dilutions.

2.6 Intracellular cytokine staining (ICS) assay

ICS assays to measure IFN- γ and TNF- α production after stimulation with optimal peptides of purified PBMC were performed. PBMC were incubated with optimal peptide pools (ranging from 100-0.001 $\mu\text{g/ml}$). After 5 hours of incubation at 37°C, brefeldin A (5 $\mu\text{g/ml}$; Sigma-Aldrich) was added, and the cells were incubated for a further 5 hours before 2 mM EDTA was added for 15 min at room temperature. Cells were then washed twice and subsequently stained with surface antibodies against CD3/Alexa Flour 700 (UCHT1, BD Pharmingen), CD4-Qdot605 (S3.5, Life Technologies), and CD8-Pacific Blue (RPA-T8, BD Biosciences). After a 20-min incubation at room temperature in the dark, cells were washed twice and resuspended in Fix/Perm solution (BD Biosciences) for 20 min at 4°C. Cells were then washed twice and resuspended in the ICS antibody mix with fluorescence-conjugated antibodies against IFN- γ /PE-Cy7 (4S.B3, eBioscience), and TNF- α /Alexa Flour 647 (MAb11, BioLegend) for 30 min at room temperature in the dark. Cells were then washed twice with PBS, re-suspended in 2% PFA, acquired on BD LSRII within 24 hours and analysed on FlowJo 8.8.7.

2.7 Tetramer synthesis

Tetramers were synthesized at the University of Copenhagen, Denmark in the laboratory of Prof. Soren Buus according to their published approach (Leisner et al. 2008). Briefly, biotinylated MHC class I monomers were tetramerised by stepwise addition (1/10th volume every 10 min) of extravidin-phycoerythrin (PE) (Sigma), streptavidin-allophycocyanin (APC) (Sigma-Aldrich, USA) at 1:1 molar ratio of biotinylated MHC to biotin binding sites. Tetramers were stored at 4°C. Tetramerisation was checked using compensation beads (BD Biosciences, UK).

2.8 Cell staining

Frozen PBMCs were thawed into R10 medium, rested for 1 h at 37°C in a 5% CO₂ atmosphere. For tetramer staining, $1-10 \times 10^5$ PBMC were re-suspended in 50 μ l/well of R10 in a 96-well plate and washed twice in cold FACS buffer (PBS, 1% BSA, 0.01% NaN₃). Cells were then re-suspended in 50 μ l/well in FACS buffer containing corresponding tet-PE or tet-APC (25 μ l of a 35nM tetramer) and incubated for 20 min at room temperature in the dark and then washed with FACS buffer. Cells were surface stained with CD3-Pacific Orange (Invitrogen, USA), CD8-Pacific Blue (BD Horizon, USA), and LIVE/Dead cell marker (Invitrogen, USA) in a mastermix of 50 μ l/well of PBS and incubated at 4°C for 30 min in the dark. Cells were then washed twice with PBS, re-suspended in 2% PFA, acquired on BD LSRII using FACSDiva software (BD Biosciences, UK) within 24 hours and analysed on FlowJo 8.8.7. The cells were hierarchically gated on singlets, lymphocytes, live cells, and distinct tetramer-specific CD8⁺ T-cells. The spectral overlap between all channels was calculated automatically by the BD FACSDiva software, after measuring negative and single-color controls. The gating strategy is shown in Fig 2.2.

2.9 Cell sorting

Peptide-MHC tetramers were generated as previously described (Leisner et al. 2008). Cryopreserved PBMC (1 million per staining condition) from the recipients were stained with PE-conjugated or APC-conjugated peptide-MHC tetramers, anti-CD3 Pacific Orange (Invitrogen, UK), anti-CD8 V450 (BD Biosciences, Oxford, UK) antibodies and near-IR Live/Dead marker (Invitrogen, Paisley, UK). The samples were sorted on BD FACSAria (BD Biosciences, Paisley, UK) and tetramer⁺ CD8⁺ T-cells were collected in 350 μ l of RLT lysis buffer (QIAGEN, Hilden, Germany).

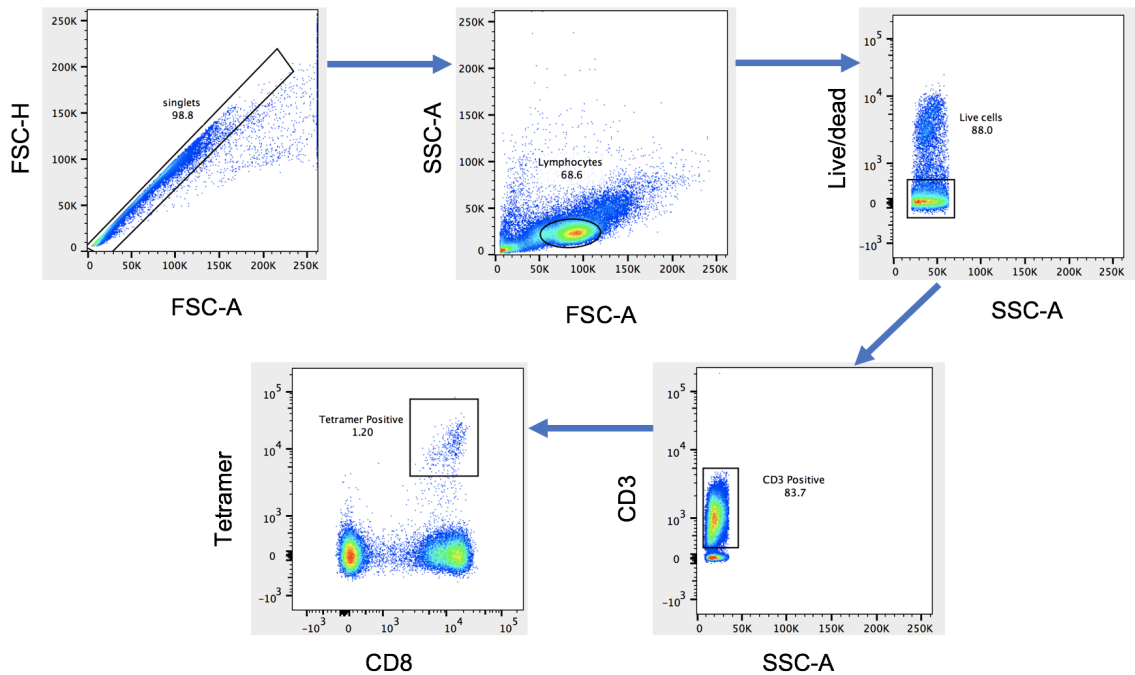


Figure 2.2: Representative example of the gating strategy. Gated on singlets (FSC-A/FSC-H), lymphocytes (FSC-A/SSC-A), live cells (SSC-A/APC-cy7.7), CD3+ cells (SSC-A/pacific orange). Tetramer positive cells expressed as a percentage of live CD3+CD8+ cells.

2.10 RNA extraction

Tetramer sorted cells were collected in 350 μ l Buffer RLT and lysed using vortex. RNA extraction was carried out using RNeasy Plus Micro kit following manufacturer's instructions (Qiagen, Germany). RNA was stored at -80°C until further use.

2.11 TCR sequencing

2.11.1 cDNA synthesis

The cDNA was synthesised using the 5'/3' SMARTer kit (Takara Bio, USA) according to the manufacturer's instructions. The SMARTer approach used a Murine Moloney Leukaemia Virus (MMLV) reverse transcriptase, a 3' oligo-dT primer and

a 5' oligonucleotide to generate cDNA templates flanked by a known, universal anchor sequence (Fig 2.3). This approach paired with a C-gene-specific 3' primer allows PCR amplification without bias. For cDNA synthesis: 1.0 μ l of 5'-CDS primer A (12 μ M) and 10 μ l of RNA was cycled as shown in Table 2.1. Subsequently, 11 μ l 5' RACE-Ready cDNA synthesis reaction mixture was added to 9 μ l master mix (Table 2.2) and

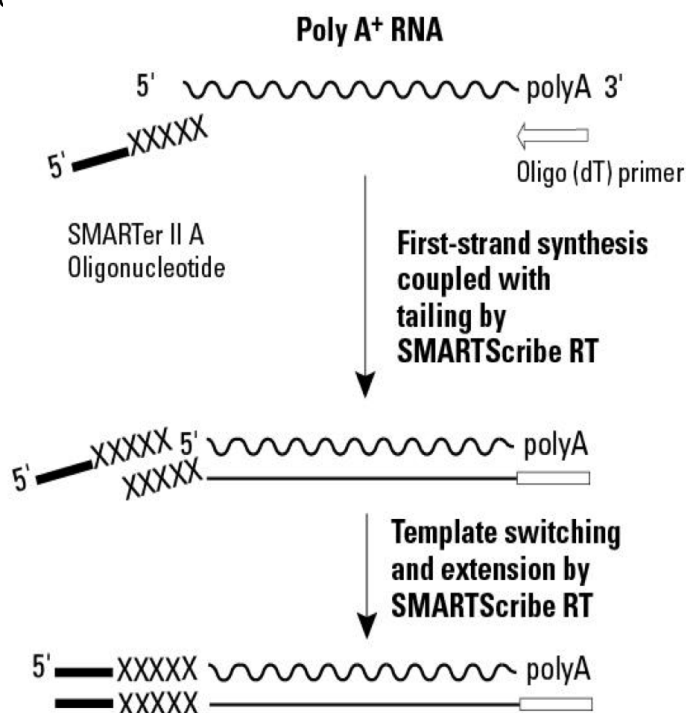


Figure 2.3: cDNA synthesis. A modified oligo (dT) primer, MMLV-derived SMARTScribe Reverse Transcriptase (RT) and the SMART II A Oligonucleotide are used for cDNA synthesis. RT adds non-templated nucleotides when it reaches the 5' end of the mRNA template. The SMART II A Oligonucleotide is enhanced with Locked Nucleic Acid (LNA) technology for increased sensitivity and specificity. It contains a terminal stretch of modified bases that then anneals to the non-templated nucleotides and serves as an extended template for RT. The RT switches template from the mRNA molecule to the SMARTer oligo (this is the template-switching step) generating a complete CDNA copy the original RNA with the additional SMARTer sequence at the end. The additional “SMART sequence” added at the end serves as a primer-annealing site for subsequent rounds of PCR, ensuring that only sequences from full-length cDNAs undergo amplification. *Source: www.clonotech.com*

2.11.2 Primer design

Primers were designed using QuikChange Primer Design online tool (<http://www.genomics.agilent.com/primerDesignProgram.jsp>) based on SMARTer RACE (Rapid amplification of CDNA ends) 5'3' kit (Clontech, USA) recommendations. The primer sequences were visualised in SnapGene Viewer software v3.1.4 and IDT OligoAnalyzer 3.1 (<http://eu.idtdna.com/calc/analyzer>) and Align Sequenced Nucleotide BLAST (NCBI) was used to analyse physical characteristics of oligo sequence, including length, GC content, melting temperature range and possible hairpins and dimers formed by my oligo. Primers were synthesised by Eurofins Genomics.

Step	Temperature	Time
1	72°C	3 minutes
2	42°C	2 minutes

Table 2.1: PCR cycling condition for first step of cDNA synthesis.

Component	Volume	Final concentration
First strand buffer	4 µl	5X
DTT	0.5 µl	100 mM
dNTPs	1 µl	20 mM
RNA inhibitor	0.5 µl	40 U/ µl
SMARTScribe Reverse Transcriptase	2 µl	100 U/ µl
SMARTer II A oligonucleotide	1 µl	24 µM

Table 2.2: cDNA synthesis master mix reaction

Step	Temperature	Time
1	42°C	90 minutes
2	70°C	10 minutes

Table 2.3: PCR cycling condition for second step of cDNA synthesis

2.11.3 PCR amplification of TCR- α and TCR- β chains

The PCR amplification was carried out in two steps. A TCR- α/β constant region-specific reverse primer (Eurofins Genomics, Germany) and an anchor-specific forward primer (Takara, USA) were used in the first PCR reaction. Subsequently, 2.5 μ L of the first round PCR products were taken out to set up a nested PCR as above, using a nested primer pair. Mastermix reaction composition used for step-out (Table 2.4) and nest amplification (Table 2.5) of TCR- α and TCR- β chains, amplification primers (Table 2.6) and PCR conditions (Table 2.7-2.8) are shown.

Component	Volume	Final concentration
High fidelity first strand buffer	10 μ l	5X
dNTP	1 μ l	20 mM
Universal primer mix	5 μ l	10X
C β /ca-round 1 primer	1 μ l	10 μ M
DMSO	0.5 μ l	1%
Phusion High fidelity Taq (Thermo fisher Scientific)	0.25 μ l	0.01 U/ μ L
Deionised H ₂ O	29.75 μ l	Not applicable
Template DNA	2.5 μ l	Not applicable

Table 2.4: Step-out PCR master mix reaction

Component	Volume	Final concentration
High fidelity first strand buffer	10 µl	5X
dNTP	1 µl	20 mM
Universal primer short	1 µl	10X
Cβ/Cα-round 2 primer	1 µl	10 µM
DMSO	0.5 µl	1%
Phusion High fidelity Taq (Thermo fisher Scientific)	0.25 µl	0.01 U/µl
Deionised H ₂ O	33.7 µl	Not applicable
PCR product from round 1	2.5 µl	Not applicable

Table 2.5: Nested PCR master mix reaction

Primer name	Primer sequence (5'-3')
Round 1- TCR-α	C CAT AGA CCT CAT GTC TAG CAC AG
Round 2- TCR-α	GGT GAA TAG GCA GAC AGA CTT GTC
Round 1-TCR-β	GA GAC CCT CAG GCG GCT GCT C
Round 2-TCR-β	TGT GGC CAG GCA CAC CAG TGT G

Table 2.6: PCR primers for amplifying TCR-α and TCR-β chains

Step	Temperature	Time
1	94°C	2 minutes
2	94°C	30 seconds
3	78°C	30 seconds
4	72°C	1.5 minutes
5	Go to step 2, 29 more times	
6	72°C	5 minutes
7	End	

Table 2.7: PCR cycling conditions for step-out PCR

Step	Temperature	Time
1	94°C	2 minutes
2	94°C	30 seconds
3	72°C	30 seconds
4	72°C	1.5 minutes
5	Go to step 2, 29 more times	
6	72°C	5 minutes
7	End	

Table 2.8: PCR cycling conditions for Nested PCR

2.11.4 Agarose gel electrophoresis

After amplification, 47 μ l of PCR products were mixed with 5x sample loading buffer (Bioline, UK) and run on agarose gel. Electrophoresis gels were prepared as 1% agarose powder (Invitrogen) dissolved in TAE buffer (Promega, USA). For DNA visualisation, Sybr Safe (Invitrogen) was added (5 μ l/250ml gel) before allowing the gels to set. Samples were allowed to run at 70V for 40 min on the gel along with 5 μ l of DNA hyperladder I (Bioline, UK) as the indicator for the size of amplicons. Gels were visualised under the Safe Imager blue light transilluminator (ThermoFisher, UK).

2.11.5 DNA gel extraction

DNA fragments of the correct size (400bp) run on agarose gels were excised using a size 11 disposable scalpel and purified using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel, Germany) following the manufacturer's instructions. DNA was stored at 4°C until further use.

2.12 Illumina high-throughput sequencing

Purified DNA products were barcoded, pooled and sequenced on an Illumina MiSeq instrument using the MiSeq v2 reagent kit (Illumina, Cambridge, UK). This was carried by Dr. Meriem Attaf at Cardiff University, Cardiff, UK.

2.13 TCR sequence analysis

Analysis of the raw TCR sequences in FASTQ format and assignment of TCR gene usage was performed using the web-based software MiXCR [<http://mixcr.milaboratory.com>]

(Bolotin et al. 2015). MiXCR employs a built-in library of reference germline V, D, J and C gene sequences for humans based on corresponding loci from the ImMunoGeneTics (IMGT) database (imgt.org). The IMGT nomenclature for TCR gene segments was used throughout the study (Lefranc et al. 2003).

2.14 Statistical analyses

All analyses were performed in Prism (v6.0c, GraphPad). Data were first run to determine the distribution was parametric or non-parametric (D'Agostino and Pearson's omnibus normality test). For two-group analyses, student's t-test (parametric) or Mann-Whitney *U* test (non-parametric) was performed; for ≥ 3 group analyses, one-way ANOVA (parametric) or Kruskal-Wallis (non-parametric) followed by post-hoc test was performed. Nominal data categorised in two distinct ways were compared by Fisher's exact test. *p* values < 0.05 were considered significant. Strength of association between two variables was analysed by Spearman (parametric) and Pearson (non-parametric) correlation. Other statistical analyses performed are specified in relevant chapters.

Chapter 3

Immunodominant cytomegalovirus-specific CD8+ T-cell responses in sub-Saharan African populations

3.1 Introduction

The prevalence of CMV infection is very high in developing countries (Cannon et al. 2010; Adland et al. 2015a) and is endemic in sub-Saharan African populations, with almost two-thirds of infants infected by 3 months of age and 85% infected by a year (Miles et al. 2008). Once adolescence is reached, CMV infection is virtually universal in sub-Saharan Africa (Cannon et al. 2010).

CMV infection is usually asymptomatic but can lead to severe clinical complications in patients who are immunocompromised (Detels et al. 1994; Deayton et al. 2004). In this study, the focus is on the CMV-specific CD8+ T-cell response pri-

marily because T-cell immunity plays an important role in controlling CMV infection and preventing symptomatic, disseminated CMV disease (Riddell and Greenberg 1997; Chen et al. 2004; Reusser et al. 2015). Secondly, the frequency of CMV-specific T-cell responses is unusually high, in adults constituting up to 10% of the CD4⁺ and CD8⁺ T-cell memory compartments in peripheral blood (Sylwester et al. 2005). The high magnitude CMV-specific T-cell responses has also been demonstrated in infancy (Huygens et al. 2014). While the high frequency T-cell responses are of paramount importance in protecting healthy people from subsequent disease, CMV develops life-long persistence and latent infection. Understanding the immune response to CMV is vital due to the prevalence of this virus, the impact it has on the immune system, and the deleterious effects it has on long-term health.

Although the prevalence of CMV is high in sub-Saharan African populations from infancy, most previous studies defining CMV-specific CD8⁺ T-cell epitopes have focused on human leukocyte antigen (HLA) class I alleles that are common in white populations (Kern et al. 1999; Akiyama et al. 2002; Wills et al. 2002; Kondo et al. 2004; Khan et al. 2007; Braendstrup et al. 2014). The main aim of this chapter was therefore to identify the immunodominant CMV-specific CD8⁺ T-cell epitopes in a study cohort of African subjects.

To identify the immunodominant CMV-specific CD8⁺ T-cell epitopes in a study cohort of African subjects, the IFN- γ responses in >250 Southern African individuals were determined by ELISpot assay against a panel of overlapping peptides spanning the phosphoprotein 65 (pp65), 72-kDa immediate early (IE1) and 86-kDa immediate early (IE2) CMV proteins. The reasons for choosing pp65, IE1 and IE2 proteins in this study were two-fold: (i) these proteins have previously been shown to be three of the most immunogenic CMV proteins (Sylwester et al. 2005)

(Fig 1.10); and (ii) I wanted to broaden the study to include IE1 and IE2 as well as pp65 since most of previous studies have focused primarily on the pp65 (Wills et al. 2002; Akiyama et al. 2002). These proteins are important for the structural assembly and replicative capacity of the virus (Crough and Khanna 2009).

3.2 Materials and methods

3.2.1 Study subjects

I studied 257 HIV-infected ART-naive adult and paediatric subjects recruited via three cohorts (Table 3.1):

1. Kimberley, South Africa, n=69 (Adland et al. 2015b)
2. Durban, South Africa, n=84 (Payne et al. 2014)
3. Thames Valley cohort (southern African subjects attending outpatient HIV clinics in the Thames Valley area of the U.K, originating from Botswana, Malawi, South Africa, Kenya, Zambia and Zimbabwe), n=104 (Payne et al. 2010).

In total, I used 62 paediatric (all from the Kimberley cohort) and 195 adult subject samples for ELISpot studies. I used an additional 37 subject samples from Durban cohort for NW8-tetramer staining. The samples were collected between 2009-2014. HIV viral load was measured using the Roche Amplicor version 1.5 assay according the manufacturer's instructions; CD4+ T-cell counts were measured by flow cytometry. The CD4% is calculated as the percentage of live CD3+ CD4+ T-cells in total lymphocytes. The CD4% in children were recorded at the recruitment site.

Cohort Name/ Location	Durban	TVC	Kimberley adults	Kimberley children	Total	Total children
Recruitment source	Antenatal clinics	HIV clinics	Paediatric clinics (mothers of HIV-infected children)	Paediatric clinics	All cohorts	Kimberley
	Positive	Positive	Positive	Positive	Positive	
HIV-status	84	104	7	62	195	62
No of individuals	5,450	3,848	34,706	22,600	4,300	22,600
HIV viral load (RNA copies/ml plasma)	Median IQR	674-15,934	25,559-43,853	4,028-106,084	816-2,300	4,028-106,084
CD4 T -cell count (cells/mm3)	Median	510	528	1018	540	1,018
	IQR	415-699	461-622	724-1,381	396-659	724-1,381
CD4 T-cell %	Median	n/a	n/a	27%	n/a	27
	IQR	n/a	n/a	22%-30%	n/a	22-30%
Age	Median	26	n/a	6	n/a	6
	IQR	23-30	n/a	3-7	n/a	3-7

Table 3.1: Cohort Characteristics

3.2.2 Epitope Prediction using *HLA-Restrictor* — a tool for prediction of HLA restriction elements and optimal epitopes within peptides

The HLA-Restrictor prediction tool (Larsen et al. 2011) was employed to predict the epitopes within the OLPs (OLP sequences are presented in appendix table A2) with a significant HLA restriction. As input it requires four-digit HLA typing and the amino acid sequence of the overlapping 15mer peptide in question. It then predicts the peptide-MHC binding avidity for all possible 8-11mer peptides within the 15mer and on this basis, calculates the likelihood of the optimal epitope.

3.2.3 Statistical analysis

Analysis was performed using Graph-Pad Prism v5.0c (GraphPad Software, Inc.). Associations between peptide recognition and expression of particular HLA class I molecules were calculated with Fisher's exact test using a 2x2 contingency table (p values shown in Table 3.2 and 3.3). In order to correct for multiple comparisons, the Bonferroni correction was applied to give a p value threshold of 0.0009 (calculated as the uncorrected p value of 0.05 divided by the number of HLA class I alleles being studied, n=55).

3.3 Results

3.3.1 Immunodominant CMV-specific CD8⁺ T-cell responses within pp65, IE1 and IE2

All but two of the 257 HIV-infected subjects studied made responses to one or more of the three CMV proteins tested. The proportion of the individuals studied

who showed responses to each of the three proteins tested was 98% for pp65 (n=152), 82% for IE1 (n=95) and 64% for IE2 (n=92) (Fig 3.1A). Fig 3.1B-D show the overlapping peptides that were targeted by more than 4% of individuals tested for pp65 and more than 3% of individuals tested for IE1 and IE2. Within each protein, certain 15mer overlapping peptides were immunodominant: within pp65, two peptides were recognised by 23% of the study population (pp65 OLP-66, and OLP-123), within IE1, one 15mer was recognised by 17% of the study cohort (IE2 OLP-215), and within IE2, one 15mer was recognised by 10% of the subjects studied (IE1 OLP-343) (Fig 3.1B-D). The OLPs adjacent to immunodominant ones were also recognised, and the majority of the subjects making a response to these were within the group who recognised immunodominant OLPs. For example, 70% of the individuals responding to OLP-67 also responded to OLP-66, 88% of the individuals responding to OLP-124 also responded to OLP-123, 92% of the individuals responding to OLP-216 also responded to OLP-215 and 50% of the individuals responding to OLP-342 also responded to OLP-343. Recognition of two adjacent OLPs indicates the location of the 8-11mer optimal epitope, at least in part, within both 15mers.

Recognition of these 15mers in the ELISpot assays was in each case strongly associated with expression of particular HLA class I molecules (Table 3.2). An epitope predictor tool HLA-Restrictor (Larsen et al. 2011) was then used to predict the optimal epitope from within the 15mers that would be presented by the relevant HLA class I molecule. This yielded 3 well-characterised and previously described epitopes (Weekes et al. 1999; Ameres et al. 2013) including the HLA-A*02:01-restricted epitope NLVPMVATV (Weekes et al. 1999) and the HLA-B*07:02-restricted epitope RIPHERNGFTVL (Weekes et al. 1999), both in pp65. Two novel epitope candidates emerged also from this analysis, an HLA-B*42:01-restricted epitope

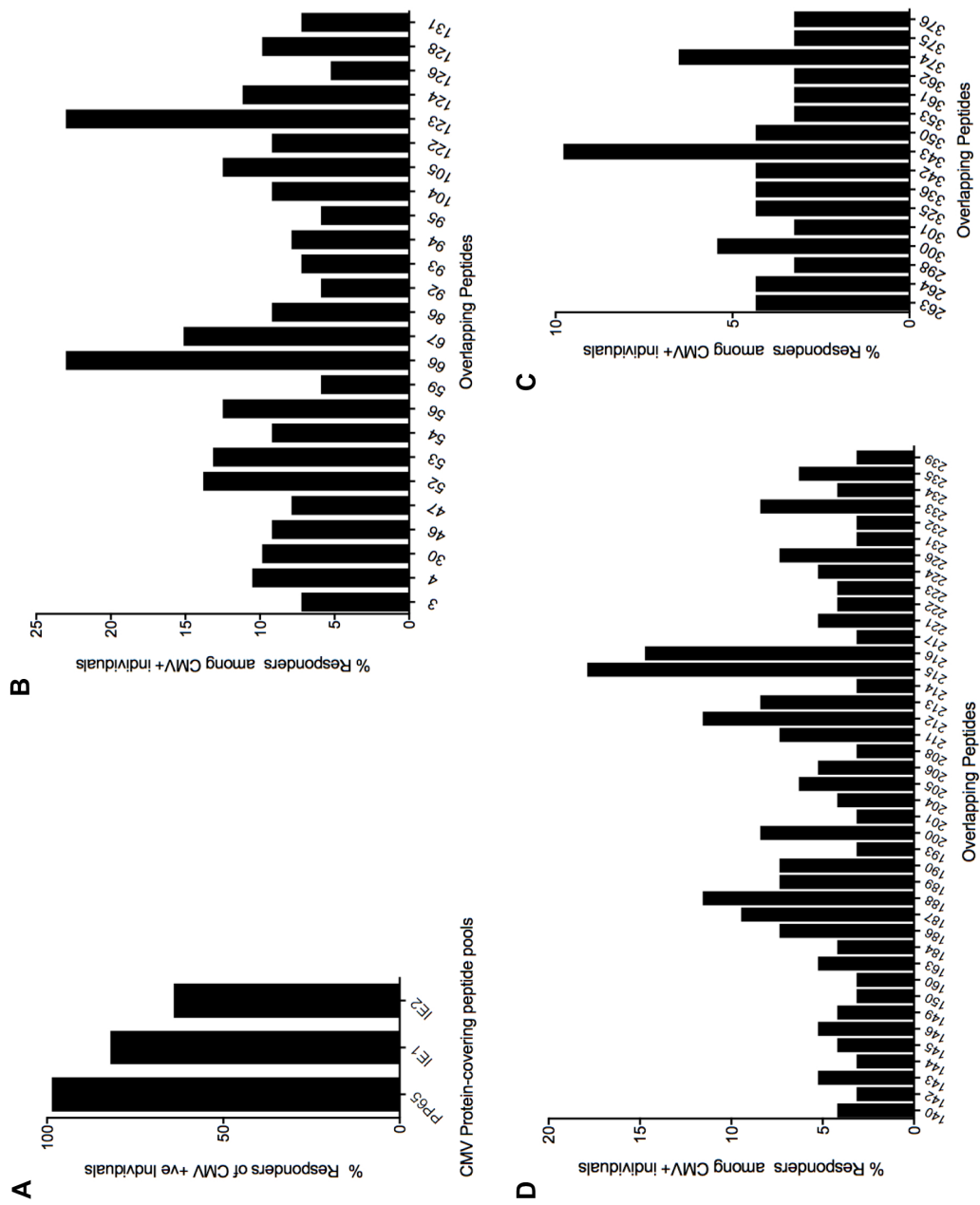


Figure 3.1: Immunodominant CMV-specific CD8+ T-cell responses within pp65, IE1 and IE2 measured by IFN- γ ELISpot assay for CMV+ subjects. (A) Percentage responders to each protein-covering peptide pool. (B) Percentage responders among 152 CMV+ individuals to pp65 15mer overlapping peptides that were targeted by >4% of the study population. (C) Percentage responders among 95 CMV+ individuals to IE1 15mer overlapping peptides that were targeted by >3% of the study population. (D) Percentage responders among 92 CMV+ individuals to IE2 15mer overlapping peptides that were targeted by >3% of the study population.

RIPHERNGFTVL in pp65 and an HLA-A*30:01-restricted epitope KVRNIMKDK within IE2. The respective peptide-MHC tetramers were synthesised and used to stain PBMC in subjects who had shown responses to the 15mer peptides; in each case, these tetramers confirmed the optimal epitope and the HLA restriction of the response (Fig 3.2). One subject per epitope was tested with the corresponding tetramer.

This approach to defining the optimal CMV-specific epitopes within a Southern African study cohort therefore was validated for the 15mer overlapping peptides that were best recognised within the three CMV proteins studied.

3.3.2 Identification of additional novel CD8+ T-cell epitopes within pp65, IE1 and IE2

Due to the encouraging results, I decided to additionally analyse the remaining 15mer peptides to which responses were observed in $\geq 4\%$ of the study cohort. This threshold was chosen because, at $< 4\%$ recognition, the number of subjects studied did not provide sufficient statistical power to detect associations between peptide recognition and expression of particular HLA class I molecules that would remain significant ($p < 0.05$) after correction for multiple tests (uncorrected p value < 0.0009 : see Methods). In addition to the five epitopes described above (Table 3.2), I identified responses to 18 predicted CMV-specific CD8+ T-cell epitopes that were significantly associated with expression of particular HLA class I molecules (uncorrected $p < 0.0009$, corrected $p < 0.05$) (Table 3.3). Using epitope predictor tool, I found that 14 predicted epitopes are novel and four have been previously described in the literature (Table 3.3). A subset of the novel epitopes was confirmed using peptide-MHC tetramers or via intracellular cytokine staining (Fig 3.3).

Protein	OLP	Sequence	HLA association	p value (Fisher's)	Predicted optimal	Described/ Novel
pp65	66/67	QPFMRPHERNGFTVL	B*07:02	<0.0001	RIPHERNGFTVL	Described (a)
		RIPHERNGFTVLCPKN	B*42:01	<0.0001	RIPHERNGFTVL	Novel
pp65	123/124	AGILARNLVPMTVATV ARNLVPMTVATVQQQN	A*02:01	<0.0001	NLVPMTVATV	Described (a)
IE-1	215/216	LSEFCRVLCCYVLEE CRVLCCYVLEEELTSVM	C*07:02	<0.0001	CRVLCCYVL	Described (b)
IE-2	342/343	LDNEKVRNIMKDKNT KVRNIMKDKNTPFCT	A*30:01	<0.0001	KVRNIMKDK	Novel

Table 3.2: HLA associations with the most targeted 15-mer OLP's in CMV pp65, IE1 and IE2. The table shows the epitopes that were identified in our preliminary investigation using only the most immunogenic OLPs (OLP 66/67, OLP123/124, OLP215/216, OLP 342/343) in each protein. I identified two novel epitopes, for which we show FACS plots in Fig 3.2. Three epitopes have been previously described [a - (Weekes et al. 1999); b - (Ameres et al. 2013)].

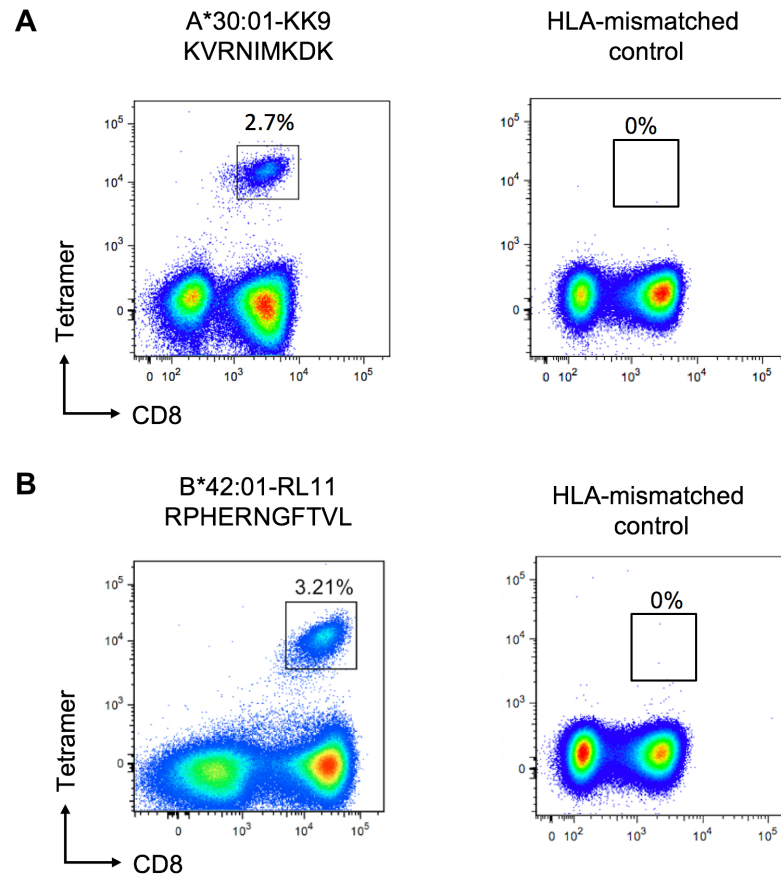


Figure 3.2: Validation of novel CD8+ T-cell epitopes in CMV using HLA-Class I tetramers. (A) Flow cytometry plot of CD8+ T-cell responses from a Durban cohort subject SK-251 (HLA-A*02:05, -A*30:01, -B*42:01, -B*58:01, -C*07:01, -C*017:01) to HLA-A*30:01-restricted IE2-342/343 epitope-KK9 and an HLA-mismatched tetramer stain as a control. (B) Flow cytometry plot from a Durban cohort subject SK-331 (HLA-A*30:01, -A*66:01, -B*42:01, -B*58:02, -C*06:02, -C*017:01) to B*42:01-restricted pp65-66/67 epitope RL11 and an HLA-mismatched tetramer stain as a control. The plots show gated live CD3+ T-cells; the number shown above each gate is the percentage of live CD3+ CD8+ cells that are tetramer-specific.

In summary, across both experiments, we found that seven of the yielded epitopes had been previously described and the remaining 16 predicted epitopes are novel. As might be expected, the previously described epitopes are presented by HLA class I molecules that are highly prevalent in white populations whilst the novel predicted epitopes are restricted by HLA class I molecules that are in many cases broadly African-specific. The PF of the HLA alleles presenting the CMV-specific epitopes identified in this study are shown in Fig 3.4 (the figure shows the PF of HLA alleles presenting the CMV-specific epitopes in a representative white population (Cao et al. 2001) and our African cohort).

3.3.3 Closely related HLA I molecules shape distinct CMV-specific CD8+ T-cell hierarchies

Comparison of the HLA class I molecules prevalent in white and sub-Saharan African populations in many cases reveals substantial differences in phenotypic frequency between the closely-related HLA alleles. Among HLA-A*02-positive white individuals, for example, 97% express HLA-A*02:01 (Cao et al. 2001), whereas approximately 50% of HLA-A*02-positive individuals in our African cohort express HLA-A*02:02 or HLA-A*02:05, HLA-A*02 molecules that differ only by 3 and 4 amino acids, respectively, from HLA-A*02:01 (Fig 3.5A). Nonetheless, they may not share the same pattern of immunodominance. For example, it is clear from this study that the immunodominant HLA-A*02:01-restricted response is towards the pp65 OLP-123/124 peptide (NLVPMVATV), as described previously (Weekes et al. 1999), whereas we suggest here that the immunodominant HLA-A*02:05-restricted response, more prevalent in Africans, is towards the pp65 OLP-46/47 peptide (FVFPTKDV) (Fig 3.5B).

Protein	OLP	Sequence	HLA association	p value (Fisher's)	Predicted optimal	Described/ Novel
pp65	4	SVLGPISGHVLKAVF	B*81:01	<0.0001	GPISGHVL	Novel
pp65	4	SVLGPISGHVLKAVF	B*39:10	<0.0001	GPISGHVL	Novel
pp65	30	PLKMLNIPSINVHHY	B*35:01	<0.0001	IPSINVHHY	Described (a)
pp65	46/47	YYTSAFVFPTKDVAL	A*02:05	<0.0001	FVFPTKDV	Novel
pp65		AFVFPTKDVALRHW				
pp65	52	VCSMENTRATKMQVI	C*06:02	<0.0001	TRATKMQVI	Described (b)
pp65	53/54	ENTRATKMQVIGDQY	A*30:02	<0.0001	KMQVIGDQY	Novel
pp65		ATKMQVIGDQYVKVY				
pp65	104/105	TERKTPRVTTGGGAMA	B*07:02	<0.0001	TPRVTTGGGAM	Described (c)
pp65		TPRVTTGGGAMAGAST				
pp65	131	YRIFAELEGVWQPAA	B*45:01	0.0008	AELEGVWQPA	Novel
IE-1	163	RHRIKEHMLKKYTQT	A*30:01	0.001	RIKEHMLKKY	Novel
IE-1	186	LQAKARAKKDELRRK	A*74:01	<0.0001	LQAKARAKK	Novel
IE-1	187	ARAKKDELRRKMMYM	B*08:01	<0.0001	ELRRKMMYM	Described (d)
IE-1	187/188	ARAKKDELRRKMMYM	B*18:01	<0.0001	DELRRKMMY	Novel
IE-1		KDELRRKMMYMCYRN				
IE-1	200	EIMAYAQKIFILDE	A*23:01	0.0004	AYAQKIFKI	Novel
IE-1	212/213	ACMMTMYGGISLLSE	B*42:01	<0.0001	TMYGGISLL	Novel
IE-1		TMYGGISLLSEFCRV				
IE-1	226	KVFAQYILGADPLRV	C*07:02	<0.0001	QYILGADPL	Novel
IE-1	233	ESDEEEAIVAYTLAT	B*35:01	0.0007	EIVAYTL	Novel
IE-1	234	EEAIVAYTLATAGVS	B*40:02	0.001	EEAIVAYTL	Novel
IE-2	374/375	TADACNEGVKAAWSL	B*44:03	<0.0001 a	NEGVKAAW	Novel
		CNEGVKAAWSLKELH				

Table 3.3: HLA associations with the remaining 15-mer OLP's in CMV pp65, IE1 and IE2 that were targeted by $\geq 4\%$ of the study cohort. The table shows additional predicted epitopes identified using the remaining 15mer peptides that were targeted by $\geq 4\%$ of the study cohort. We identified an additional 14 novel predicted epitopes, and four previously published epitopes (a - Gavin et al. 1993; b - Longmate et al. 2001; c - Elkington et al. 2003; d - Rist et al. 2009).

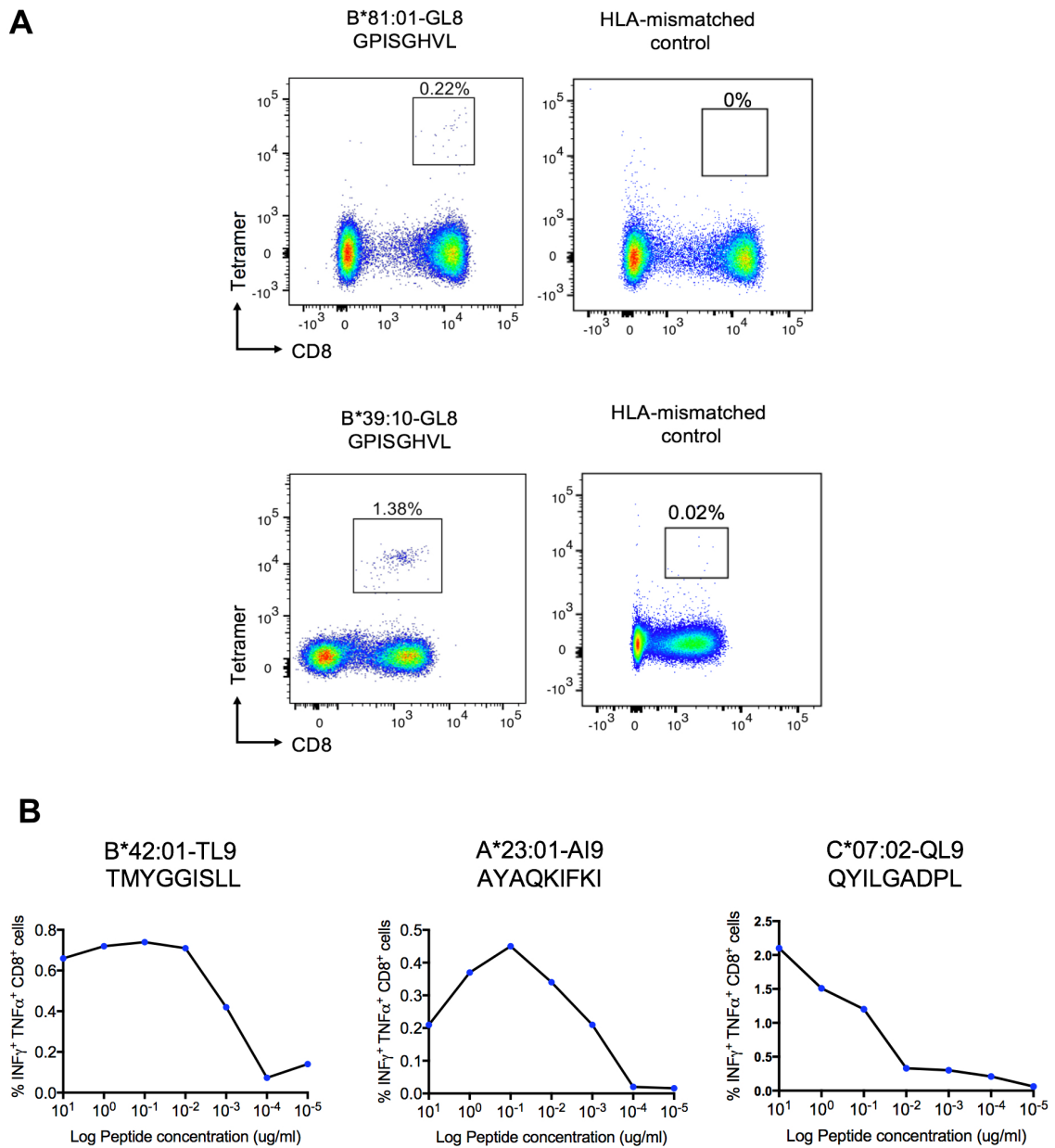


Figure 3.3: Validation of additional novel CD8+ T-cell epitopes in CMV using HLA-Class I tetramers and intracellular cytokine assay. (A) Flow cytometry plots of CD8+ T-cell responses from HLA matched and HLA mismatched donors to B*81:01-restricted pp65-4 epitope GL8 and HLA-B*39:10-restricted pp65-4 epitope GL8. The plots show gated CD3+ T-cells; the number shown above each gate is the percentage of total lymphocytes that are tetramer-specific (tetramer positive cells expressed as a percentage of live CD3+CD8+ cells). **(B)** Titration curves of IFN- γ and TNF- α cytokine responses from HLA matched donor to HLA B*42:01 restricted IE1-212/213 epitope TL9, HLA A*23:01 restricted IE1-200 AI9 and HLA C*07:02 restricted IE1-226 QL9. The titration curves show the magnitude of responses to optimal peptides ranging from 100-0.0001 $\mu\text{g/ml}$ in an intracellular cytokine assay.

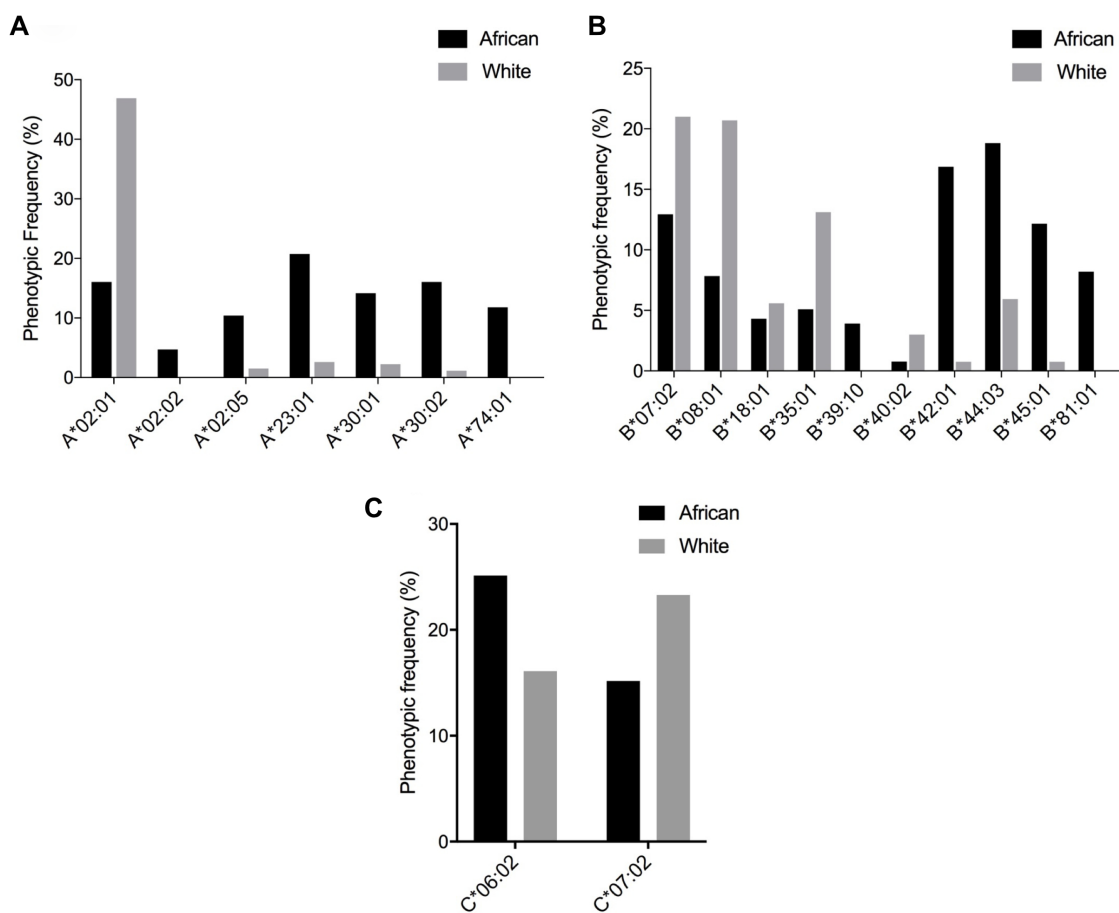


Figure 3.4: Phenotypic frequencies of the class I HLA molecules in the African study cohort and a representative white population. (A) Phenotypic frequencies of HLA-A molecules, (B) Phenotypic frequencies of HLA-B molecules, (C) Phenotypic frequencies of HLA-C molecules; Phenotypic frequencies of HLA alleles highlighted in Table 3.2 and Table 3.3.

Similarly, among HLA class I molecules within the B7 superfamily, only HLA-B*07:02-restricted responses have been well characterised to date. The HLA-B*07:02 allele is the most common allele in B7 superfamily in white populations (Cao et al. 2001). There are two pp65-specific HLA-B*07:02-restricted epitopes that are both targeted by >50% of subjects expressing HLA-B*07:02. However, one of these two epitopes is clearly immunodominant among HLA-B*42:01-restricted CMV-specific responses, and neither is significantly targeted by subjects expressing either HLA-B*39:10, HLA-B*81:01 or HLAB*35:01 (Fig 3.5D). HLA-B*39:10, HLA-B*81:01 and HLA-B*35:01 alleles are part of the B7 superfamily and relatively common African alleles. Thus, as might be expected and shown previously in relation to other virus-specific CD8+ T-cell responses (Leslie et al. 2006; Carlson et al. 2012), the difference of even a few amino acids between closely-related HLA class I molecules can have a substantial impact on the CD8+ T-cell immunodominance hierarchy observed.

3.3.4 Definition of a high frequency HLA-B*44:03-restricted IE2-specific response

One of the novel epitopes identified in these studies is the epitope NEGVKAAW (NW8) in IE2 restricted by HLA-B*44:03. The responses to this epitope were unusually high, even for CMV-specific CD8+ T-cell responses that are generally high magnitude responses, in one individual accounting for 18.8% of CD8+ T-cells (Fig 3.6A). As controls, I show representative FACS plots of a HLA-B*44:03 expressing individual stained with an HLA-mismatched tetramer and a non-HLA-B*44:03 expressing individual stained with the NW8-tetramer (3.6B-C). I tested a subset of 37 individuals from the Durban cohort with the NW8-epitope tetramer, that were

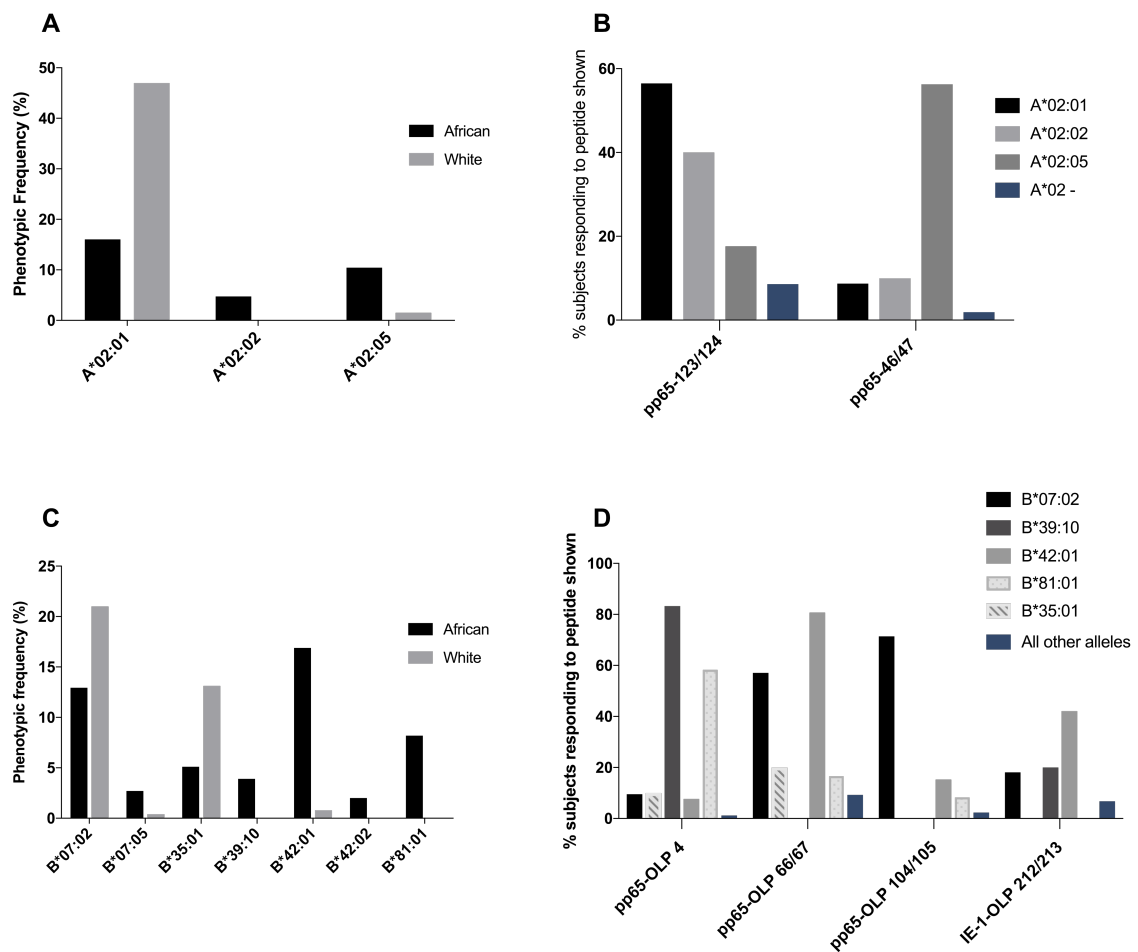


Figure 3.5: Phenotypic frequencies of closely related HLA I molecules and their impact on CMV specific CD8+ T-cell hierarchies. (A) The phenotypic frequency of some of the HLA class I molecules in A2 superfamily in white populations (Cao et al. 2001) and our African cohort. **(B)** The percentage of subjects with HLA alleles A*02:01 (n=23), A*02:02 (n=10) and A*02:05 (n=16), responding to epitope (NLVPMVATV) in pp65-123/124 peptide and predicted epitope (FVFPKTDV) in pp65-46/47 peptide. **(C)** The phenotypic frequency of the HLA class I molecules in B07 superfamily in white populations (Cao et al. 2001) and our African cohort. **(D)** The percentage of subjects with B*07:02 (n=21), B*35:01 (n=10), B*39:10 (n=6), B*42:01 (n=26) and B*81:01(n=12) alleles in our African cohort responding to predicted epitopes in pp65, GPISGHVL (OLP-4), RIPHERNGFTVL (OLP-66/67) and TPRVTGGGAM (OLP-104/105). The percentage of subjects in our African cohort responding to predicted epitope in IE1, TMYGGISLL (OLP-212/213) with HLA-B*07:02 (n=11), B*35:01 (n=6) B*39:10 (n=5), B*42:01 (n=19) and B*81:01(n=1) alleles.

not part of the ELISpot studies. Among the 37 subjects tested, 35 (95%) made responses above the positive threshold of $>0.02\%$ of CD8+ T-cells to this epitope, with a mean CD8+ T-cell response of 3.4% (Fig 3.6D).

3.3.5 Relative contribution of HLA-A, -B and -C alleles to immune recognition of CMV

As described above, I identified significant associations between expression of the particular class I alleles and recognition of certain CMV-specific peptides in pp65, IE1 and IE2. These 20 statistically significant HLA-peptide associations reflect true HLA class I restricted CD8+ T-cell responses, as confirmed by tetramer staining. I then calculated from these data the relative contribution of HLA alleles in immune recognition of CMV (Fig 3.7). The majority (12/20) of the 20 CMV-specific responses identified in the study cohort were restricted by HLA-B alleles. The HLA-B-restricted responses were more frequent than those restricted by HLA-C (12/20 versus 2/20, p value = 0.01, two-tailed unpaired t-test; mean peptide associations per allele, 0.75 versus 0.15) and also than those restricted by HLA-A alleles (12/20 versus 6/20, p value = 0.08; mean peptide associations per allele, 0.75 versus 0.3). There was no difference observed between HLA-A and HLA-C (p value = 0.23). Overall, our data suggested that HLA-B alleles contribute more than the HLA-A and -C alleles to immune recognition of CMV.

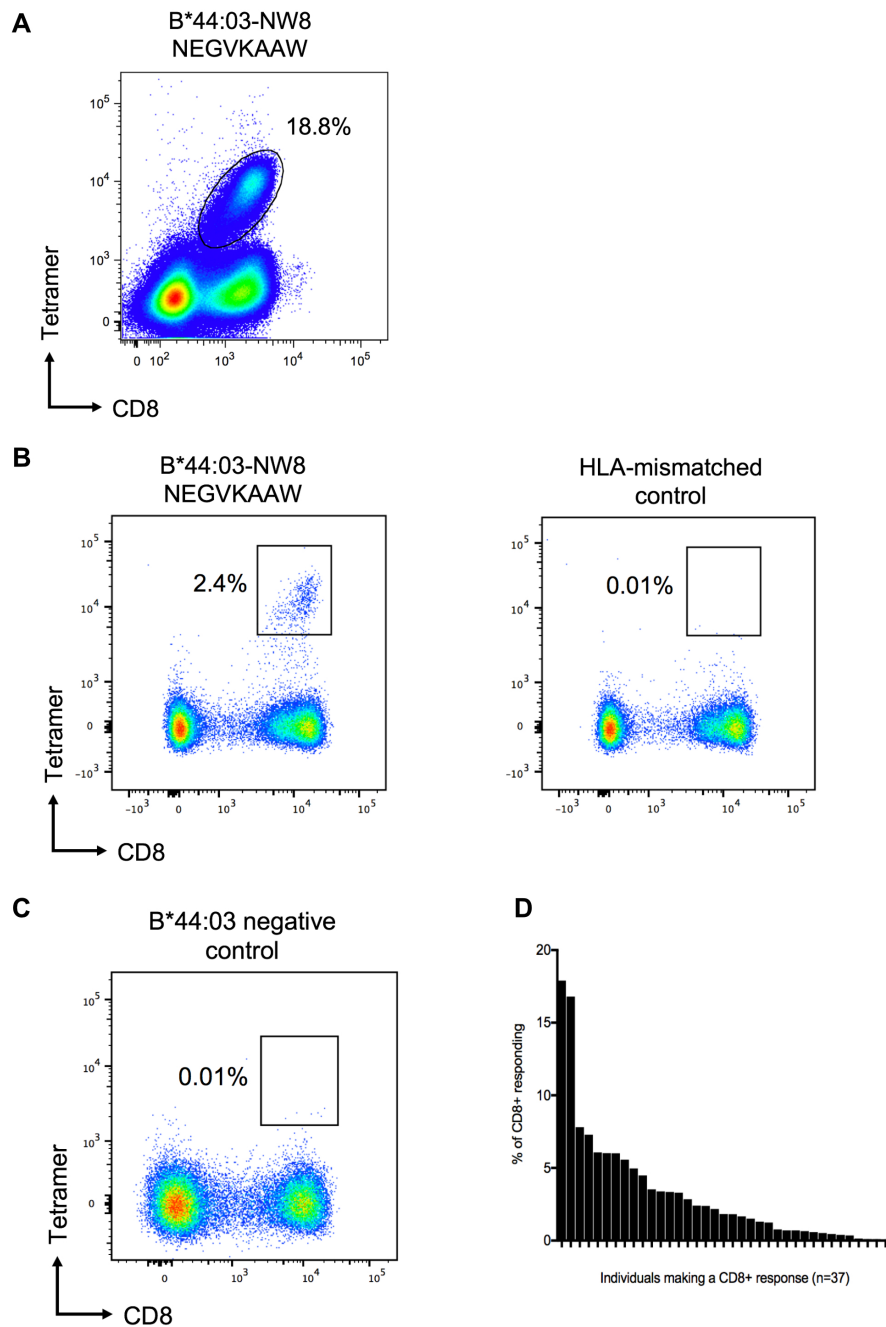


Figure 3.6: Individuals making CD8+ T-cell responses against CMV IE2 epitope HLA-B*44:03-NW8. (A) Representative FACS plot of a Durban cohort subject 202-30-0064 (expressing HLA-B*44:03/B*58:01) to HLA-B*44:03-restricted NW8 tetramer. (B) FACS plot of subject R048 (HLA-A*29:02, -A*6802, -B*15:03, -B*44:03, -C*02:10, -C07:01) with NW8 tetramer and HLA-mismatched tetramer. (C) FACS plot of a TVC subject R112 (HLA-A*23:01, -A*3002, -B*18:01, -B*18:01, -C*02:02, -C07:04) to HLA-B*44:03-restricted NW8 tetramer, as B*44:03-negative control. The FACS plots show tetramer-specific cells expressed as the percentage of live CD3+ CD8+ T-cells. (D) Percentage of live CD3+ CD8+ tetramer-specific T-cells in individuals expressing HLA-B*44:03 measured using the B*44:03-NW8 tetramer.

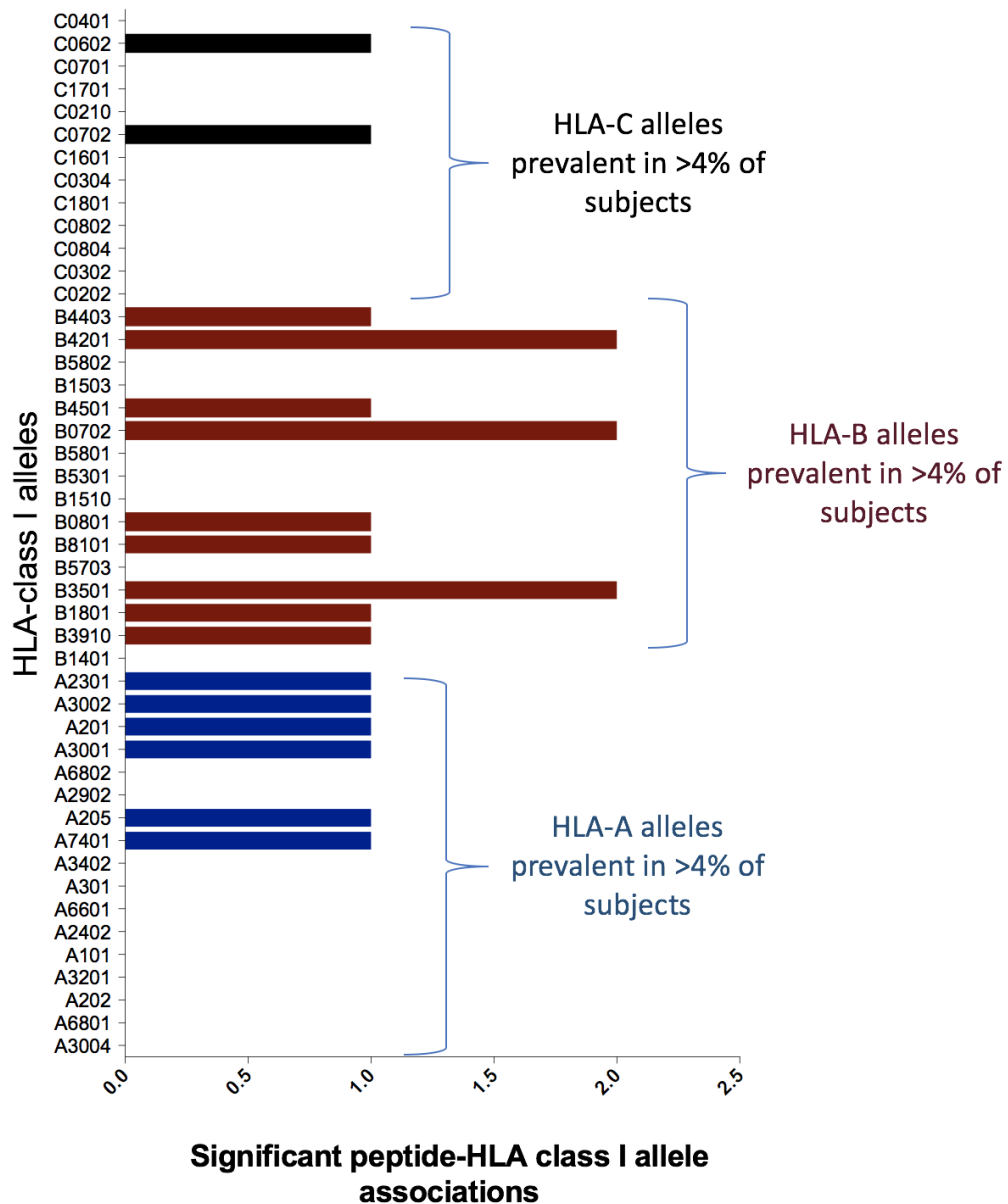


Figure 3.7: Number of peptides recognised in association with individual HLA allele expression. The figure shows the HLA-A (blue), HLA-B (red) and HLA-C alleles (black) that were targeted by more than 4% of the study population. The number of peptides recognised in association with these HLA alleles with statistical significance (p value=0.0009) are shown. From a total of 20 peptide-HLA associations identified, 12 were HLA-B allele restricted, six were HLA-A allele restricted and two were HLA-C allele restricted.

3.4 Discussion

One aim of this study was to define the immunodominant CMV-specific CD8⁺ T-cell responses in a study cohort comprising Southern African individuals, focusing on three of the most immunogenic CMV proteins pp65, IE1 and IE2. The strategy of using a functional screen first and a bioinformatics screen second, was an efficient approach to the identification of CD8⁺ T-cell epitopes and their HLA restriction elements. I validated some of these immunodominant responses using peptide-MHC-class I tetramers and intracellular cytokine assays. In this way, I identified 16 novel CMV-specific epitopes within CMV pp65, IE1 and IE2 proteins. Using a population of HIV-infected subjects, all of whom have sub-Saharan African origin, it has been possible, for the first time, to investigate HLA restrictions for the most prevalent alleles in sub-Saharan Africa. The resulting epitopes will underpin further understanding of CMV specific T-cell immunity, which is of particular interest considering the high frequency of CD8⁺ T-cells that recognise these epitopes.

One of the validated epitopes was an IE2 epitope restricted by the predominantly African allele, HLA*B*44:03, that induced consistently high CD8⁺ T-cell responses in individuals expressing HLA-B*44:03, in one individual approaching 19% of their CD8⁺ cells. Typically, during viral infections, the establishment of a memory T-cell population in response to an antigen is thought to involve expansion followed by contraction into a stable pool of memory cells, once the infection is cleared (O'Hara et al. 2012). However, in CMV infection, CD8⁺ T-cells specific for certain epitopes do not contract but instead are maintained at high frequencies. This increase in CMV specific CD8⁺ T-cell response has been termed "memory inflation" (Karrer et al. 2003). "Memory inflation" is a term derived from studies of murine cytomegalovirus (MCMV) that detected unusually high numbers of MCMV-specific

CD8⁺ T-cells and eventual stabilization of memory CD8⁺ T-cells at high frequencies, in immune competent mice (Karrer et al. 2003). The underlying mechanisms of memory T-cell inflation in humans are not yet fully elucidated. However, one feature described is restricted TCR usage (Weekes et al. 1999; Klenerman and Oxenius 2016), and investigating the underlying TCR sequences would help in the understanding of this unusual phenomenon. We looked for evidence of memory inflation in our cohort but found no association between the size of CD8⁺ T-cell responses in relation to the age of the individuals. This may be due to the adults in our study cohort being relatively young (median age of 26 years).

HLA-B*44:03, a common allele in African populations, forms a part of HLA-B*44 supertype family. The predominant HLA-B*44 allele in white populations is HLA-B*44:02. Although these two HLA-B44 allotypes differ only at one position, residue 156 (HLA-B*44:02: 156-Asp; HLA-B*44:03: 156-Leu), the impact of this micro-polymorphism appears to be that HLA-B*44:03 presents virtually all the peptides that are presented by HLA-B*44:02, plus some additional peptides not presented by HLA-B*44:02 (Macdonald et al. 2003). The one HLA-B*44:02-positive subject tested made no IFN- γ ELISpot response to the overlapping peptides containing NW8.

Examining the influence of HLA alleles on immune recognition of CMV revealed a somewhat dominant role for HLA-B alleles in CMV infection. This is in line with previously reported data that HLA-B-restricted CD8⁺ T-cell responses are the most important in shaping the evolution of HIV infection (Kiepiela et al. 2004). Greater diversity in peptides that bind HLA-B alleles might explain the mechanism for the dominant HLA-B allele effect. In contrast to HLA-A alleles, the amino acids binding the B, F, C and D pockets of HLA-B are more varied. This diversity in HLA-B alleles

arises from the intralocus recombination events within exon 3 which mainly affects the F, C and D pockets (Marsh et al. 1999). The one amino acid difference at position 156 between HLA-B*44:02 and B*44:03, for example, would be predicted only to affect the D pocket (Macdonald et al. 2003). The greater functional diversity of HLA-B alleles means that CMV is less likely to evade immune recognition by them.

These studies were undertaken in a cohort of HIV-infected African subjects, prompting the question of whether the CMV-specific CD8⁺ T-cell responses might be affected. It is well established that, at extremely low absolute CD4 count (<50 CD4 T-cells/mm³) HIV-infected individuals are at risk of opportunistic CMV infections such as CMV retinitis (Crough and Khanna 2009). However, the HIV-infected individuals studied did not include such severely immunocompromised subjects (the mean absolute CD4 count among the adults studied was 540 T-cells/mm³). Indeed, HIV-infected subjects who are infected with CMV typically have a higher magnitude CMV response than their HIV-uninfected counterparts (Kovacs et al. 1999). For these reasons, I believe the identification of CMV-specific CD8⁺ T-cell epitope would not be substantially affected in such a study cohort. Further evidence for the validity of the overall approach comes from the identification of the previously well-characterised HLA-A*02:01-restricted NLVPMVATV and the HLA-B*07:02-restricted RIPHERNGFTVL epitopes (Weekes et al. 1999), as described above.

Although the work presented here involves a comprehensive study of peptide immunodominance against 3 CMV-encoded proteins in a large African cohort, there were some limitations beyond our control. I was not able to test all subjects for all three CMV proteins due to a lack of PBMC availability. Testing each protein for all subjects could help us identify more novel epitopes in pp65, IE1 and IE2. The

ELISpot assays were performed without CD8+ T-cell enrichment so CD4+ T-cell cross reactivity with the overlapping peptides cannot be ruled out, however, the tetramer staining for some of the epitopes is unequivocal evidence for CD8+ T-cell specificity. The HLA class I associations with recognition of particular peptides is further evidence against these being CD4+ T helper responses. A CMV negative control to test the tetramers would have been desirable but it is not feasible to find a CMV negative subject in an African cohort.

These new findings facilitate further analysis of CMV-specific CD8+ T-cell responses in African populations where most infants are CMV-infected within the first few months of life and approaching 90% are CMV-infected by 1 year of age. CMV is a major cause for concern in sub-Saharan African populations, due to high prevalence of HIV infection in these populations. The immunodeficiency caused by chronic HIV infection influences the consequences of co-infection, increasing the risk of long-term morbidity and mortality. Identification of novel epitopes reported here that are restricted by common HLA alleles could pave the way to immune-mediated prevention and/or therapy of CMV infection.

Chapter 4

TCR bias within the immunodominant HLA-B*44:03-restricted, CMV-specific CD8+ T-cell response

4.1 Introduction

Human cytomegalovirus (CMV) is a ubiquitous and highly prevalent β -herpesvirus (Miles et al. 2008; Cannon et al. 2010; Adland et al. 2015a). Although CMV infection can lead to severe clinical complications in immunocompromised patients such as in the setting of AIDS (Detels et al. 1994; Deayton et al. 2004; Griffiths 2006), in transplant patients on immunosuppressive regimens (Brown et al. 2010; McGoldrick et al. 2013) or following congenital infection (Boppana and Britt 1995; Fowler and Boppana 2006), it is usually asymptomatic. CD8+ T-cells play a central part of protection against CMV disease (Riddell et al. 1992; Walter et al. 1995). Adoptive transfer of donor derived CMV-specific CD8+ T-cells into immunodeficient

recipients effectively restores antigen-specific CD8+ T-cell immunity and reduces CMV viraemia and prevents disease. Conversely, lack of CD8+ T-cell-mediated responses correlates with CMV replication and disease progression. The frequency of CMV-specific CD8+ T-cells following natural infection is unusually high, constituting up to 10% of the CD8+ T-cell memory compartments in peripheral blood in adults (Gillespie et al. 2000; Sylwester et al. 2005; Huygens et al. 2014). These responses are maintained for decades, and in the case of certain specificities may increase in frequency even further over the course of time (Karrer et al. 2003; O'Hara et al. 2012; Klenerman and Oxenius 2016).

The overall impact of the size of this CMV-specific T-cell response on the ability of the immune system to mount effective responses against other pathogens is unclear, but it is apparent that HIV disease progression is substantially more rapid in the case of HIV-CMV co-infection (Kovacs et al. 1999). The contribution of CMV to immunosenescence and increased susceptibility in the elderly to mortality from infectious diseases (Looney et al. 1999) remains a topic of active debate (Pawelec 2014). In view of the fact that the human T-cell receptor (TCRs) repertoire is finite, and limited to approximately 25 million TCRs (Attaf et al. 2015), it is possible that, in individuals in whom 10% or more of CD8+ T-cell responses are dedicated to control of CMV infection, the availability of optimal TCRs to control non-CMV infections might be constrained.

Studies analysing the TCR usage following CMV infection are relatively few, especially in populations where CMV is endemic from infancy (Wynn et al. 2008; Babel et al. 2009; Yang et al. 2015; Chen et al. 2017). In sub-Saharan Africa, most children become CMV infected in the first 3 months of life and CMV infection approaches 90% in infants reaching 12 months of age (Miles et al. 2007; Cannon

et al. 2010). I here studied the TCR repertoire of CD8+ T-cells specific for an immunodominant CMV-specific response in African populations that is restricted by HLA-B*44:03 (Malik et al, 2017[currently under review]). This HLA class I molecule is expressed in approximately 15% of African populations (Kiepiela et al. 2004) and differs by a single amino acid from the closely-related HLA-B*44:02 (Macdonald et al. 2003), the more prevalent HLA-B*44 subtype in White populations. Previous studies have shown that, among individuals expressing HLA-B*44:03, up to 18% of CD8+ T-cells are specific for the IE2 epitope NEGVKAAW (NW8) (IE2 residues 466-473).

In this chapter, I present the TCR deep sequencing data of tetramer-sorted NW8-HLA-B*44:03-specific CD8+ T-cells from 20 HLA-B*44:03+ CMV-infected African individuals. It was observed that the TCR repertoire of NW8-specific CD8+ T-cells displayed limited clonal diversity and usage of restricted set of V and J genes. The response was characterised by the presence of identical or 'public' TCRs in multiple unrelated individuals (Venturi et al. 2008b).

4.2 Materials and methods

4.2.1 Study subjects

20 HIV-infected ART-naive adult subjects from a previously described cohort in Durban (South Africa) (Payne et al. 2014) were studied. The subjects studied had a mean age of 23 years (IQR, 21-31 years), a mean absolute CD4+ T-cell count of 477/mm³ (IQR, 205-107,500) and a median HIV viral load of 20,500 copies/ml (IQR 816-23,000). HIV viral load was measured using the Roche Amplicor version 1.5 assay according the manufacturer's instructions; CD4+ T-cell counts were mea-

sured by flow cytometry. Ethics approval was given by the KwaZulu-Natal Review Board (Durban cohort). All subjects provided written informed consent.

4.2.2 Statistical analysis

TCR repertoire clonality is given by the Shannon evenness index (J'):

$$J' = \frac{H'}{\ln(n)} \quad , \quad H' = \sum_{i=1}^n p_i \ln(p_i) \quad (4.1)$$

Where p_i is the frequency of the i -th clonotype in a population of n clonotypes.

Compositional similarity between TCR repertoires were assessed using the Morisita-Horn similarity index (CMH) (Venturi et al. 2008b):

$$C_{MH} = \frac{\sum_{i=1}^c f_i g_i}{\sum_{i=1}^c (f_i^2 + g_i^2)} \quad (4.2)$$

where $f_i = n_{1i}/N_1$ and $g_i = n_{2i}/N_2$, n_{1i} and n_{2i} are the number of copies of the i -th clonotype in samples 1 and 2, and N_1 and N_2 are the total number of TCRs in samples 1 and 2, respectively. The summations in the numerator and the denominator are over the unique clonotypes (c) in both samples. The similarity indices range in value from 0 (minimal similarity) to 1 (maximal similarity). The Morisita-Horn similarity index accounts for both the number of common clonotypes and the distribution of clone sizes, and it is most sensitive to the clone sizes of the dominant clonotypes. The similarity analyses were performed using the numpy package in Python.

4.2.3 Data analysis

All analyses were performed in Prism v5.0 (GraphPad, San Diego, USA) unless stated otherwise. Strength of association between two variables was analysed by Spearman's rank test. P values <0.05 were considered significant.

4.3 Results

4.3.1 NW8-specific T-cells display limited TCR diversity

The TCR- α and TCR- β chain repertoires of NW8-specific CD8+ T-cells sorted from 20 HLA-B*44:03 positive individuals from Durban, South Africa was deep sequenced. The mean size of the tetramer+ population was 3.88% of CD3+ CD8+ T-cells (IQR 0.36-6.05) and the number of cells sorted for sequencing was a mean of 101,69 cells (IQR 327-117,91) (Table 4.1). In total, 1,650,000 reads from TCR- α chain samples and 500,000 from TCR- β chains were generated using the Illumina MiSeq platform from these samples. This translated into 316,149 functional, in-frame TCR- α chain sequences and 90,864 TCR- β chains.

A total of 65 TCR- α chain clonotypes were identified in 16 individuals and 51 TCR- β clonotypes in 18 individuals (Figs 4.1-4.2). The number of clonotypes per patient varied widely across the cohort, from 1 to 43. Although in part this can be attributed to variation in the number of sorted cells, the number of TCR- α and TCR- β chains clonotypes was not significantly associated with the number of cells sequenced ($r=0.381$ ($p=0.181$), $r=-0.082$ ($p=0.76$) respectively; Spearman correlation). TCR clonality was evaluated using the Shannon evenness index (J'). J' is 0 when the sample is monoclonal and is 1 when all clonotypes have the same frequency. Three out of 16 TCR- α and four out of 18 TCR- β repertoires were strictly

Patient ID	A		B	C	
	HLA-B	HLA-B	tet+CD8+(%)	No of TCR- α seq.	No of TCR- β seq.
Patient 1	44:03	49:01	0.39	1551	3373
Patient 2	44:03	82:02	7.29	67609	N/A
Patient 3	44:03	81:01	0.02	N/A	16716
Patient 4	07:02	44:03	0.44	2328	N/A
Patient 5	44:03	82:02	6.07	57489	5311
Patient 6	44:03	58:01	16.8	6173	1580
Patient 7	44:03	81:01	0.01	2319	1580
Patient 8	35:01	44:03	2.85	5552	1388
Patient 9	15:10	44:03	16.8	3976	2210
Patient 10	42:01	44:03	0.35	N/A	21163
Patient 11	44:03	58:01	2.39	12729	5221
Patient 12	42:01	44:03	1.51	20277	7197
Patient 13	44:03	81:01	7.81	N/A	2252
Patient 14	27:05	44:03	1.3	13746	216
Patient 15	15:01	44:03	6.02	5812	18066
Patient 16	44:03	45:01	0.7	19	1550
Patient 17	42:01	44:03	0.09	N/A	903
Patient 18	44:03	58:01	4.49	47916	1611
Patient 19	08:01	44:03	2.18	48902	237
Patient 20	42:01	44:03	0.08	19751	290

Table 4.1: Subject Characteristics. All subjects (n=20) were recruited from Durban, South Africa, with a mean age of 23 years (IQR, 21-31 years). **(A)** Frequency of tetramer+ cells in CD8+CD3+ T-cells. **(B)** Number of sequencing reads aligning to the trb locus. **(C)** Number of sequencing reads aligning to the tra locus.

monoclonal (Figs 1-2). Other repertoires showed evidence of preferential expansion of a limited set of clonotypes, as suggested by low J values. Thus, both the TCR- α and the TCR- β repertoires on NW8-specific CD8⁺ T-cells displayed evidence of clonotypic bias.

4.3.2 The NW8-specific CD8⁺ T-cell repertoire is characterised by restricted V and J segment usage

Antigen-driven selection of TCR clonotypes often leads to skewed distributions of V and J genes (Moss et al. 1991; Argat et al. 1994; Lehner et al. 1995). NW8-specific responses displayed such bias (Fig 4.3). Out of the 16 individuals sequenced, 14 expressed TRAV20 gene segment. Similarly, TRBV19 gene segment was found in 15 out of the 18 individuals sequenced. Nearly a third of all TCR- α chains were encoded by the TRAV20 gene segment (32% of sequences) (Fig 4.3A). Similarly, TCR- β chains were heavily biased towards TRBV19 (43% of sequences) (Fig 4.3B). As expected, certain V-J gene combinations were more frequent than others. Among TCR- α chain repertoires (Fig 4.3C), the TRAV20-TRAJ39 and TRAV5-TRAJ7 pairs were highly prevalent (34% and 15% of all pairings, respectively). Among TCR- β sequences (Fig 4.3D), TRBV19-TRBJ2.2 and TRBV20.1-TRBJ1.4 were common pairings (27% and 21%, respectively). Thus, the NW8-specific CD8⁺ T-cell repertoire is characterised by restricted V and J segment usage.

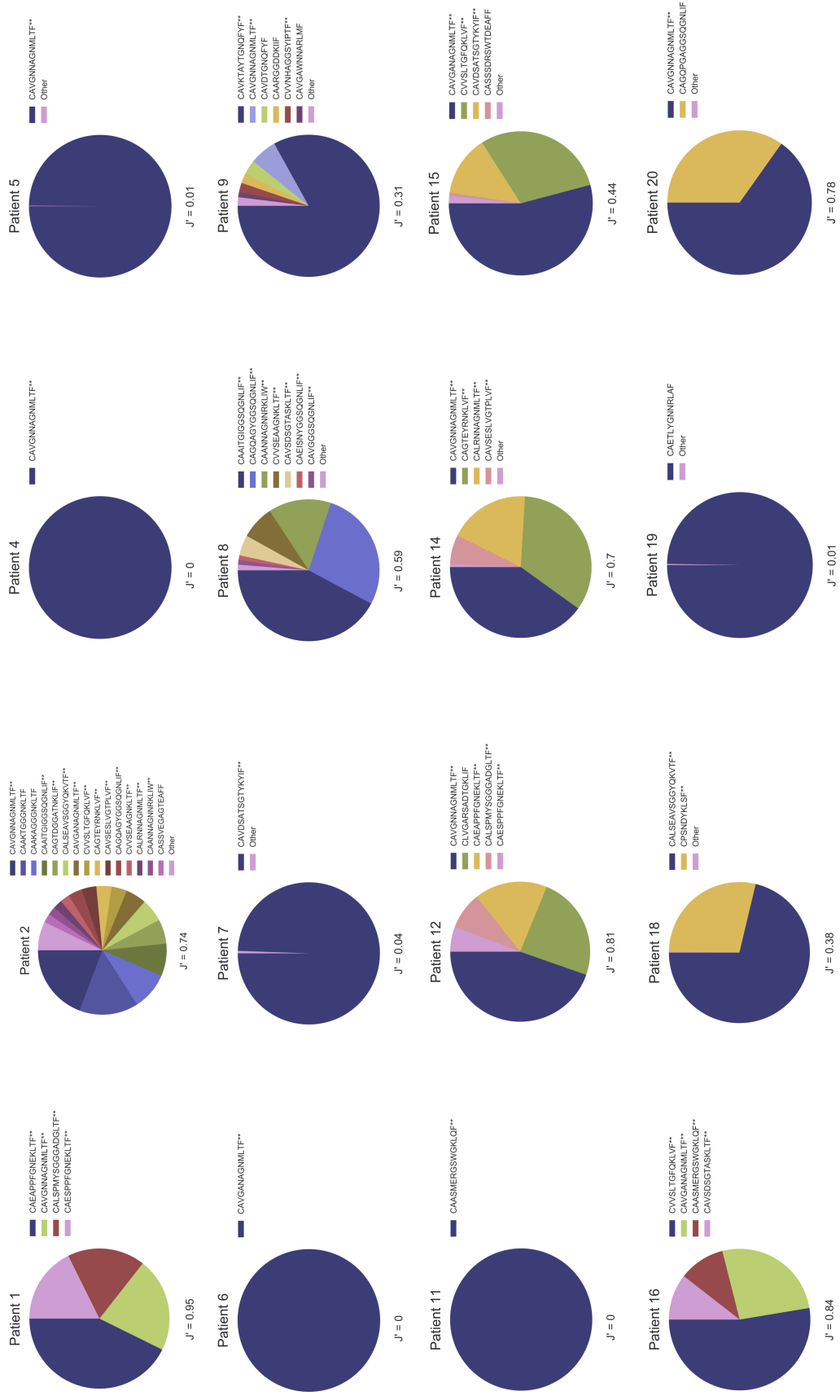


Figure 4.1: Distribution of TCR- α chain clonotypes in HLA-B*44:03 NW8-specific CD8+ T-cells. The frequency of TCR- α chain amino acid sequences obtained from (n=16) CMV-infected individuals are shown as pie charts. Frequencies are calculated as a percentage of aligned sequencing reads. Public clonotypes are shown with asterisks (**). Shannon's evenness index (J') is shown under each pie chart. Colours are assigned randomly and do not correspond to a fixed a sequence.

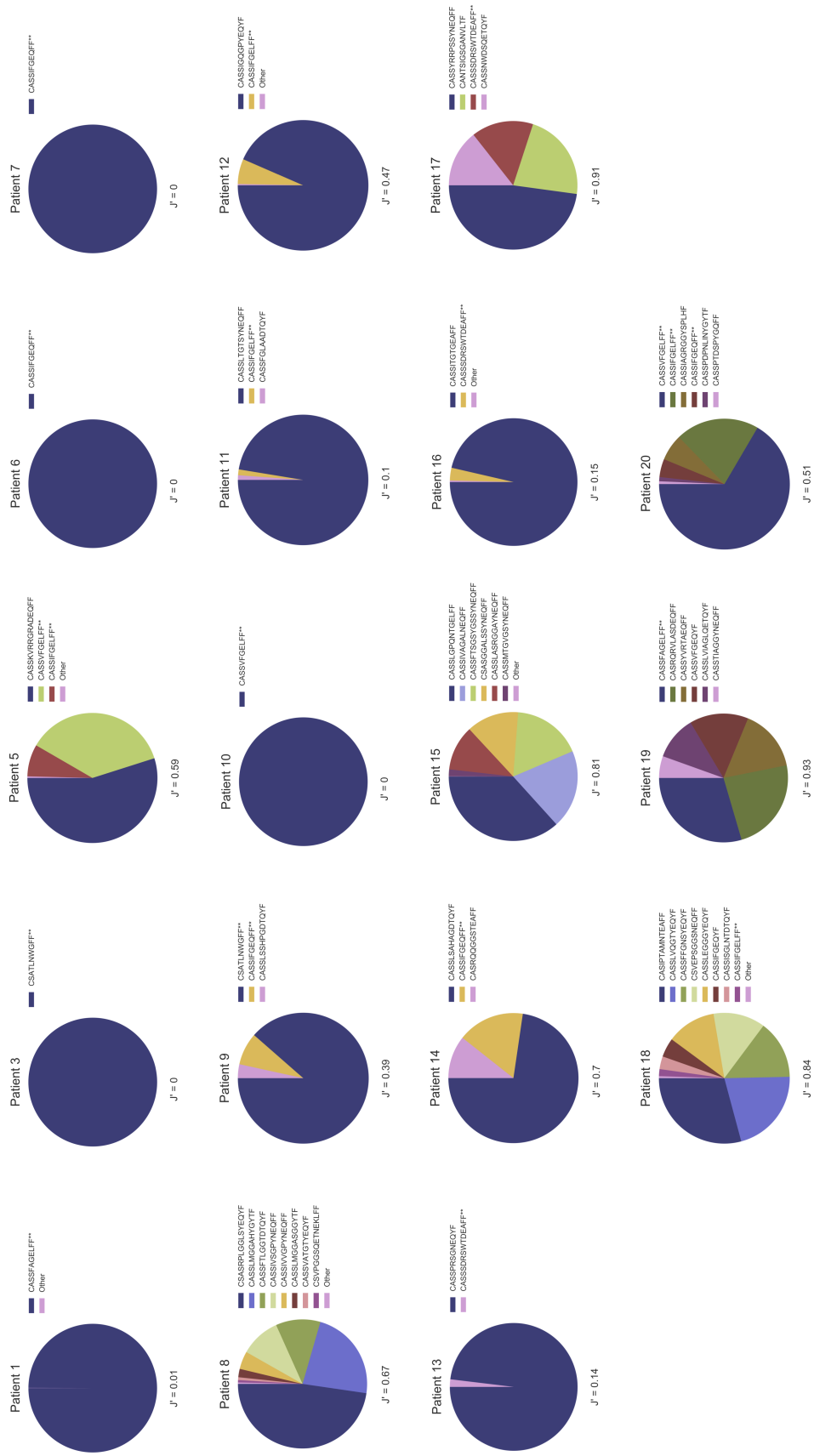


Figure 4.2: Distribution of TCR-β chain clonotypes in HLA-B*44:03 NW8-specific CD8+ T-cells. The frequency distribution of TCR-β chain amino acid sequences obtained from (n=18) CMV-infected individuals is shown as pie charts. Frequencies are calculated as a percentage of aligned sequencing reads. Public clonotypes are shown with asterisks (**). Shannon's evenness index (J') is shown under each pie chart. Colours are assigned randomly and do not correspond to a fixed a sequence.

4.3.3 TCR sharing in NW8-specific T-cell responses

The Morisita-Horn similarity-index (CMH) (Venturi et al. 2008a) was used to perform pairwise comparisons between every combination of two donors, as a measure of compositional similarity, or overlap (Fig 4.4). CMH values lie between 0 (absolutely different populations; shown in white colour) and 1 (identical populations; shown in black colour). The CMH value for every combination of two donors (120 pairwise combinations for TCR- α chain and 153 pairwise for TCR- β chain calculated using equation 4.3) was calculated and the mean CMH value of the TCR- α chain repertoire was 0.20 ± 0.02 (Fig 4.4A) and the mean CMH value of the TCR- β repertoire was 0.09 ± 0.1 (Fig 4.4B). These results indicated that there was a degree of overlap between NW8-specific repertoires and that the TCR- α chain repertoires were more similar between individuals compared to the TCR- β repertoires (0.20 ± 0.02 vs 0.09 ± 0.01 $P < 0.0001$; Mann-Whitney U test).

$$x = n(n - 1)/2 \quad (4.3)$$

4.3.4 High-frequency TCR- α chains are highly shared amongst HLA-B*44:03-infected individuals

Multiple public TCRs shared amongst different subsets of individuals were identified (Fig 4.1 and 4.2). 35 out of 52 TCR- α chain clonotypes and 6 out of 48 TCR- β chains were public. The public clonotype, CAVGNNAGNMLTF, was the most abundant clone (30.4% of all clones sequenced). This and another closely related clonotype, CAVGANAGNMLTF (both TRAV20/TRAJ39), were also the most extensively shared. CAVGNNAGNMLTF was found in 12/16 of subjects studied and CAVGANAGNMLTF in 7/16. Similarly, the most abundant public TCR- β clonotypes, CASSIFGEQFF and CASSIFGELFF (TRBV19/TRBJ2-1 and TRBV19/TRBJ2-2,

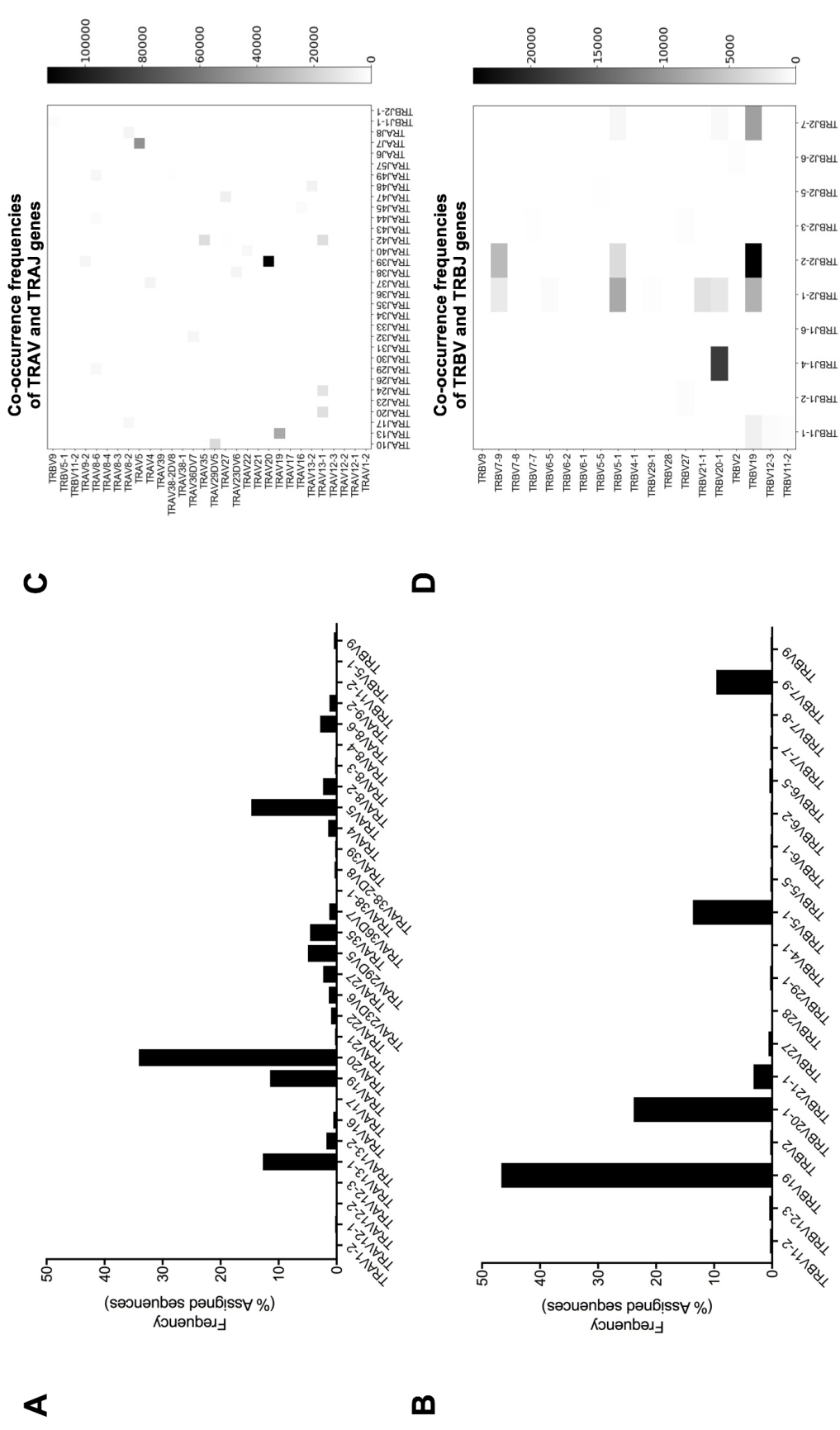


Figure 4.3: V and J gene usage in TCR- α and - β chain repertoires of HLA-B*44:03/NW8-specific CD8+ T-cells. V gene usage is shown for (A) TCR- α chains (n=16 individuals) and (B) TCR- β chains (n=18). The frequency of V genes found in NW8-tetramer-sorted cells is expressed as a frequency of all assigned sequences. V-J pair frequencies are represented as heat maps for (C) TRAV-TRAJ combinations and (D) TRBV-TRBJ combinations.

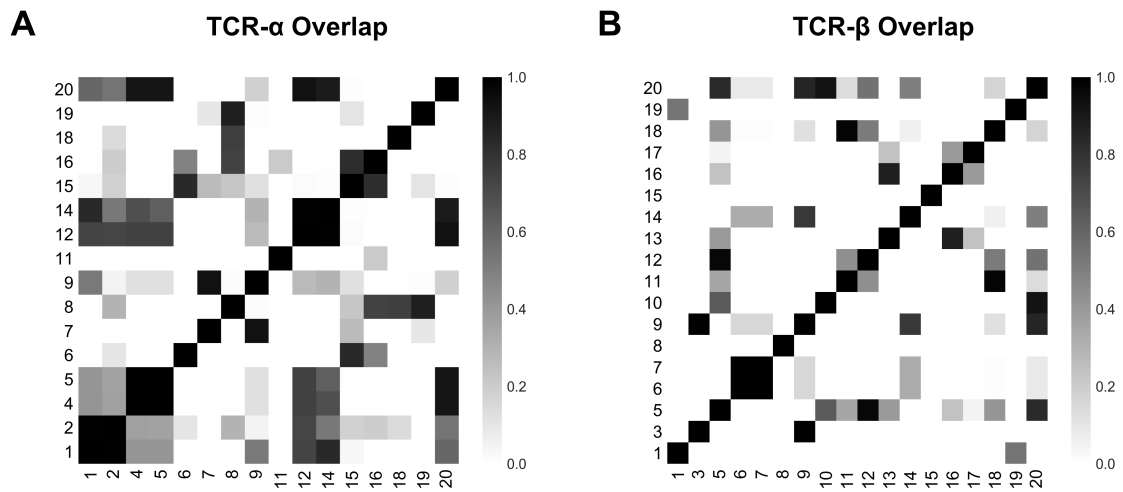


Figure 4.4: TCR sharing amongst NW8-specific CD8⁺ T-cells from CMV-infected individuals. Compositional similarity amongst (A) TCR- α and (B) TCR- β chains was assessed using the Morisita-Horn (CMH) index and illustrated as a heat map. Patient IDs are shown along the x and y-axes. CMH values range from 0 (no overlap) to 1 (perfect overlap).

respectively), were shared by 6/18 and 5/18 of individuals studied, respectively. As expected, and as previously observed (Venturi et al. 2008b), the higher the frequency of a given clonotype, the greater the likelihood that this would be a public TCR sequence. The correlation between clonotype frequency and publicity of the TCR was statistically significant for the TCR- α and TCR- β chains (Fig 4.5).

4.3.5 Public NW8-specific TCRs are produced by convergent recombination

The public amino acid sequences were often encoded by more than one distinct nucleotide sequence. For example, the most prevalent public TCR- β clonotype, CASSIFGEQFF, was encoded by 4 different nucleotide sequences. Similarly, the most frequently occurring TCR- α public clonotype, CAVGNNAGNMLTF, was encoded by 4 different nucleotide sequences. The clonotypes that occurred in a greater number of individuals tended to also be encoded by a greater number of

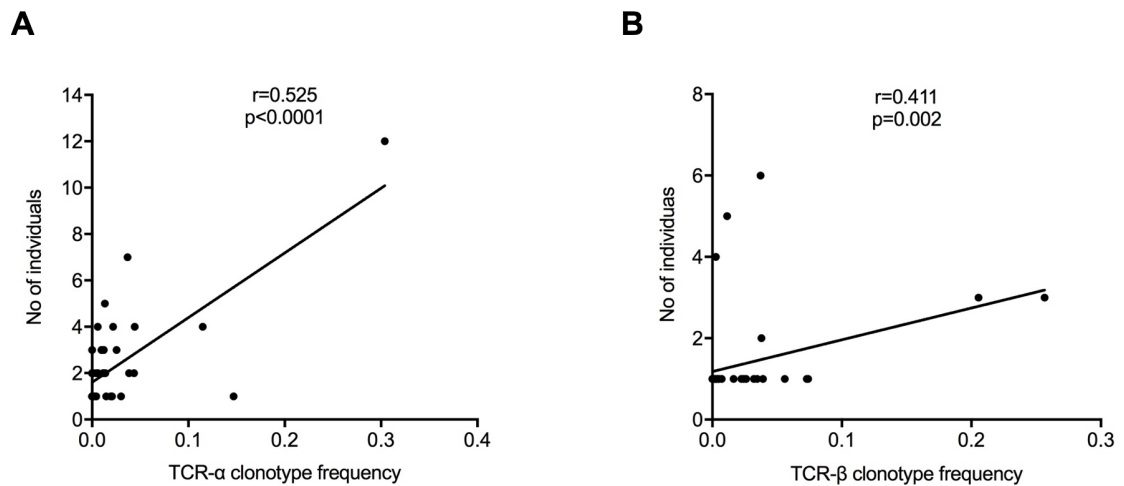


Figure 4.5: Clonal dominance and TCR publicity. The frequency of a given clonotype is plotted against the number of individuals sharing that clonotype, for **(A)** TCR- α chains ($n=65$) and **(B)** TCR- β chains ($n=51$). Spearman's rho and p-values are shown above each graph.

distinct nucleotide sequences. For TCR- β chains, this association was statistically significant ($r=0.798$, $p<0.0001$, Fig 4.6). This may be explained by recombinant convergence, whereby multiple nucleotide sequences can encode the same CDR3 amino acid sequence (Venturi et al. 2006). This is further evidence of the strong selection for particular TCR amino acid sequences by the HLA-B*44:03-NW8 peptide-MHC complex.

4.3.6 Public CDR3 motifs are strictly conserved amongst HLA-B*44:03-positive individuals

TCR repertoire bias often includes the conservation of amino acid motifs within CDR3 loops (Venturi et al. 2006; Li et al. 2012). Several TCR- α chain clonotypes contained the sequence motif Asn-Ala-Gly (referred to as “NAG” thereafter) or Gly-Gly-Ser (“GGS”) were observed. TCR- α chains with the “NAG” motif were shared across more individuals than those with no motif ($p<0.0001$) or with the “GGS” motif ($p<0.0011$) (Fig 4.7A). This also held true for TCR- β chains contain-

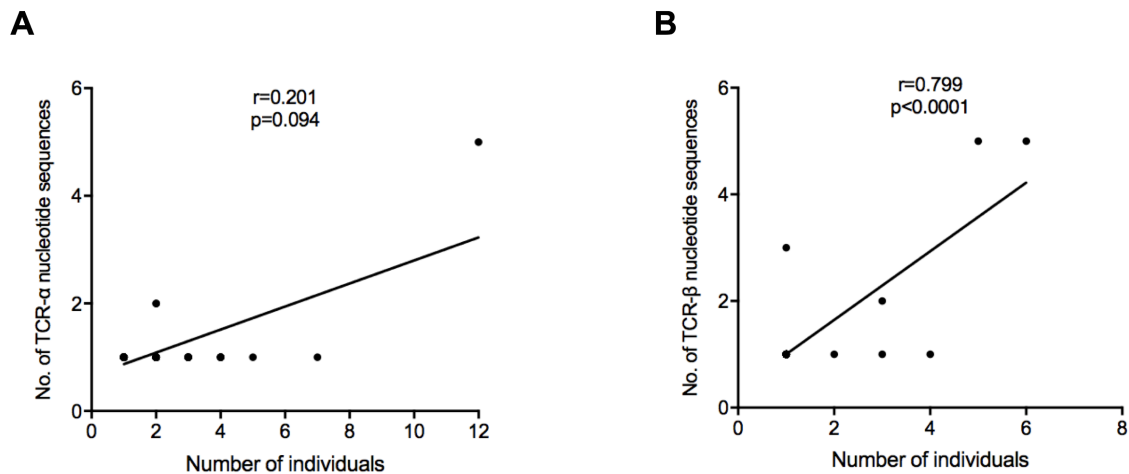


Figure 4.6: Number of nucleotide sequences and TCR publicity. The number of nucleotide sequences encoding a given clonotype is plotted against the number of individuals sharing that clonotype, for **(A)** TCR- α chains and **(B)** TCR- β chains. Spearman's rho and p-values are shown above each graph.

ing the Ile-Phe-Gly motif (“IFG”, Fig 4.7B). Most TCR chains with a conserved motif were public. Conversely, private TCR chains displayed no apparent motif. The “NAG” motif was entirely encoded by the TRAJ region, specifically the TRAJ39 segment in 93.5% of clonotypes. Similarly, the “GGS” motif was largely encoded by TRAJ42. Strikingly, V gene usage was not conserved. “NAG” clonotypes were encoded by three different TRAV genes and “GGS” by as many as seven TRAV genes (Fig 4.7C). This excludes the possibility of a germline-encoded interaction between TCR- α CDR1 and -2 loops with MHC and suggests a more prominent role for CDR3 loops in NW8 recognition.

TCR- β chains clonotypes with the “IFG” motif were strictly encoded by TRBV19, while TRBJ usage was less conserved (Fig 4.7D). The central phenylalanine residue at position 6 was not germline-encoded in all cases and resulted from two codons. The conservation of a non-germline residue again argues for structurally-imposed, antigen-driven selection of clonotypes bearing this amino acid at this position.

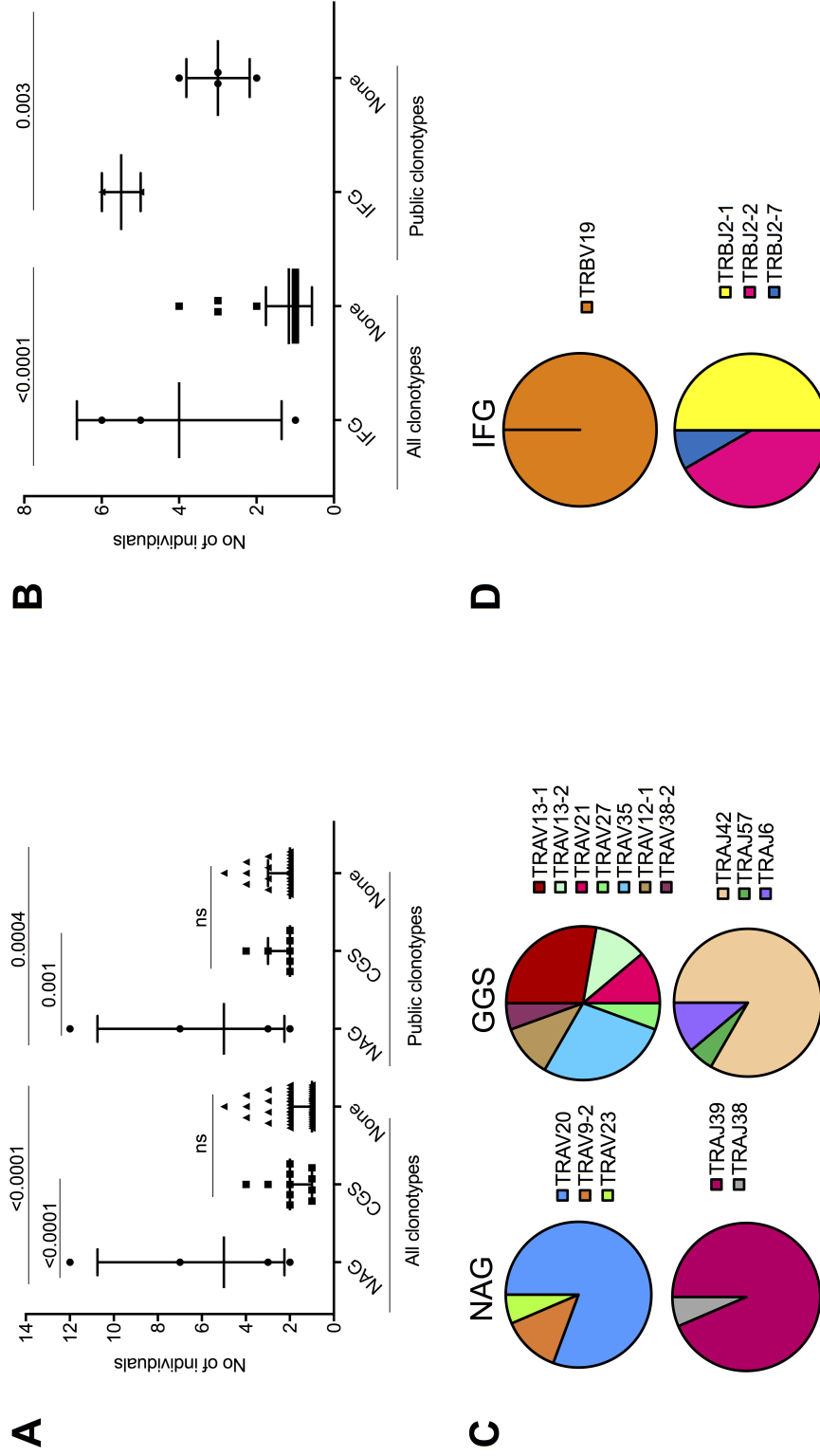


Figure 4.7: Amino acid motif conservation in public clonotypes from HLA-B*44:03 NW8-specific CD8+ T-cells. (A) The number of individuals sharing “NAG” or “GGS”-containing TCR- α chains is shown. **(B)** The number of individuals sharing “IFG”-containing TCR- β chains is shown. **(C)** TRAV and TRAJ gene usage in “NAG” or “GGS”-containing TCR- α chains is illustrated as pie charts. Frequencies are calculated as fraction of “NAG” or “GGS”-containing amino acid sequences. **(D)** TRBV and TRBJ gene usage in “IFG”-containing TCR- β chains is illustrated as pie charts. Frequencies are calculated as fraction of “IFG”-containing amino acid sequences. Horizontal bars represent the median and IQR. Statistical significance is shown where $p < 0.05$ (ns = not significant). Pie charts illustrate the frequency of a V or J gene. Colours are assigned randomly and do not correspond to a fixed a sequence.

4.4 Discussion

In this chapter, I sought to characterise the TCR repertoire of CD8⁺ T-cells responding to an immunodominant CMV epitope in HLA B*44:03⁺ individuals. Using high throughput sequencing, the NW8-specific TCR- α and - β chain repertoires of a total of 20 CMV-infected subjects was dissected. These repertoires were found to be highly skewed, as indicated by preferential use of V and J genes, extensive TCR sharing and conservation of CDR3 motifs.

TCR diversity and the specificity it underlies are the distinguishing feature of $\alpha\beta$ T-cells. Diversity in T-cell responses has long been thought of as a strict requirement for the provision of protective immunity and full immune coverage. Nonetheless, it is now well established that human viral infections can give rise to narrow and skewed repertoires, often associated with TCR sharing in different individuals. Thus, the repertoire is often shaped towards a single or focused solution for the recognition of peptide-HLA complexes (Li et al. 2012; Venturi et al. 2008b). Reports of TCR publicity in immune responses have accumulated in the literature, such that TCR sharing across individuals can now almost be taken as a rule, rather than the exception.

The occurrence of public TCRs is nearly 10^3 times more likely than expected if all rearrangements were equiprobable, irrespective of HLA type (Robins et al. 2010). A TCR sequence will indeed have a higher likelihood of being produced if it is encoded by a nucleotide sequence with few additions and deletions or if multiple nucleotide sequences give rise to the same TCR sequence. It has been suggested that V(D)J recombination events 'converge' to give rise to the same TCR (Venturi et al. 2008b; Miles et al. 2010; Venturi et al. 2006). There are three types of con-

vergent recombination events. The first involves different splicings of the germline V(D)J gene segments with random addition of nucleotides that “converge” with the same resultant nucleotide sequence. Second, many different nucleotide sequences “converge” to encode the same amino-acid sequence. The last level of convergent recombination is at the level of TCR repertoire in which the V(D)J recombination process facilitate some of the amino acids to conform to the amino-acid motif which are more frequently associated with public T-cell responses.

Redundancy in the genetic code also explains why TCR chains containing residues encoded by multiple codons (such as Gly or Ala) are very commonly produced and can occur at high frequency (Venturi et al. 2006; Li et al. 2012).

TCR production frequency is an important determinant of TCR sharing (Venturi et al. 2008b). Indeed, public chains described in this study were encoded by convergent recombination and the public chains were nearly entirely germline-encoded or involved very limited changes at the CDR3 junction. This was effectively mirrored by the conservation of amino acid motifs. The TCR- α chain “NAG” motif described here was found to be entirely encoded by the J segment, and the “IFG” motif found in TCR- β chains only contained one added amino acid.

The germline nature of these public chains again highlights the importance of TCR generation efficiency. In the case of “NAG” TCR- α chains, the conservation of the J segment, but not the V segment, excludes a potential role for CDR1 and CDR2 in recognition of HLA-B*44:03. Rather, this emphasises the starring role of the CDR3 loop in antigen recognition and structural constraints favouring the selection of public clonotypes with these conserved motifs.

The other distinguishing feature of NW8-specific repertoires was their extremely limited diversity. In some individuals, TCR- α and TCR- β chain repertoires were even monoclonal. This was notably the case of patient 6, with a tetramer response of 16.8% (equivalent to 31,480 sorted cells) composed of a single TCR, CAVGANAGMLTF-CASSIFGEQFF. This, to our knowledge, is the largest reported T-cell response mounted by a single public T-cell clone in humans, which initially led us to envisage a positive relationship between TCR publicity and the magnitude of the T-cell response. However, this was not the case, as the number of public TCRs did not correlate with the percentage of tetramer-positive cells or with the number of sorted cells. Instead, the frequency of a given clonotype positively correlated with the extent to which it is shared (i.e. the number of individuals harbouring that clonotype). Altogether, these results suggest that publicity in T-cell responses does not necessarily equate to superior immunity. Rather, the dominance of public TCRs within an individual and across individuals can be explained merely by their high precursor frequency, or in other words, by a numerical competitive advantage.

Intuitively, that highly prevalent pathogens such as CMV are best dealt with by TCRs which are highly prevalent across a population appears to be a beneficial trait. Whether public TCRs represent a built-in bank of receptors which can readily be mobilised against pathogens remains to be determined. Understanding the molecular rules which govern how TCRs are generated and selected to seed the periphery will be crucial in our understanding of public T-cell responses.

Chapter 5

Studying changes in CMV-specific TCR repertoires in mother-child pairs over a span of 10+ years

5.1 Introduction

Changes in the TCR repertoire of antiviral CD8⁺ T-cell responses over the course of an infection have been studied for various systems in humans. However, most of these studies have been done before the next generation sequencing era, limiting their scope. Primary infection with HIV results in oligoclonal expansions with the use of a limited subset of V gene segments (Pantaleo et al. 1994). Longitudinal studies reveal that HIV-1-specific CD8⁺ T-cell repertoires are characterised by changes in clonal composition, with expansions and contractions of distinct clonotypes (Pantaleo et al. 1997; Meyer-Olson et al. 2004; Costa et al. 2015). Another study, however, reported that the HIV-specific TCR repertoire is stable and distinct HIV-specific clonotypes can persist for many years (van Bockel et al. 2011).

Similarly, primary infection with EBV also results in oligoclonal expansions, with restricted TCR usage (Callan et al. 1996). Some longitudinal repertoire studies have indicated that the EBV-specific dominant CD8⁺ T-cell clones are stable over time (Silins et al. 1996; Levitsky et al. 1998), whereas another suggests that there are changing patterns of clonal composition over the first two years following EBV infection (Annels et al. 2000).

Relatively little is known about the changes in the epitope-specific T-cell repertoire following CMV infection, particularly in children. CMV infection is characterised by marked oligoclonal, or in some cases monoclonal, expansions, with restricted TCR usage, as reported earlier (Khan et al. 2002; O'Hara et al. 2012) and also as shown in chapter 4. One longitudinal study carried out on CMV-infected individuals, using the traditional clonal sequencing method, revealed a high degree of clonal focusing, with selection of public TCRs, following primary infection. These were maintained long-term, with some being detectable 5 years after primary infection (Day et al. 2007). Similarly, in other studies, CMV-specific repertoires were shown to be highly stable, some clones detected early during primary infection being maintained at high frequencies even 5 years later (Iancu et al. 2009; Klarenbeek et al. 2012). In a study of elderly individuals (nonagenarians), the clonal compositions of CMV-specific CD8⁺ T-cells were reported to be stable over a time period of two years (Hadrup et al. 2006).

The HLA-B*44:03-restricted NW8 epitope elicits a potent CD8⁺ T-cell response in CMV-infected individuals, which allows easy identification of epitope-specific CD8⁺ T-cells, thus facilitating longitudinal analysis. In this chapter, I studied the changes in TCR usage of HLA-B*4403-restricted CMV-specific NW8 response over time in mother-child pairs, with the aim of addressing the following questions:

1. Is TCR usage stable over time following CMV infection?

Hypothesis: The CMV-specific TCR usage will remain stable over time following CMV infection.

2. Is TCR usage more stable over time in adulthood, decades after infection occurred, than in childhood?

Hypothesis: The TCR usage in childhood will undergo variations, whereas in mothers, the TCR usage will remain largely stable.

3. Does the TCR usage fluctuate randomly over time or is there a clear tendency for particular public TCRs to expand over time and dominate the repertoire?

Hypothesis: Public TCRs will expand over time and dominate the repertoire.

4. Is there any difference in the TCR usage (the CDR3 sequences) between children (as a group) and the mothers (as a group)?

Hypothesis: The overall TCR-CDR3 sequence usage in children will harbour fewer public TCRs compared to those in mothers.

5. Are particular mother-child pairs more likely to share public TCRs compared to unrelated individuals?

Hypothesis: The TCR repertoire of related mother-child pairs will show greater degree of overlap in comparison to unrelated individuals.

In order to address these questions, the NW8-epitope tetramer was used to stain and sort virus specific CD8⁺ T-cells in four mother-child pairs over the span of 10+ years, followed by high-throughput sequencing.

5.2 Materials and methods

5.2.1 Study subjects

We studied four mother-child pairs over the course of 10+ years (Table 5.1). All the children and their mothers were of sub-Saharan African origin. The study participants were HIV infected, ART-naive individuals. The table 5.1 shows the age, HLA type, viral load, CD4 T-cell count and CD3+ CD8+ tet+ positive percentages measured at different time points for children and their mothers.

5.2.2 TCR sequencing

TCR sequencing and analysis were conducted using the techniques detailed in section 2.11.

5.3 Results

5.3.1 Analysis of NW8-specific CD8+ T-cell responses

The magnitude of the NW8-specific CD8+ T-cell response in children (age 1-11 years, n=5) and their mothers (age 17-44 years, n=4) was assessed with the NW8-tetramer (Table 5.1). In addition to the four mother-child pairs, in one family there was an HIV-uninfected, HLA-B*44:03+ sibling of the child that was also studied. All of the nine HLA-B*44:03+ donors studied made a detectable CMV-specific CD8+ T-cell response, with frequencies ranging from 0.8 to 6.32% of the total CD8+ subset (mean 1.7%) (Table 5.1). CD8+ T-cell responses varied between the early and late time points without a common tendency to increase or decrease.

Subjects	HIV status	Sampling date	D.O.B	Age (yrs)	CD4 count (cells/mm ³)	CD4 (%)	VL (RNA copies/ml)	HLA-A	HLA-A	HLA-B	HLA-B	HLA-C	HLA-C	CD3+CD8+tet+ (%)
Mother-1	infected	21/07/2017	09/07/1973	44.0	n/a	n/a	n/a	A*02:05	A*33:01	B*42:01	B*44:03	C*02:10	C*17:01	1.420
Child-1	infected	21/07/2017	06/09/2005	11.9	923	36	63,000	A*02:05	A*29:02	B*42:01	B*44:03	C*02:10	C*17:01	0.740
		25/05/2006		23.0	518	21	2,940							3.890
Mother-2	infected	18/02/2008	14/05/1983	24.8	427	17	n/a	A*29:02	A*30:04	B*15:10	B*44:03	C*02:02	C*03:04	3.060
		13/10/2009		26.4	271	16	n/a							2.440
		19/08/2013		30.3	654	25	< 20							4.870
		15/08/2006		1.7	1,716	28	< 50							6.320
Child-2	infected	18/06/2007	23/11/2004	2.6	1,032	33	< 50	A*02:05	A*30:04	B*42:01	B*44:03	C*02:02	C*17:01	2.990
		19/01/2010		5.2	447	16	13,340							1.440
		19/08/2013		8.7	850	32	< 20							3.130
		27/06/2006		27.7	257	26	256							2.370
		30/01/2008		29.2	318	20	< 50							3.060
Mother-3	infected	27/08/2013	01/11/1978	34.8	319	26	10,000	A*24:02	A*29:02	B*07:02	B*44:03	C*07:01	C*07:02	1.270
		03/06/2016		37.6	536	36	n/a							0.441
		15/05/2017		38.5	n/a	n/a	n/a							0.673
		18/09/2006		1.5	1,349	35	< 50							0.002
		22/07/2008		3.4	611	33	< 50							2.320
		20/07/2010		5.4	772	39	< 20							0.212
Child-3	infected	17/07/2012	01/03/2005	7.4	917	33	< 20	A*29:02	A*68:02	B*14:01	B*44:03	C*07:01	C*08:02	0.133
		19/08/2014		9.5	1,077	38	< 20							0.081
		03/06/2016		11.3	833	36	1,300							0.225
		20/02/2017		12.0	685	33	< 20							0.102
		05/04/2005		17.6	733	28	n/a							1.810
		18/04/2007		19.7	598	31	n/a							1.180
		09/09/2008		21.0	642	30	n/a							2.370
Mother-4	infected	30/11/2010	22/08/1987	23.3	n/a	n/a	n/a	A*03:02	A*29:02	B*35:01	B*44:03	C*04:01	C*07:01	1.210
		28/01/2014		26.4	582	22	1,100							1.140
		30/10/2015		28.2	533	31	< 20							0.836
		28/01/2016		28.4	667	27	< 20							1.110
		03/11/2005		1.1	1,534	44	< 50							3.610
		31/05/2006		1.7	1,209	37	< 50							3.420
Child-4	infected	08/06/2010	23/09/2004	5.7	463	17	13,579	A*29:02	A*30:02	B*41:02	B*44:03	C*07:01	C*17:01	1.720
		28/07/2014		9.8	532	28	< 20							1.220
		30/10/2015		11.1	604	26	340							2.650
Sibling-4	uninfected	13/08/2013	05/05/2006	7.3	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	0.216

Table 5.1: Table 1. Subject Characteristics.

5.3.2 Characteristics of NW8-repertoires in children and adults

The TCR- α and - β chain repertoires of NW8-specific CD8+ T-cells were sorted and sequenced from four mother-child pairs and one HIV-negative sibling, all expressing HLA-B*44:03. In total, 1,650,000 reads from TCR- α chain samples and 900,000 from TCR- β chains were generated using the Illumina MiSeq platform from these samples. This translated into 605,219 functional, in-frame TCR- α chain sequences and 339,573 TCR- β chains. A total of 76 TCR- α chain clonotypes and 132 TCR- β clonotypes were identified.

Skewing of the TCR repertoire after the acute/early CMV infection, and long-term persistence of a number of CMV-specific clones has been reported in the past (Day et al. 2007; Klarenbeek et al. 2012). Our data also reflected that most repertoires were preferentially expanded with a limited set of clonotypes, which was also shown previously in chapter 4 (Fig 5.1). The NW8 CMV-specific TCR repertoire was invariably oligoclonal, often with a single TCR- α / β chain comprising over 50% of the repertoire of several individuals (Figure 5.1A). In general, a single primary TCR- α / β chain dominates more than half of the repertoire of an individual. In some cases, there seems to be a secondary TCR- α / β chain that acts in competition to the primary chain (Fig 5.1). For example, in Child-2, TRBV12-4/TRBJ2-3 and TRBV19/TRBJ2-1 appear to compete with each other and in Mother-2, TRAV1-2/TRAJ10 similarly competes with TRAV4/TRAJ6.

Figure 5.1A: Mother-Child Pair 1

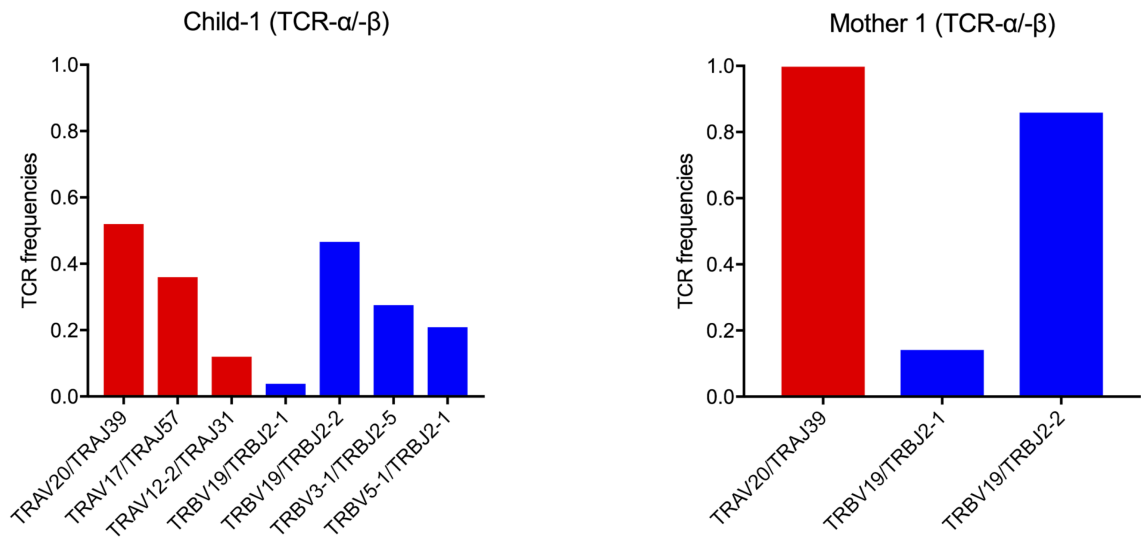


Figure 5.1B: Mother-Child Pair 2

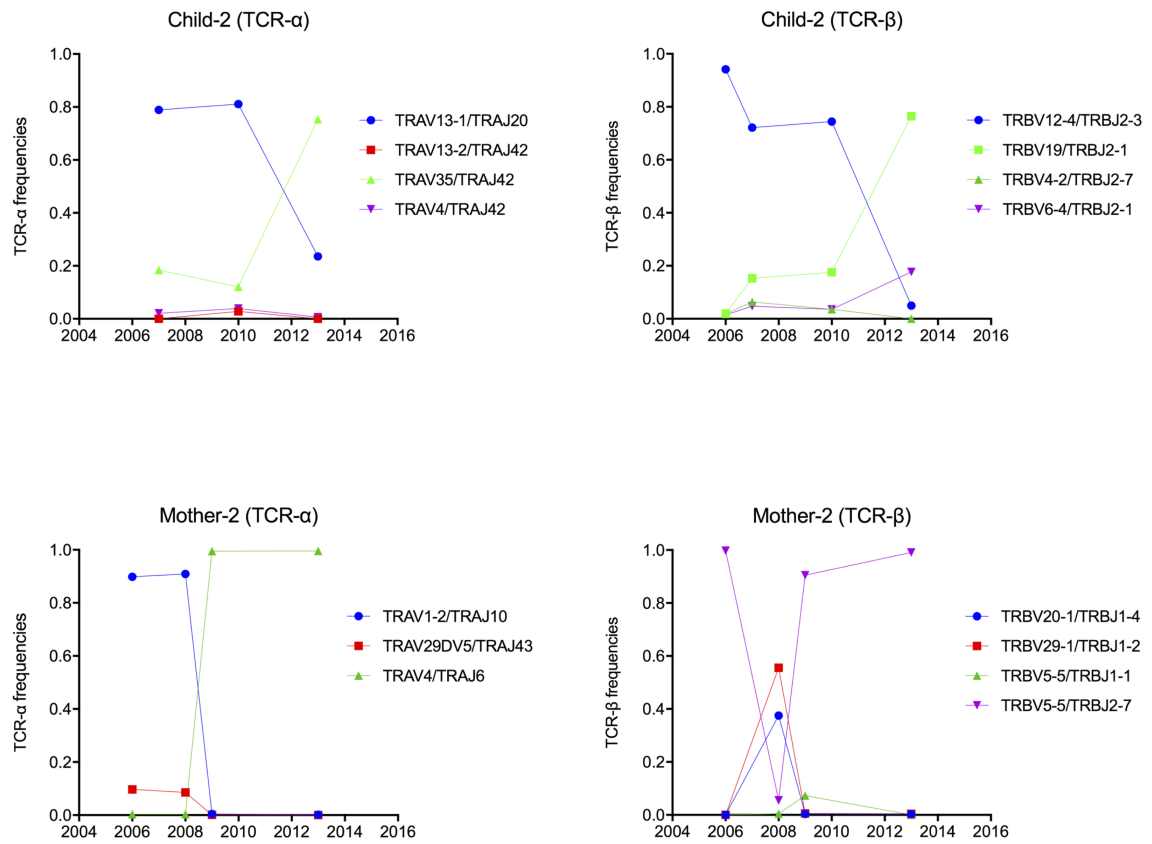


Figure 5.1C: Mother-Child Pair 3

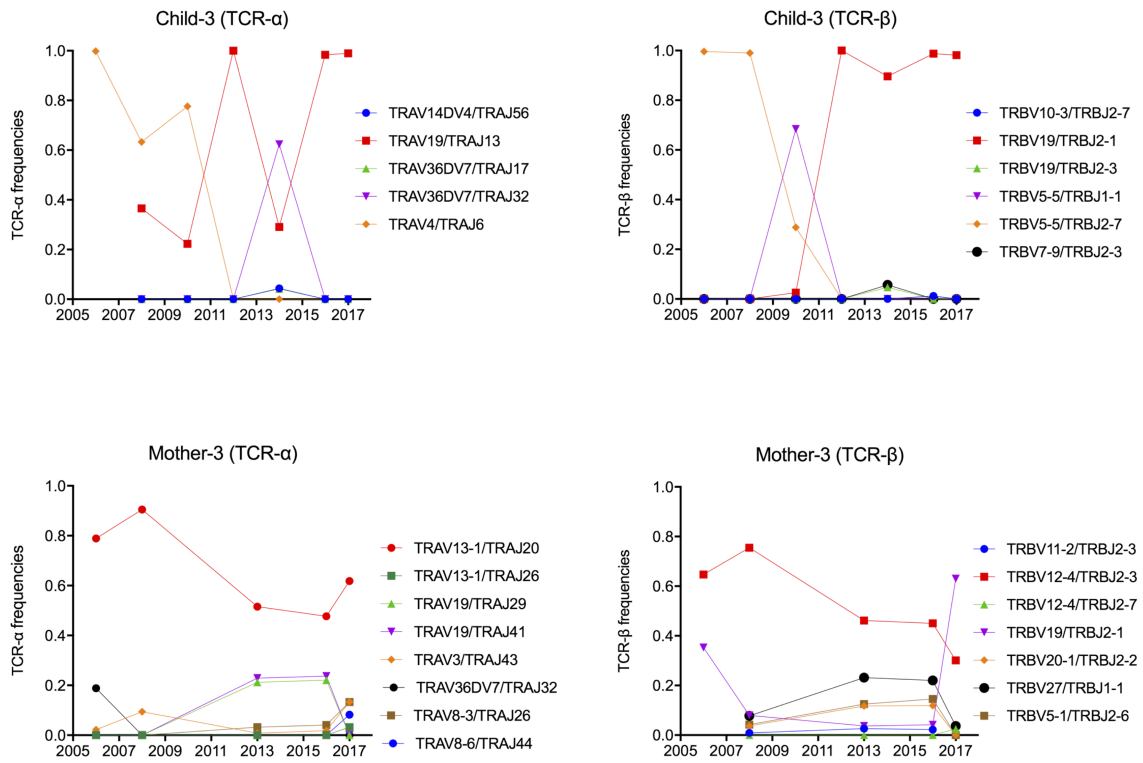


Figure 5.1D: Mother-Child Pair 4

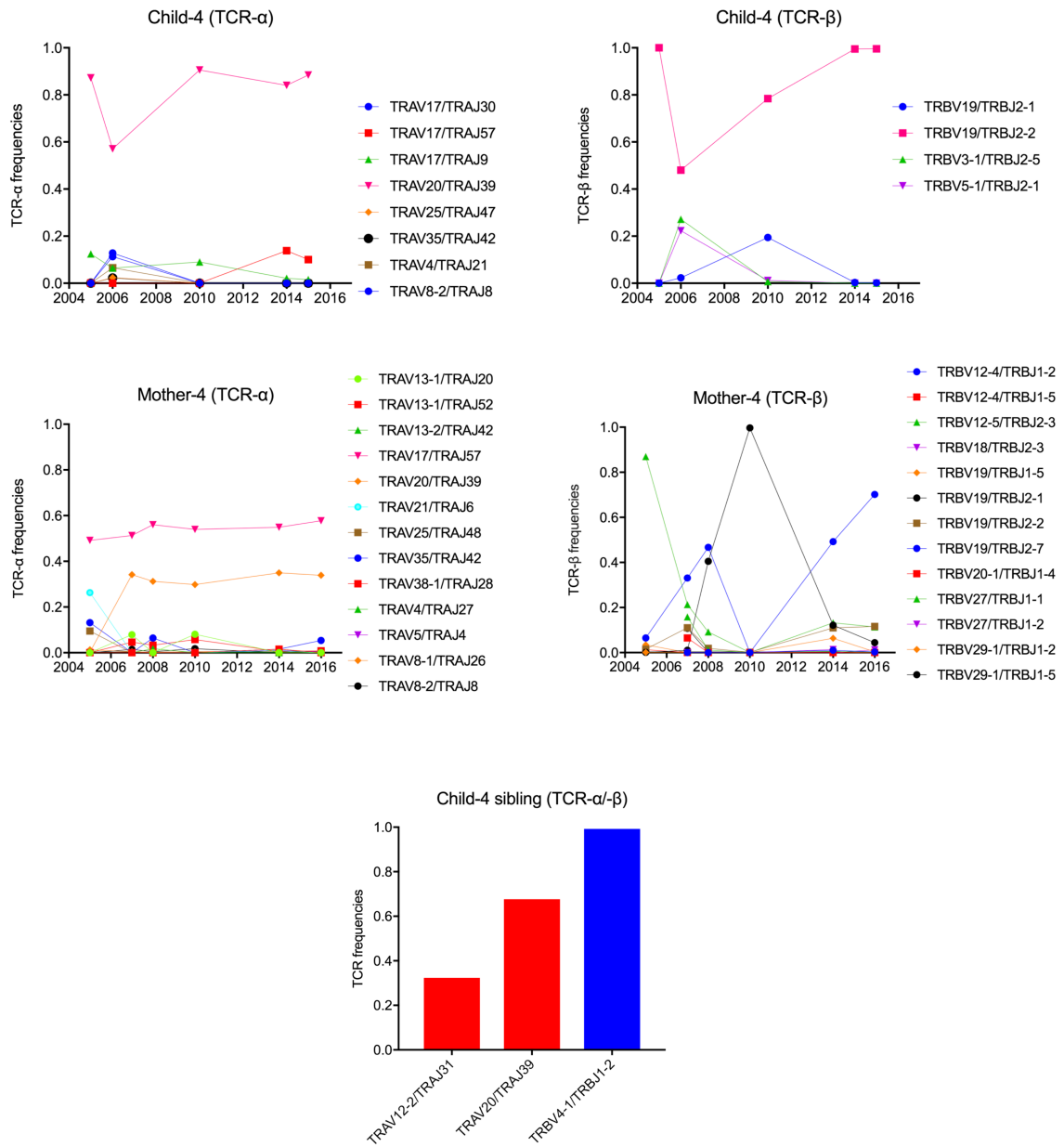


Figure 5.1: Distribution of TCR- α and TCR- β chains in mother-child pairs over the course of 10+ years. Frequency of TCR- α and - β chains obtained from children (n=5) and their mothers (n=4). **(A)** Mother-Child Pair 1, **(B)** Mother-Child Pair 2, **(C)** Mother-Child Pair 3, **(D)** Mother-Children Pair 4. Frequencies are calculated as a percentage of aligned sequencing reads.

Consistent with the data in chapter 4, we found a skewed distribution of V and J genes. Out of the nine individuals sequenced, all nine expressed the TRAV20 gene segment. Similarly, the TRBV19 gene segment was found in all nine out of the nine individuals sequenced. As expected, certain V-J gene combinations were more frequent than others. Among TCR- α chain repertoires (Fig 5.2A), the TRAV20-TRAJ39 pairs were highly prevalent (22% of all pairings, respectively). Among TCR- β chains (Fig 5.2B), TRBV19-TRBJ2.2 and TRBV19-TRBJ2.1 were common pairings (18% and 17%, respectively). Discordant patterns of J and V gene usage were apparent within each subject. For example, the TRAV17/TRAJ57 genes in Mother-4 showed minimal changes in clonotypic representation over time, whereas the rest of the TRAV and TRAJ genes varied.

In some cases, the changes in frequency of an α chain over time were mirrored by the by a β chain (Fig 5.3), suggesting that these particular α and β chains are pairs. Based on this observation, the following TCR- α and - β chain pairs were inferred (Fig 5.3):

- TRAV20/TRAJ39 (CAVGNNAGNMLTF) with TRBV19/TRBJ2-2 (CASSIFGELFF) in Mother-1, Child-1 and Child-4;
- TRAV13-1/TRAJ20 (CAARANDYKLSF) with TRBV12-4/TRBJ2-3 (CASTERGVLTDTQYF) in Child-2; and with TRBV12-4/TRBJ2-3 (CASSLRGVWTDQYF) in Mother-3;
- TRAV35/TRAJ42 with TRBV19/TRBJ2.1 in Child-2. In this example, however, the CDR3 sequence was variable even though the V-J gene segments appeared likely pairings.

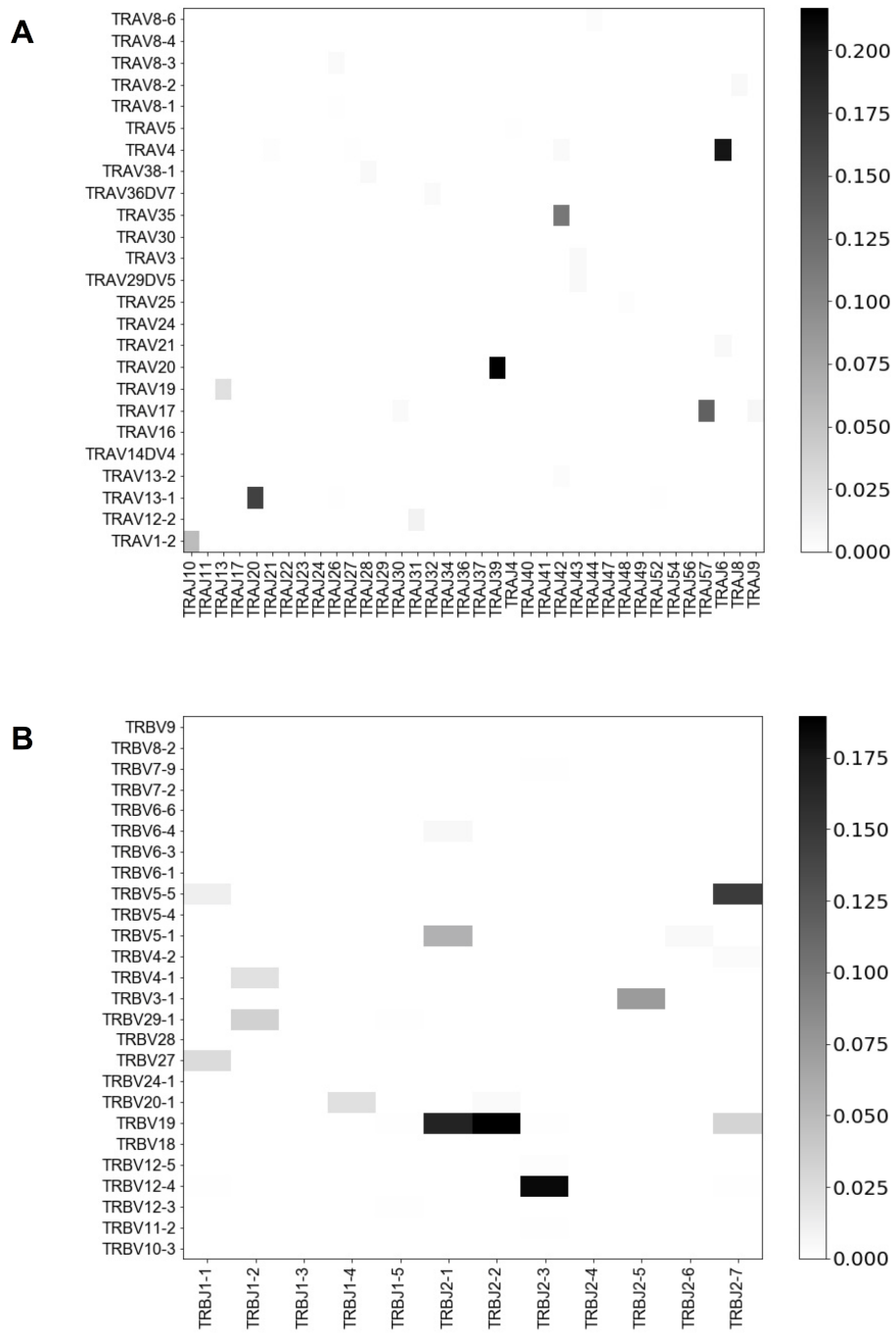


Figure 5.2: V and J gene usage in TCR- α and - β chain repertoires of HLA-B*44:03-NW8-specific CD8+ T-cells. (A) TRAV-TRAJ co-occurrence frequencies. (B) TRBV-TRBJ co-occurrence frequencies. V-J pair frequencies are represented as heats maps.

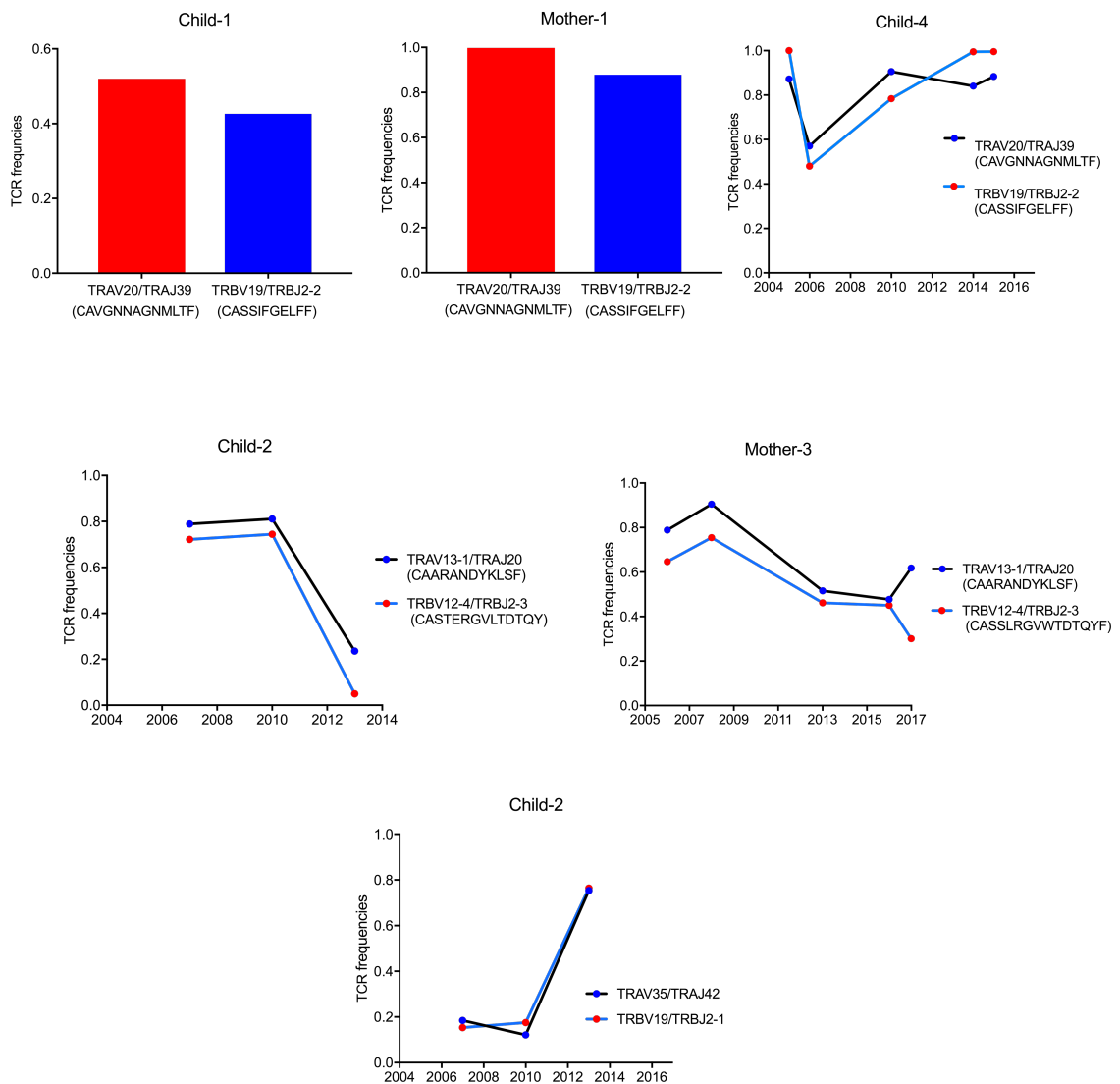


Figure 5.3: TCR- α and - β chain pairings. The individual cases of predicted TCR- α and - β chain pairs ($n=3$) are shown in six individuals. The TCR frequencies are shown on the y-axis. Frequencies were calculated as a percentage of aligned sequencing reads.

In the example of the most commonly observed combination observed here, TRAV20/TRAJ39 (CAVGNNAGNMLTF) with TRBV19/TRBJ2-2 (CASSIFGELFF), this was seen also in the unrelated adults studied (Patient 5, Patient 14, Patient 18, and Patient 20 in Chapter 4). However, in these subjects, the frequency of these particular pairs does not correspond especially well, suggesting that these α and β chains do not necessarily need to be paired.

5.3.3 TCR publicity

As expected and previously shown in Chapter 4, TCR sequence analysis showed that the most expanded clones were often public TCRs. I observed that the higher the frequency of a given clonotype, the greater the likelihood that this would be a public TCR sequence. This relationship was statistically significant for TCR- α and - β chains (Fig 5.4A). The average frequency of public TCR- α and - β chain clonotypes was significantly higher than the private clonotypes (Fig 5.4B).

Multiple public TCRs were identified: 25 out of 76 TCR- α chain clonotypes and 35 out of 132 TCR- β chains were public. The clonotype CAVGNNAGNMLTF (TRAV20/TRAJ39), was the most abundant clonotype and was also the most extensively shared. CAVGNNAGNMLTF was found in 8/9 (22% of all assigned clonotypes) of subjects studied. Similarly, the most abundant public TCR- β clonotypes, CASSIFGELFF and CASSIFGEQFF (TRBV19/TRBJ2-2 and TRBV19/TRBJ2-1, respectively), were both shared by 7/9 of individuals studied. The average number of public TCR- α chains were more prevalent in mothers compared to children, although this didn't reach statistical significance (Fig 5.4C).

The high degree of sharing of common clonotypes arise from multiple public clonotypes produced by convergent recombination (Venturi et al. 2006, Venturi et al.

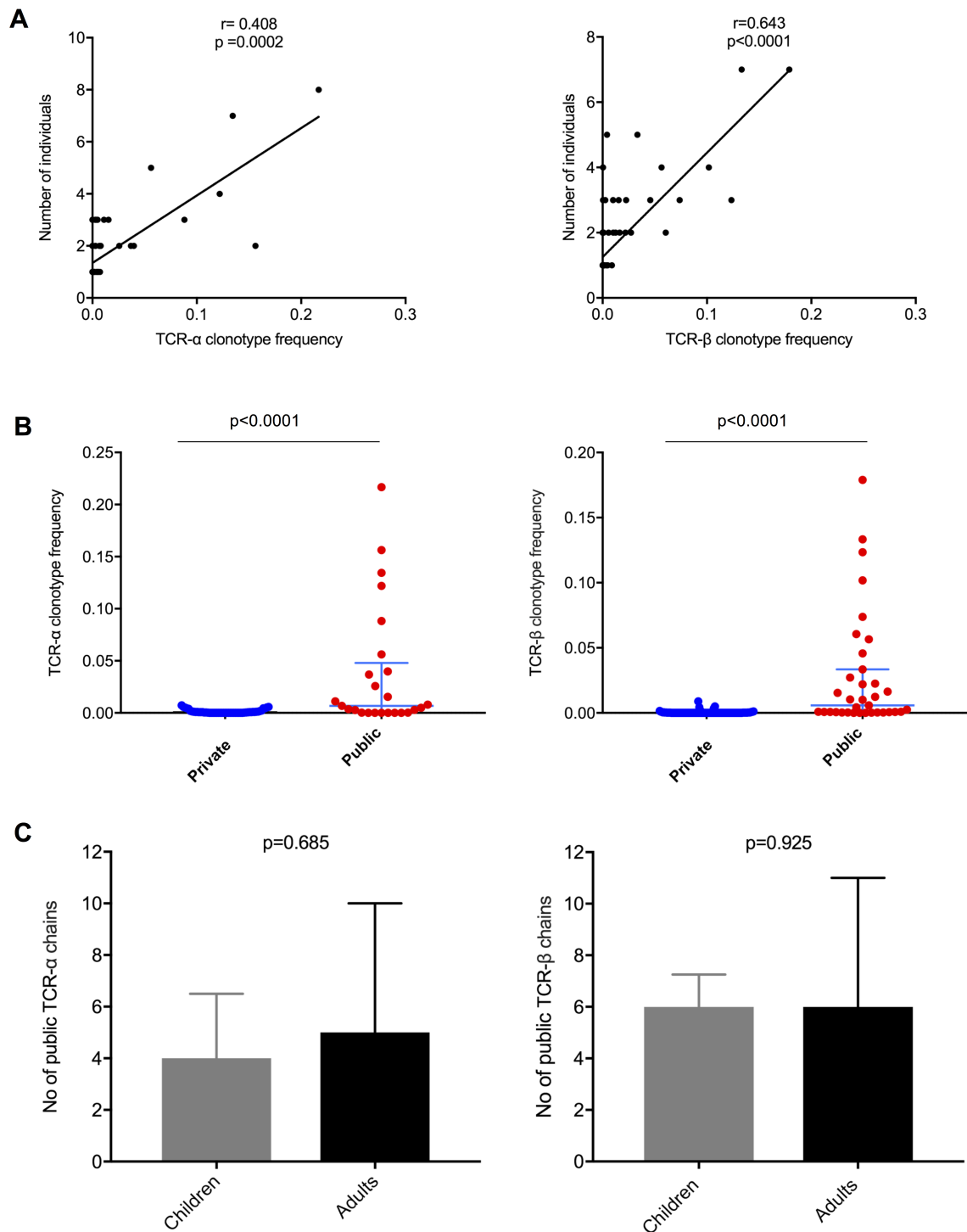


Figure 5.4: Clonal dominance and TCR publicity. (A) Average frequency of a given clonotype plotted against the number of individuals sharing that clonotype for TCR- α chains ($n=76$) and TCR- β chains ($n=132$). (B) Frequencies of private and public TCR- α and TCR- β chain clonotypes. The horizontal bar represents IQR. (C) Average number of TCR- α and TCR- β public clonotypes in children and adults. p-values in (B) and (C) are calculated using Mann-Whitney U test.

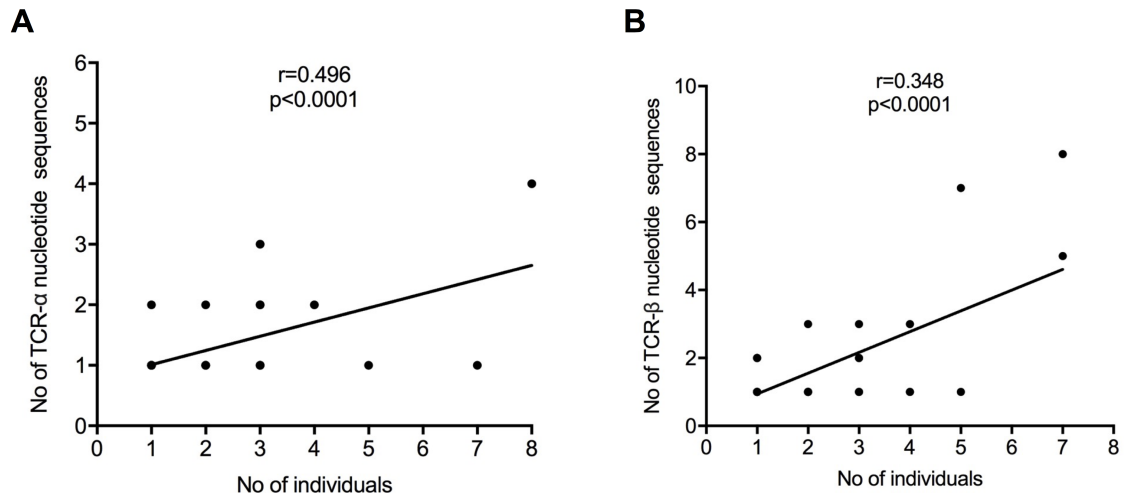


Figure 5.5: Number of nucleotide sequences and TCR publicity. Number of TCR- α and - β nucleotide sequences encoding a given clonotype plotted against the number of individuals sharing that clonotype. Spearman's rho and p-values are shown above each graph.

2008a). Indeed, we found clonotypes that were more frequently shared in individuals were also encoded by a greater number of distinct nucleotide sequences (Fig 5.5).

5.3.4 Changes in NW8-specific TCR repertoire diversity

The TCR repertoire diversity is known to change with exposure to viral antigens (Monteiro et al. 1995; Messaoudi et al. 2004). To address how the TCR usage changes over time following CMV infection and to examine whether the TCR usage is more stable over time in adulthood than in childhood, the diversity of TCR- α and - β chain clonotypes was calculated at every time point using the Shannon evenness index (J') (as explained previously in chapter 4). Briefly, J' is 0 when the sample is monoclonal and is 1 when all clonotypes have the same frequency. Most repertoires were preferentially expanded with a limited set of clonotypes, as suggested by low J' values. The TCR usage varied from individual to individual, with public clones often occupying a greater homeostatic space in the NW8-specific

repertoire. When observing the changes in TCR repertoire diversity in a single child and mother over the span of 10 years, significant fluctuations were seen in both (Fig 5.6A).

Overall, the TCR- α chain repertoire diversity undergoes more fluctuations in children, than mothers. However, the variation in TCR- β chain repertoire diversity is the same in children and mothers (Fig 5.6B). A general trend of increase in diversity of TCR- α and TCR- β repertoire with age was observed. For TCR- β chains, this association was statistically significant ($r=0.39$, $p=0.02$, Fig 5.6C).

5.3.5 The relationship between CD8+ T-cell response magnitude and TCR repertoire diversity in children and mothers

The TCR diversity has been related to the magnitude of the CD8+ T-cell response (Costa et al. 2015). Therefore, I studied how changes in TCR diversity (measured as the Shannon evenness index) related to longitudinal changes in CD8+ T-cell response magnitude in children and mothers (Fig 5.7A-B). In mothers, a significant negative correlation was observed between TCR- α repertoire diversity and CD8+ response magnitude ($p=0.002$, $r=0.710$, Spearman's rank correlation; Fig 5.7B). This indicates that the TCR- α chain repertoires of mothers became less diverse with CD8+ T-cell expansion and more diverse with contraction. On this basis, the longitudinal CD8+ T-cell responses per person and the corresponding TCR- α / β repertoire compositions were compared. The shifts in CD8+ T-cell response magnitude over time were associated with inflation or deflation of particular clonotypes in some, but not all mothers (Fig 5.8). For example, the decreasing response in Mother-3 (between 2008-2013) was accompanied by deflation of a dominant TCR-

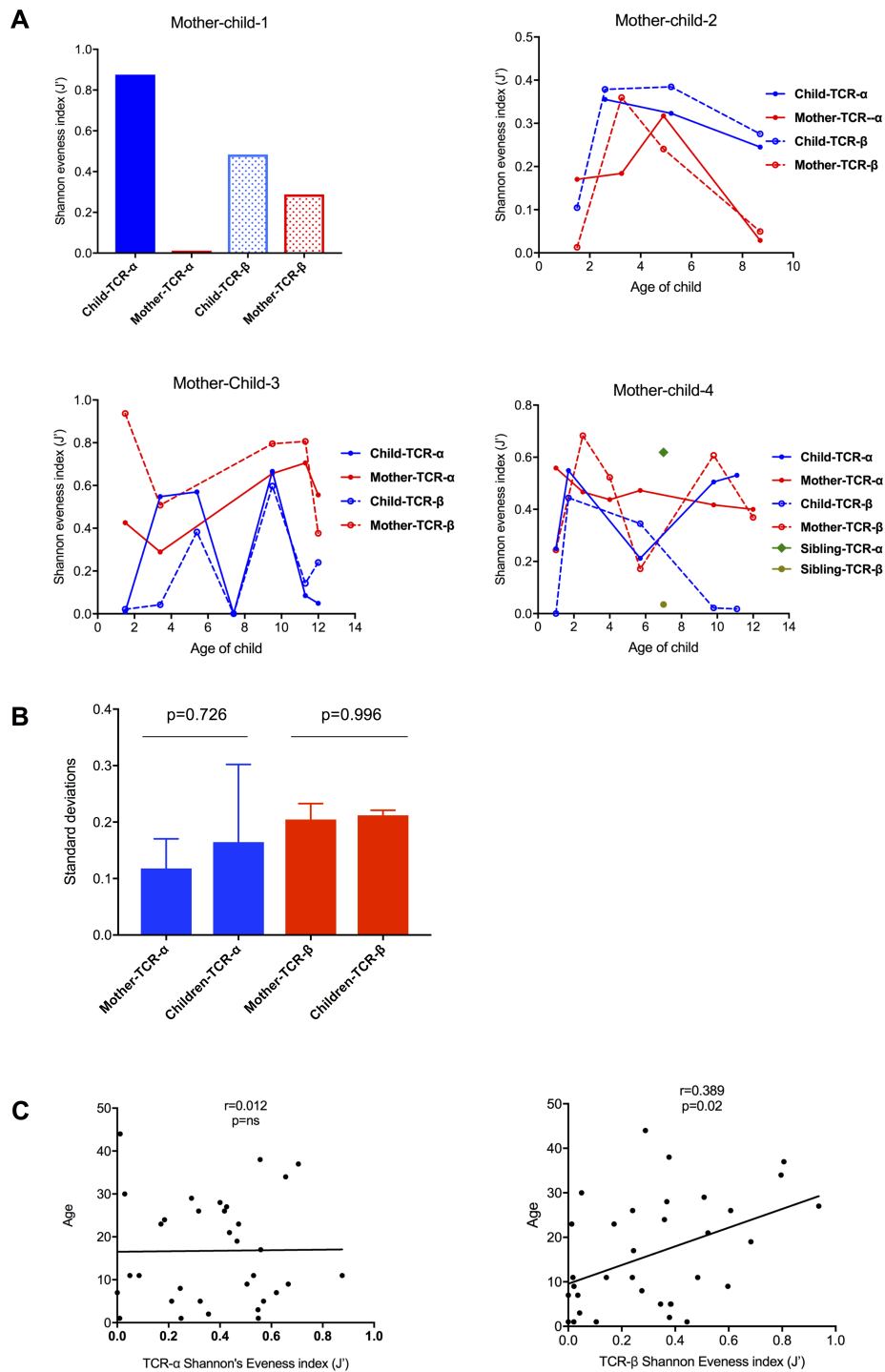


Figure 5.6: Changes in diversity of TCR- α and - β chain clonotypes in children and their mothers over the course of 10+ years. (A) The diversity of TCR- α and - β chain repertoires of children ($n=5$) and their mothers ($n=4$). **(B)** Variation in TCR repertoire diversities expressed by standard deviations of Shannon diversity indices (J') in children ($n=3$) and mothers ($n=3$). **(C)** The diversities of TCR- α and - β chain repertoires plotted against the age. Diversities are calculated using Shannon evenness index. J' is 0 when the sample is monoclonal and is 1 when all clonotypes have the same frequency.

α (CARSNDYKLSF) and $-\beta$ clonotype (CASSLRGVWTDQYF) (Fig 5.8C) and the increasing response (between 2006-2008) was accompanied by inflation of these dominant clonotypes. This relationship was not found at all in the children studied.

5.3.6 Overlap of TCR- β repertoires for related and unrelated mother-child pairs

In the past, a number of studies have revealed that unrelated individuals widely share TCR repertoires (Venturi et al. 2008a, 2008b, 2011; Li et al. 2012) However, it is not clear whether the repertoires of haploidentical individuals (i.e. mother-child) are characterised by a higher level of overlap compared to unrelated individuals. To address this, the degree of overlap in TCR- α and $-\beta$ chain repertoires of related and unrelated individuals was compared using Morisita-Horn similarity-index (CMH) (explained in chapter 4). There was a greater degree of overlap in the TCR- α and $-\beta$ repertoires of related individuals compared to unrelated individuals, however, it did not reach statistical significance (Fig 5.9; Mann-Whitney U test). The greatest relative overlap of TCR repertoires was observed among high-frequency clonotypes, defined as those present at $>10\%$. This finding can be attributed to favourable selection of high-frequency clonotypes during T-cell development, which leads to higher representation of these common clonotypes both within and between individuals (Venturi et al. 2011).

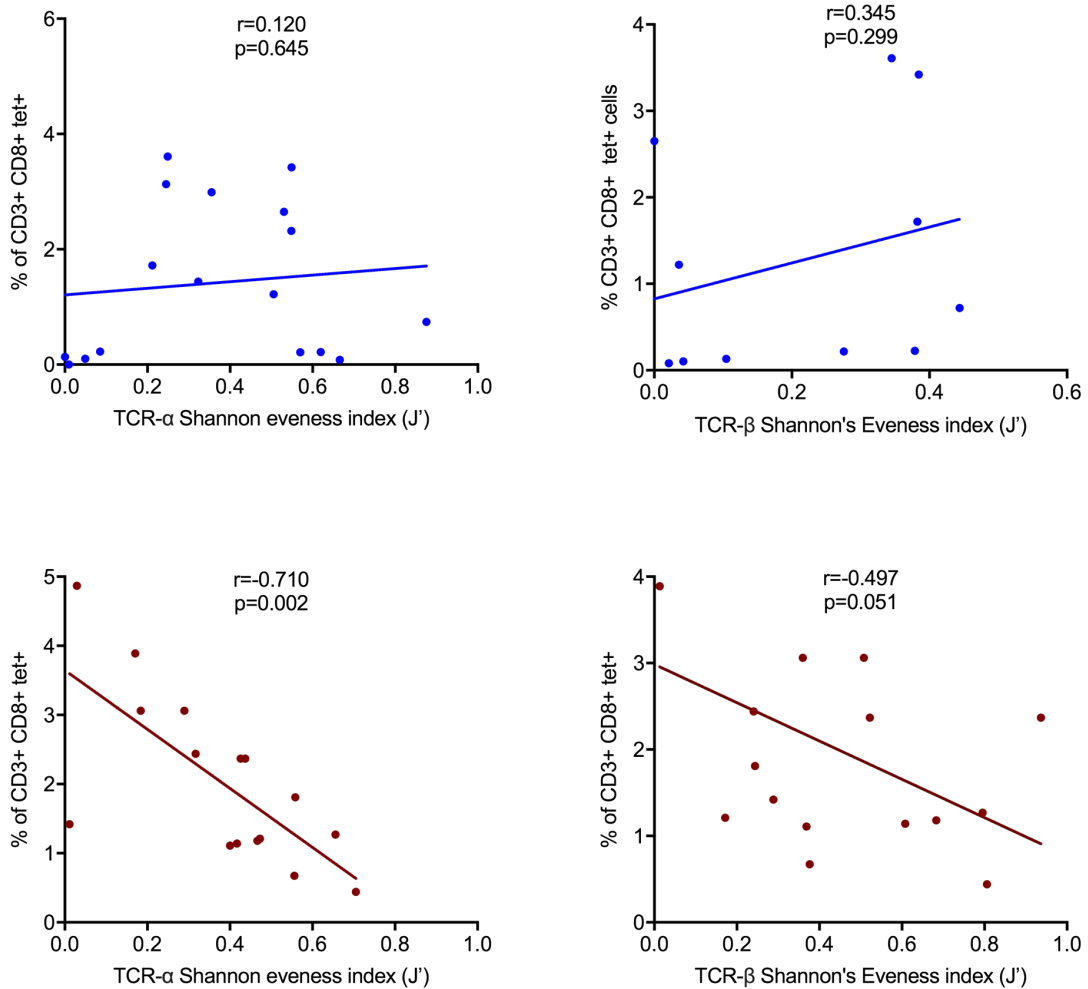


Figure 5.7: The relationship between CD8+ T-cell response magnitude and TCR repertoire diversity in children and mothers. (A) The diversity of TCR- α and - β chain repertoires is plotted against the magnitude of CD8+ T-cell response in children. **(B)** The diversity of TCR- α and - β chain repertoires is plotted against the magnitude of CD8+ T-cell response in adults. Diversities are calculated using Shannon evenness index (J'); J' is 0 when the sample is monoclonal and is 1 when all clonotypes have the same frequency. Spearman's rho and p-values are shown above each graph.

Figure 5.8A: Mother-Child Pair 1

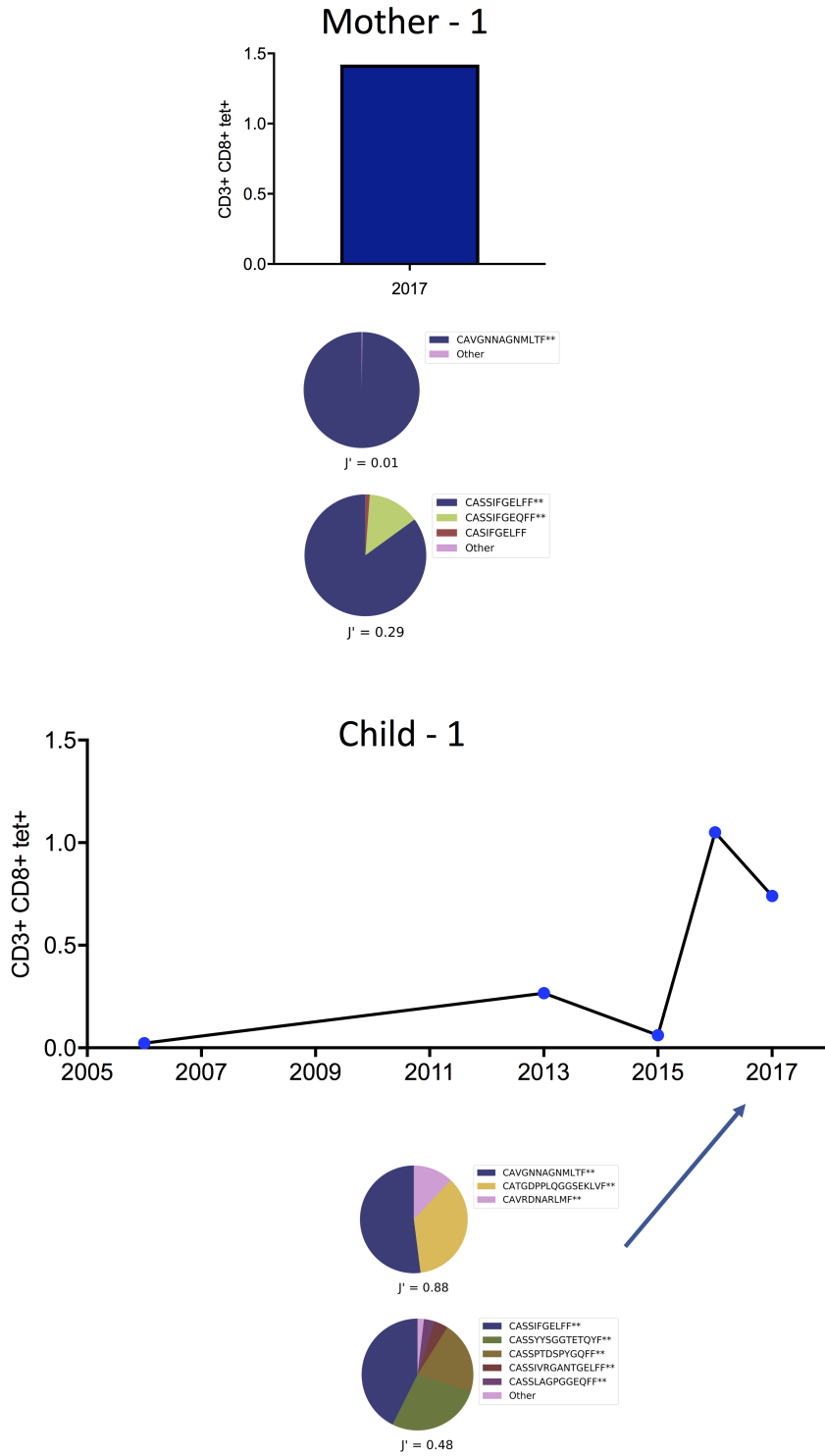


Figure 5.8B: Mother-Child Pair 2

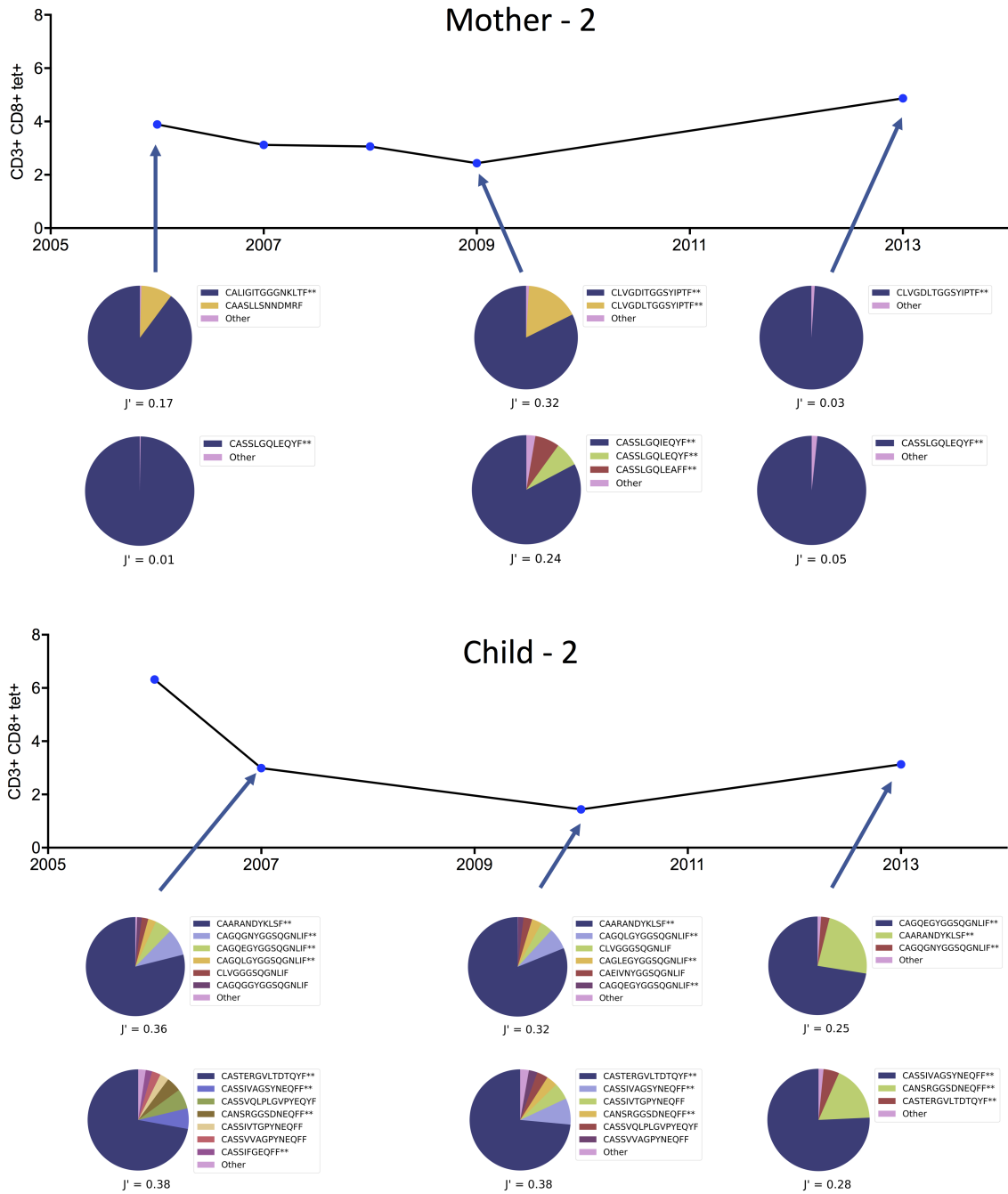


Figure 5.8C: Mother-Child Pair 3

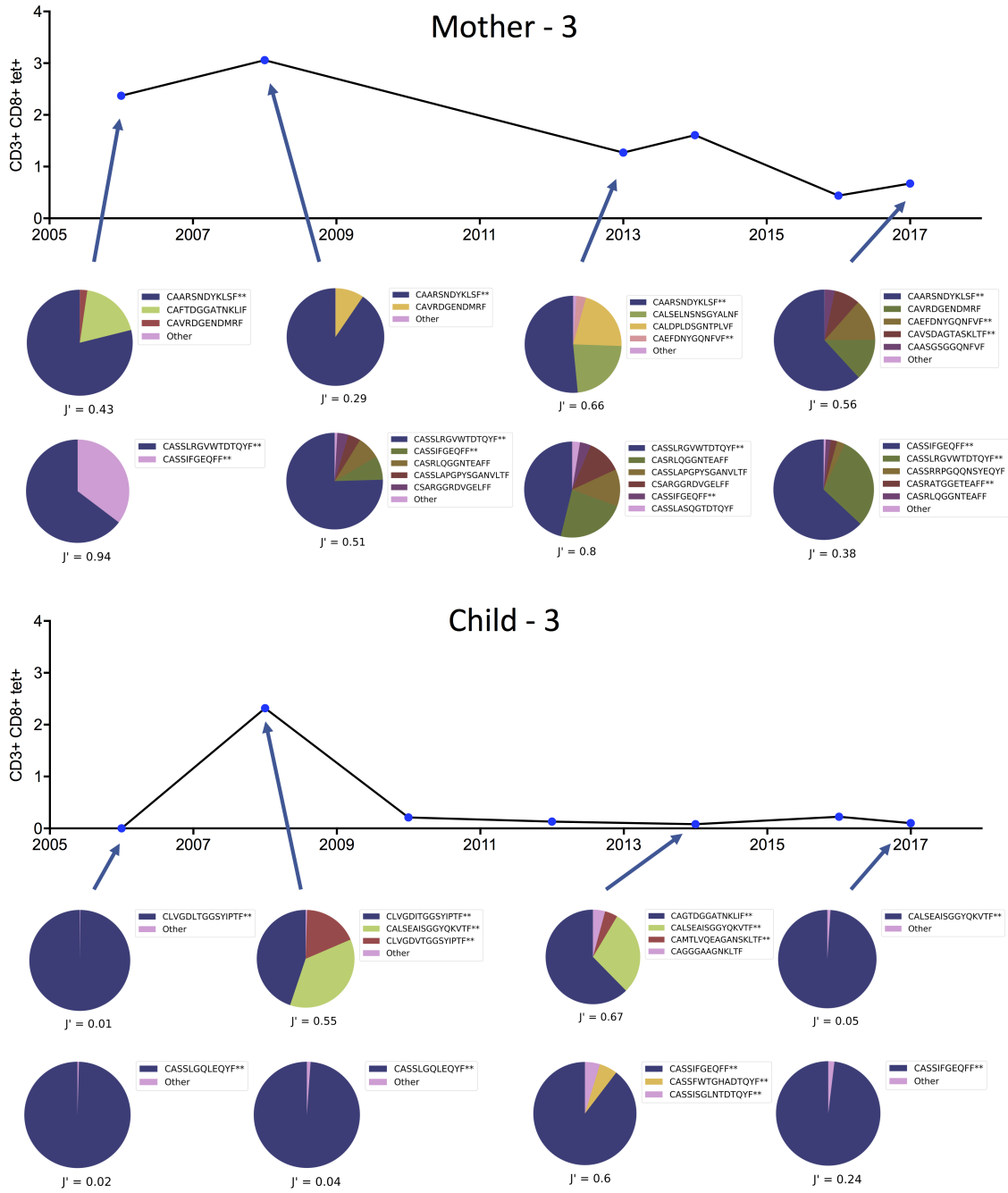


Figure 5.8D: Mother-Child Pair 4

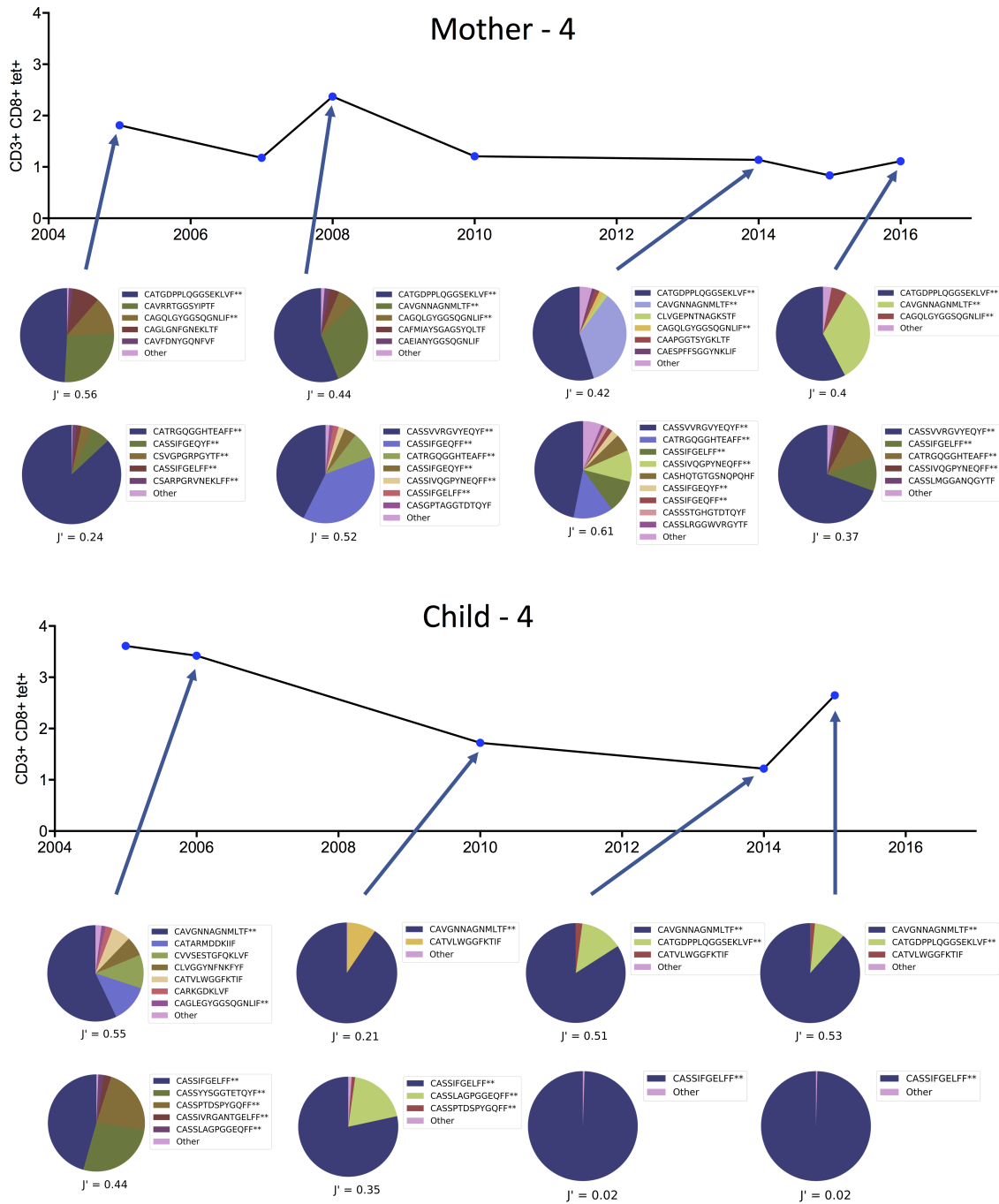


Figure 5.8: CD8+ T-cell response magnitude and composition of TCR- α and - β chain clonotypes. The magnitude of CD8+ T-cell response is shown in the top panel, composition of TCR- α chain clonotypes, in the middle panel, and TCR- β chain clonotypes, in the lower panel. Public clonotypes are shown with (**). Shannon evenness index (J') is shown under each pie chart. Colours are assigned randomly and do not correspond to a fixed a sequence.

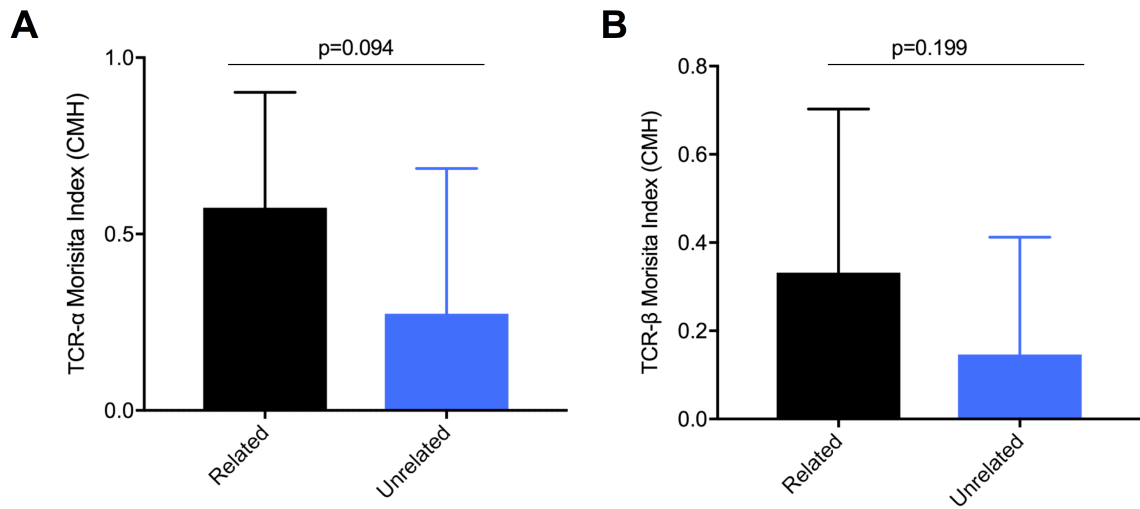


Figure 5.9: TCR sharing amongst NW8-specific CD8+ T-cells from related versus unrelated mother-child pairs. (A) Compositional similarity amongst TCR- α clonotypes in related and unrelated mother-child pairs. **(B)** Compositional similarity amongst TCR- β clonotypes in related and unrelated mother-child pairs. The compositional similarity is assessed using the Morisita-Horn (CMH) index. CMH values range from 0 (no overlap) to 1 (perfect overlap). p-values are calculated using Mann-Whitney U test are shown above each graph.

5.4 Discussion

The evolution over a span of 10 years of the clonal composition of CD8+ T-cells responding to an immunodominant CMV epitope was assessed in children and their mothers. All the individuals made a detectable response to the CMV NW8-epitope at all time points. Contrary to previous findings (O'Hara et al. 2012), we found no increase in the magnitude of CMV-specific CD8+ T-cell response with age. In the children studied, the highest responses were observed during the first the 1-2 years of life.

Is TCR usage stable over time following CMV infection?

Overall, there was a great degree of fluctuation in the TCR usage over time following CMV infection, both in children and mothers. Some individual cases did exhibit

stability in TCR usage, for example, the TCR- α chain usage in Mother-4 and in Child-2. However, this stability was not reflected in the TCR- β chains of these individuals. Public TCRs often occupied the greatest space in the TCR repertoire. However, the public TCR usage varied from individual to individual.

Is TCR usage more stable over time in adulthood, decades after infection occurred, than in childhood?

The TCR usage in children varied greatly over time. Whilst this is expected in children that were recently exposed to CMV, this variation was also observed in mothers who have been exposed to CMV for multiple decades. The individual TCR- α chain repertoires of children (aged 1-11 years) undergo more variations compared to mothers, however, the overall variation in TCR- β chain repertoire is similar. Despite the presence of strong inter-individual variation in the overall features of the TCR repertoires in all of the individuals, there was a trend of increase in repertoire diversity of TCR- β chains with age. Studies in the past have reported contrary findings, whereby, the TCR repertoire diversity is either well maintained till the age of 65 followed by dramatic contraction (Naylor et al. 2005) or there is a linear decrease in repertoire diversity with age (Britanova et al. 2014; Britanova et al. 2016).

Does the TCR usage fluctuate randomly over time or is there a clear tendency for particular public TCRs to expand over time and dominate the repertoire?

A study by Day et al. (2007) had shown that public TCRs rapidly expand following acute CMV infection and dominate the CMV-specific TCR-repertoire. Consistent with these findings, the TCR- α and - β chain repertoires were found to be highly skewed in both children and mothers in the study described herein, as indicated by preferential use of V and J genes and extensive TCR sharing. Public TCRs were

often the highest frequency clonotypes within individuals, however, there were variable patterns of public TCR usage for each individual.

Is there any difference in the TCR usage (the CDR3 sequences) between children (as a group) and mothers (as a group)?

The comparison between children and mothers demonstrates that individual adult and child NW8-specific TCR repertoires have the same basic features. Namely, the overall NW8-specific TCR- α and - β chain repertoires of children and adults are largely determined by public TCR clonotypes that are often shared between children and mothers. For example, the most frequent TCR- α chain clonotype CAVGNNAGNMLTF was shared by all of the five children and three out of the four mothers. Similarly, one of the most frequent TCR- β chain clonotype CASSIFGEQFF was shared by three out of six children and all of the four mothers.

Is there any similarity in the TCR usage within particular mother-child pairs compared to between unrelated individuals?

A higher number of shared clonotypes were observed in related versus unrelated mother-child pairs, with the greatest overlap amongst high-frequency clonotypes. These findings are consistent with previously published studies (Robins et al. 2010; Putintseva et al. 2013). The higher overlap among the high-frequency clonotypes can be attributed to their higher production frequency, which is an important determinant of TCR sharing (Venturi et al. 2008a).

The involvement of shared clonotypes in NW8 repertoires can be explained by the number of ways shared clonotypes are produced more favourably and efficiently, such as recombinant recombination (Venturi et al. 2008b, Venturi et al. 2011) or if encoded by a nucleotide sequence with few additions and deletions

(Venturi et al. 2006; Li et al. 2012). Indeed, there was an increase in the level of convergent recombination with the increase in sharing of a given clonotype. Furthermore, the public chains were nearly entirely germline-encoded or involved very limited changes at the CDR3 junction. Thus, it is possible that the clonotypes that are more frequently produced with higher efficiency become overrepresented as thymic output decreases in older individuals (Rudd et al. 2011). Public TCRs are of higher affinity (Kloverpris et al. 2015) and because TCR signalling dampens with age, it is also possible that high affinity public TCRs are required to obtain sufficient levels of TCR signalling strength for both homeostatic maintenance and activation (Murasko et al. 1987; Rudd et al. 2011). Given that this is the case, it may provide an explanation for the emergence of more public clonotypes in older individuals.

Certain TCR- α and - β chain pairs were observed where changes in α/β frequencies mirrored each other. These were chosen conservatively by observing the frequencies over time in individuals. There may be more pairs, e.g. TRAV4/TRAJ6 with TRBV5-5/TRBJ2-7 in Mother-2 but more time points are needed to strengthen the hypothesis and single cell TCR sequencing could confirm these hypotheses. Accurate assignment of paired alpha and beta genes would be very useful to track TCR “signatures” associated with specific human diseases. This can be a particularly useful approach in diseases where pathogenesis is usually slow, particularly autoimmunity, where it takes several years from symptoms to make a clinical diagnosis. With the recent advances in high-throughput sequencing, the study of TCR repertoires in response to infections has been revolutionised (Miconnet 2012). These new advances will facilitate practical TCR-based diagnoses.

Overall, the NW8-specific TCR repertoires exhibited a greater degree of change among adults than expected, and similar features among children and adults. Con-

trary to previous findings, the CMV-specific TCR repertoire of young children was highly skewed with limited repertoire diversity.

Chapter 6

PARV4 and Hepatitis-B prevalence, CD8+ T-cell responses and co-infection with HIV and CMV in African children and adults

6.1 Introduction

The main focus of the work described in my thesis has been to further the understanding of CD8+ T-cell responses to a chronic virus, CMV, in African populations. Although my research has therefore primarily been concentrated on CMV, I have also contributed to a wider effort to characterise the immune responses to other chronic viral infection, namely, Human parvovirus 4 (PARV4) and Hepatitis B (HBV), which, like CMV, have been, and to some extent still are, common childhood infections in sub-Saharan African populations. This is of particular relevance because of the high prevalence of HIV infection in these same populations. The immunodeficiency caused by chronic HIV infection influences the consequences

of co-infection, and vice versa, increasing the risk of long-term morbidity and mortality. Despite these concerns, the burden and impact of HBV and PARV4 in sub-Saharan Africa have not been well characterised in adults and even less so in children.

6.1.1 Parvovirus 4 (PARV4)

PARV4 is a single-stranded DNA virus in the Parvoviridae family, first reported in 2005 in a HBV-positive injecting drug user (IDU) (Jones et al. 2005). There are number of studies that suggest that PARV4 operates as a chronic latent infection (Simmons et al. 2011, 2013; Yang et al. 2011; Chen et al. 2015). Studies have demonstrated a sustained high-magnitude ex vivo CD8+ T-cell response to PARV4 non-structural (NS) peptides, which suggests a persistent or relapsing exposure to viral antigen (Simmons et al. 2011; Simmons et al. 2013). Furthermore, in a study of Taiwanese healthcare workers, PARV4 IgG antibody was detected in 60% of individuals, and PARV4 DNA was detected in half of these, suggesting persistent PARV4 replication or reactivation of a latent PARV4 infection (Chen et al. 2015). The clinical impact or risk factors for transmission for PARV4 have not been clearly determined. However, individuals with evidence of past or current PARV4 infection exhibit a range of clinical symptoms (Matthews et al. 2014b) (Table 6.1).

A number of previous studies have reported an increased prevalence of PARV4 IgG in people infected with HIV (Manning et al. 2007; Simmons et al. 2012) and HCV (Fryer et al. 2007; Simmonds et al. 2007; Lurcharchaiwong et al. 2008; Schneider et al. 2008; Simmons et al. 2011; Servant-Delmas et al. 2014) in North American and European populations. In individuals infected with PARV4, a faster progression to HIV CDC-B symptoms has been observed (Simmons et al. 2012), although this effect is not consistent across cohorts and may be confounded by HCV co-

infection. The strong association of PARV4 with HIV and the potential clinical impact of PARV4 on HIV reported in European studies are of particular concern in a setting where these viruses are co-endemic. The general rate of PARV4 infection in sub-Saharan African populations is high but varies depending on location [Table 6.2] (Simmonds et al. 2008; Panning et al. 2010; Sharp et al. 2010b; Drexler et al. 2012; Lavoie et al. 2012; May et al. 2012; Matthews et al. 2015b). Given that a strong association between PARV4 and HIV has been reported in previous European studies, I have contributed to the further investigation of the prevalence of PARV4/HIV co-infection and investigated the impact of PARV4 infection on HIV disease progression and HIV-specific CD8⁺ T-cell responses in sub-Saharan African populations.

PARV4 can be sub classified into three genotypes: genotypes 1 and 2 are predominant in Europe, North America, and Asia and genotype 3 is the most widespread in Africa (Simmonds et al. 2008). PARV4 contains two open reading frames (ORFs), ORF1 and ORF2, that encode NS and viral capsid (VP) proteins, respectively (Fig 6.1). NS is essential for viral replication and potential cytopathic effects (Lou et al. 2012) and thought to be the preferential target of CD8⁺ T-cells (Simmons et al. 2011, 2013). VP1 and VP2 are structural capsid proteins and therefore expected to be the main target of an antibody response (Lou et al. 2012). Previous studies have defined CD8⁺ T-cell responses to PARV4 NS protein in PARV4 IgG positive adults (Simmons et al. 2011, 2013). However, in some of the PARV4 IgG positive individuals, no PARV4 NS-specific T-cell responses were observed (Simmons et al. 2011, 2013). It was speculated that this may be in part due to VP-specific T-cells rather than NS-specific T-cells, however, no VP-specific responses were studied (Simmons et al. 2011). To further our understanding of PARV4-specific responses, I built upon previous work in three aspects: 1) the study of the entire PARV4 pro-

teome instead of just the NS protein, 2) the inclusion of CD8+ T-cell responses to PARV4 proteome in children as well as adults, and 3) a focus on African subjects.

6.1.2 Hepatitis B virus (HBV)

Hepatitis B virus (HBV) is a hepatotropic virus of the Hepadnaviridae family that causes variable degrees of liver disease in humans (Busca and Kumar 2014). HBV is a partly double-stranded DNA virus with several serological markers (Trepo et al. 2014) (Table 6.3). The presence of HBV surface antigen (HbsAg) is the characteristic hallmark of HBV infection. Infection with HBV can be either acute or chronic. The outcome of HBV infection depends on the age and immune status of the individual; infection in immunocompetent adults has a relatively low rate of chronicity, while neonatal infections usually have a high persistence rate (Trepo et al. 2014). Chronic HBV is characterised by fluctuations in liver enzymes, mainly alanine aminotransferase (ALT) (Trepo et al. 2014). Recurrent ALT flares reflect episodes of inflammation that increase the risk of cirrhosis and hepatocellular carcinoma (Chang and Liaw 2014).

Throughout the world there is a high prevalence of HBV, with an estimated 257 million people living with HBV infection (WHO) (<http://www.who.int/mediacentre/factsheets/fs204/en/>). In the sub-Saharan African population, chronic HBV is considered to be highly endemic with $\geq 8\%$ population prevalence in adults (Matthews et al. 2014a). Due to the high prevalence of HIV in the same population, HIV/HBV co-infection is a major concern (Ocama et al. 2011; Stabinski et al. 2011; Hawkins et al. 2013; Ladep et al. 2013). The existing literature on HBV infection is generally based on small studies of disparate populations (Matthews et al. 2014a). Therefore, I contributed to a study of a large cohort of

Virus	Prevalence in HIV+ population	Prevalence in HIV- population	Transmission routes	Clinical manifestations	CD8+ T-cell responses	Adult vs paediatric infection
CMV	100%	100%	Saliva, sexual contact, placental transfer, breastfeeding, blood transfusion, SOT or HSC	Described in table 1.2	Strong CD8+ T-cell responses; UL48, PP65, IE1 and IE2 being the most immunogenic proteins	Prolonged viral excretion in young children compared to adults (upto 5 years in children vs 3–6 months in adults)
PARV4	51.80%	24.40%	Potentially via direct contact with blood or respiratory secretions. Route of transmission unclear	Flu-like symptoms, encephalitis, transient rash, foetal hydrops, or gastroenteritis	Strong CD8+ T-cell responses to entire PARV4 proteome; high magnitude responses to NS and ARF proteins	Not known
HBV	6.90%	8.30%	Sexual contact, placental transfer, sharing needles, blood transfusion or any direct contact with blood	Flue-like symptoms, jaundice, inflammation of the liver, cirrhosis, hepatocellular carcinoma	Weak CD8+ T-cell responses during chronic infection	25-50% children exposed before the age of 5 develop chronic HBV vs 5-10% of adults

Table 6.1: Prevalence and characteristics of CMV, PARV4, and HBV infection. The prevalence of PARV4, HBV and CMV was calculated based on the adults and children recruited via our African cohorts based in South Africa and Botswana. Source: Busca and Kumar 2014; Matthews et al. 2014b; Simmons et al. 2011; Trepo et al. 2014)

Study Location	Sample type	Subjects	Characteristics of the cohorts	Assay	PARV4 positive %	References
sub-Saharan Africa	Bone marrow/lymph node	Adults	HIV+	PCR	15.3%	Simmonds et al. 2008
Ghana	Blood	Children	On preventative Malaria trial	PCR	8.6%	Panning et al. 2010
Burkina Faso	Serum	Adults	Blood donors	ELISA	37.1%	Sharp et al. 2010
Cameroon	Serum	Adults	General populations	ELISA	24.8%	Sharp et al. 2010
Democratic Republic of the Congo	Serum	Adults	Military populations	ELISA	35.3%	Sharp et al. 2010
South Africa	Serum	Adults	Blood donors (HIV+)	ELISA	36.4%	Sharp et al. 2010
South Africa	Serum	Adults	Blood donors	ELISA	4.4%	Sharp et al. 2010
Ghana	Nasal swab	Children	With respiratory infections	PCR	0.8%	Drexler et al. 2012
Ghana	Fecal	Children	With respiratory infections	PCR	0.4%	Drexler et al. 2012
Ghana	Plasma	Children	On preventative Malaria trial	PCR	8.9%	May et al. 2012
Cameroon	Serum	Adults	High rate of HCV	ELISA	17.5%	Lavoie et al. 2012

Table 6.2: Prevalence of PARV4 reported in previously published studies of African populations.

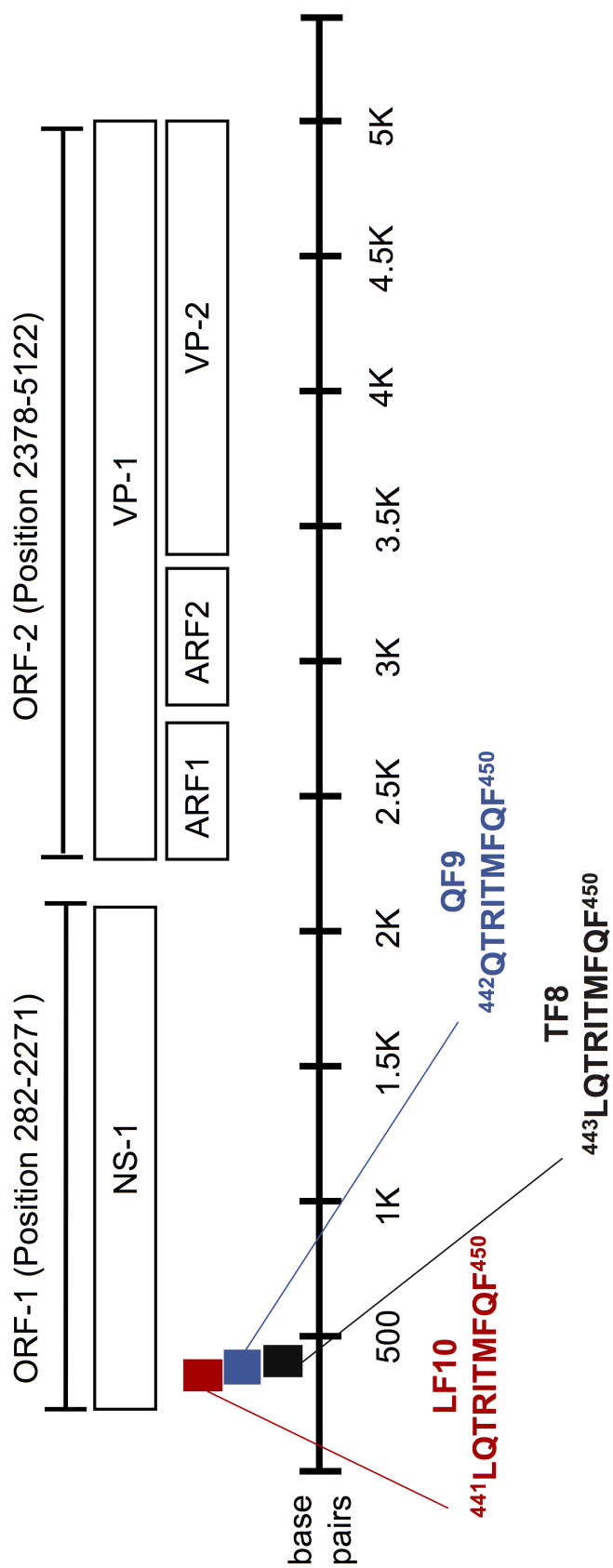


Figure 6.1: Schematic diagram of PARV4 genome and schematic views of HLA-B*57:03-restricted epitope variants in PARV4 genome. Open Reading Frame 1 (ORF-1) encodes Non-Structural Protein 1 (NS-1). Open Reading Frame 2 (ORF2) comprises overlapping proteins Viral Protein 1 (VP-1) and Viral Protein 2 (VP-2). Protein lengths are shown as number of amino acids (aa). Additional Reading Frames (ARF) are conserved across PARV4 genotypes; shown as ARF-1 and ARF-2. The coloured boxes indicate three different variants of a putative novel HLA-B*57:03 epitope.

	Acute HBV	Chronic HBV	Cleared HBV	Vaccination
HBsAg	+	+	-	-
HBcAb IgM	+	-	-	-
HBcAb IgG	+	+	+	-
Anti-HBs	-	-	+	+
HBeAg	+	+/-	-	-
Anti-Hbe	-	+/-	+/-	-
HBV DNA	High/Low	Low/High	-	-

Table 6.3: Serological markers associated with HBV infection. Source: (<http://www.oxfordmedicaleducation.com>)

HIV-positive African women and children and a small control group of HIV-negative individuals to characterise the prevalence, rate of co-infection with HIV/PARV4 and the impact of HBV infection on HIV disease outcome.

The Sustainable Development Goals initiative and the World Health Organisation (WHO) have set the goal to eliminate HBV by 2030. Based on the current prevalence rate, this will require an increase in research to find a cure and improve robust administration of prophylactic vaccinations. Currently, the only therapeutic option for chronic HBV is prolonged (often lifelong) anti-viral treatment to suppress virus replication [<https://www.nice.org.uk/guidance/cg165>]. Therefore, it is important to work towards a therapeutic vaccine, underpinning the urgent need for more research on HBV-specific immune responses. The immune determinants for successful clearance of HBV are not fully understood, but both innate and adaptive immune responses may be important. HBV replication itself is not directly cytotoxic to cells, but host immune responses to HBV antigens and to viral replication in infected hepatocytes are likely to be the main determinants of hepatocellular injury and HBV pathogenesis (Chisari et al. 2010).

Strong CD8⁺ T-cell responses to several HBV epitopes in different regions of the HBV genome are associated with recovery from acute HBV infection (Ferrari et al. 1990; Jung et al. 1991; Maini et al. 1999; Thimme et al. 2003; Trepo et al. 2014). In chronic HBV, HBV-specific T-cell responses are significantly reduced (Boni et al. 2007; Yang et al. 2010), making it difficult to study CD8⁺ T-cell responses. It is evident that clearance of HBV depends on robust and sustained CD8⁺ T-cell activity (Thimme et al. 2003), therefore, therapeutic vaccines that can stimulate or mimic the multi-epitope specific responses may help resolve chronic HBV infection (Barnes 2015). Advancing our understanding of HBV-specific T-cell responses that are associated with containment and/or clearance are important for the development of a therapeutic vaccine.

The HIV Molecular Immunology Database is a public database on HIV epitopes (https://www.hiv.lanl.gov/content/immunology/ctl_search). This database enables increased collaborations and knowledge pooling in the field of HIV T-cell immunology. There are no such public databases documenting HBV epitopes, which hinders the study of HBV-specific T-cell immunology. We recognised the need for an equivalent resource for HBV. Therefore, a project involving the process of collating HBV class I epitopes was initiated, with the aim to provide a detailed and comprehensive database of epitopes that have been identified to date.

6.2 Materials and methods

6.2.1 Patient cohorts

- i. **Durban HIV-positive cohort:** HIV-positive women at Cato Manor and Sinikithemba attending antenatal clinics in Durban, South Africa (Brumme et al. 2010; Kiepiela

et al. 2007).

- ii. **Kimberley HIV-positive cohort:** HIV-positive women recruited as mothers of children attending paediatric HIV clinics in Kimberley, South Africa (Matthews et al. 2011); 2012.
- iii. **Gaborone HIV-positive cohort:** HIV-positive women recruited via the Mma Bana Study (Shapiro et al. 2010), attending antenatal clinics in Gaborone, Botswana.
- iv. **HIV-negative adults and children:** HIV-negative women were recruited from antenatal clinics in Durban (Masibambisane cohort) and HIV-negative children were recruited via Kimberley Respiratory Cohort (KReC). KReC comprises children aged between 9 and 48 months who were seen in the paediatric department at Kimberley Hospital following respiratory tract infection.

6.2.2 PARV4 IgG assay

Indirect ELISA assays were carried out by Dr. Colin Sharp at Edinburgh University (Sharp et al. 2010a; Matthews et al. 2015b) . The samples were screened in duplicate using baculovirus-expressed VP2 and control antigens; arbitrary unit (AU) values were calculated relative to a control sample. Due to a high background reactivity observed in our cohorts, an additional stipulation that positive samples must demonstrate a VP2-to-control optical density ratio (ODR) greater than 1.2 was added; samples falling below this cut-off were considered negative.

6.2.3 HBV testing

The South African cohorts underwent testing for Hepatitis B surface antigen (HBsAg) retrospectively from frozen sera using the Biokit enzyme immune assay (Barcelona,

Spain), following the manufacturer's instructions. For the Botswana cohort, HBsAg results were performed in Gaborone, using the Murex HBsAg v3 (DiaSorin) assay.

6.2.4 PARV4 IFN- γ ELISpot assays

Using methods as previously described (Section 2.5), IFN- γ responses to a bank of PARV4 overlapping peptides (OLPs) spanning PARV4 NS, VP and ARF proteins were quantified. The overlapping peptides in the reactive NS pools were tested in a confirmatory ELISpot assay and HLA-predictor tool was employed to predict the optimal epitopes (Larsen et al. 2011). OLPs in the reactive VP and ARF pools were not tested due to limited PBMC availability. Based on high magnitude responses by HLA-B*5703-positive subjects, a putative epitope within OLPs 9.6 and 9.7 was tested. Three truncations of this epitope (supplied by Schafer-N, Denmark; >80% purity) were synthesised as follows: 8-mer TRITMFQF, 9-mer QTRITMFQF, and 10-mer LQTRITMFQF that most closely matched the binding motif for HLA-B*57:03. Using cells from a PARV4 IgG-positive subject recruited from the Thames Valley Cohort (Patient ID N087), I tested IFN- γ ELISpot responses to serial dilutions of these three putative optimal epitope truncations.

6.2.5 PARV4 DNA

DNA was extracted from pooled serum samples (50 μ l each of 10 samples) using the DNeasy blood and tissue kit (Qiagen), according to the manufacturer's protocol. PCR reactions were performed in Edinburgh by Dr. Colin Sharp (Sharp et al. 2012).

6.2.6 HIV status

HIV-status had been ascertained prior to recruitment and was recorded prospectively. KReC children were deemed to be HIV-negative at the point of presentation

to hospital. The majority of these were also screened for HIV infection during their hospital admission episode, with the exception of three children for whom we did not confirm HIV status (these children were included in the HIV-negative group for analysis, based on the clinical data available at the time of admission).

6.2.7 HIV-viral load and CD4 T-cell count

HIV-1 RNA viral load was determined by Roche Amplicor Version 1.5 assay (Rotkreuz, Switzerland) or Abbott Laboratories m2000 platform (Abbott Park, IL, USA). CD4+ T-cell counts and percentages were measured by flow cytometry as part of routine clinical diagnostics at the recruitment centre.

6.2.8 Systematic literature review for “Hepitopes” database

A systematic literature review using two online databases, Medline and Embase was carried out to identify HBV-specific CD8+ T-cell epitopes. All the epitopes published between 1946-2016 were included in the search. Using this approach, in total 447 papers were identified. From this initial set, duplicates (n=113), papers that were not relevant or missing essential data (n=177), publications not written in English (n=26) and those that we were unable to access (n=19) were excluded, resulting in a final set of 112 papers. The following details were extracted from each of the papers in the final set:

- Citation
- Epitope location and sequence
- HLA restriction
- Experimental approach used to confirm the epitope
- Any documented association with stage of HBV disease

6.2.9 Statistical analysis

Data were analysed using GraphPad Prism v.6.0 and Fisher's exact test was used to compute significance for categorical variables in 2x2 contingency tables and Mann-Whitney *U* test for continuous non-parametric data.

6.3 Results

6.3.1 Determination of the level of PARV4, HBV, CMV and HIV co-infection in South Africa

A total of 632 adults and children from the Durban, Kimberley and Botswana cohorts were studied for PARV4 IgG prevalence and detection of PARV4 DNA. Of these 632 individuals, 593 adults and children were also tested for HBsAg prevalence. An additional set of 651 women from the Kimberley, Durban and Botswana cohorts were included in the study for HBsAg prevalence.

The overall prevalence of PARV4 IgG (43%) across all our cohorts was broadly comparable with that reported in the past for sub-Saharan African populations (Sharp et al. 2010b). The prevalence of PARV4 IgG was significantly higher in adults (55%) compared to children [$p < 0.0001$, Fisher's exact test (Fig 6.2A)]. The prevalence of PARV4 IgG was higher in the HIV positive group vs. HIV-negative sub-group [52% vs. 24% (Table 6.1)]. A significant association between PARV4-IgG status and HIV infection was observed in adults [$p = 0.002$, Fisher's exact test (Fig 6.2B)]. Although there was a trend towards higher PARV4-IgG positivity in children in the context of HIV infection, it did not reach statistical significance [$p = 0.1$; Fisher's exact test (Fig 6.2B)]. PARV4 viraemia was detected in five out of the 695 subjects tested. Four of these were children and one was an adult woman.

All but one of the children tested [221/222, (99.6%)] were HBsAg negative. The exception was an HIV uninfected child from the KReC cohort and the Hepatitis B vaccination status of this child was not known. The child was 20 months old at the time of sampling, therefore the child should have been vaccinated by this stage. However, the first dose of HBV vaccine in South Africa is not administered until age 6 weeks, by which time vertical transmission may already have occurred. The overall prevalence of HBsAg in adults (7%, Fig 6.2C) was in line with previously reported data from African populations (Matthews et al. 2014a). The HBsAg status was not significantly different between HIV-positive vs. HIV-negative sub-groups [6.9% vs. 8.3%; $p = 0.6$, Fisher's exact test; (Fig 6.2D)]. HBV infections were not statistically associated with PARV4 infection (Figure 6.3). Although a different subset of cohort for CMV was tested, the prevalence of CMV in sub-Saharan Africans was close to 100% (Table 6.1) which is also similar to previously reported data (Cannon et al. 2010), thus almost all the individuals infected with PARV4 and HBV would be co-infected with CMV (Fig 6.3).

6.3.2 Impact of PARV4 and HBV on HIV disease progression

To assess what effect PARV4 status has on HIV disease progression in children and adults, the CD4+ T-cell counts and HIV-viral loads were assessed in the context of PARV4 IgG status. PARV4 IgG positivity had no impact on CD4% or HIV-viral load ($p=0.6$, $p=0.7$, respectively; Mann-Whitney U test) in HIV infected children. The adults infected with PARV4 had a lower CD4+ T-cell count (median 310 vs. 373, $p=0.01$, Mann-Whitney U test) and a higher HIV-viral load (median 30600 vs. 13635, $p=0.0003$, Mann-Whitney U test] compared to adults who were PARV4 uninfected.

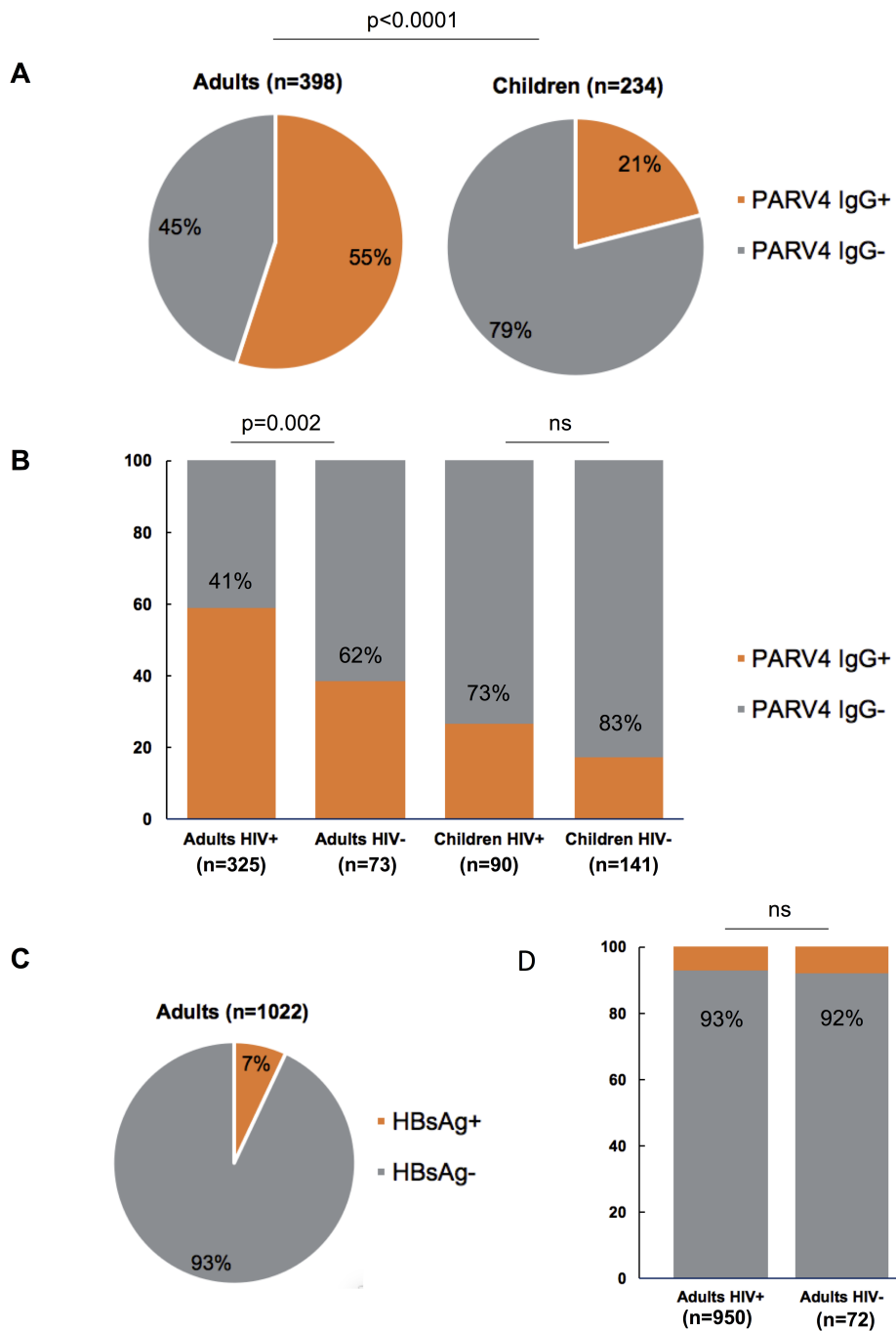


Figure 6.2: Relationship of PARV4 IgG status and HBsAg status with age and HIV status. (A) The pie charts show the prevalence of PARV4 IgG in adults and children. (B) The plot shows the relationship between PARV4 IgG status and HIV status in adults and children. (C) The pie chart shows the prevalence of HBV in adults. (D) The plot shows the relationship between HBsAg status and HIV status in adults. The denominator stated on each panel varies according to availability of relevant data.

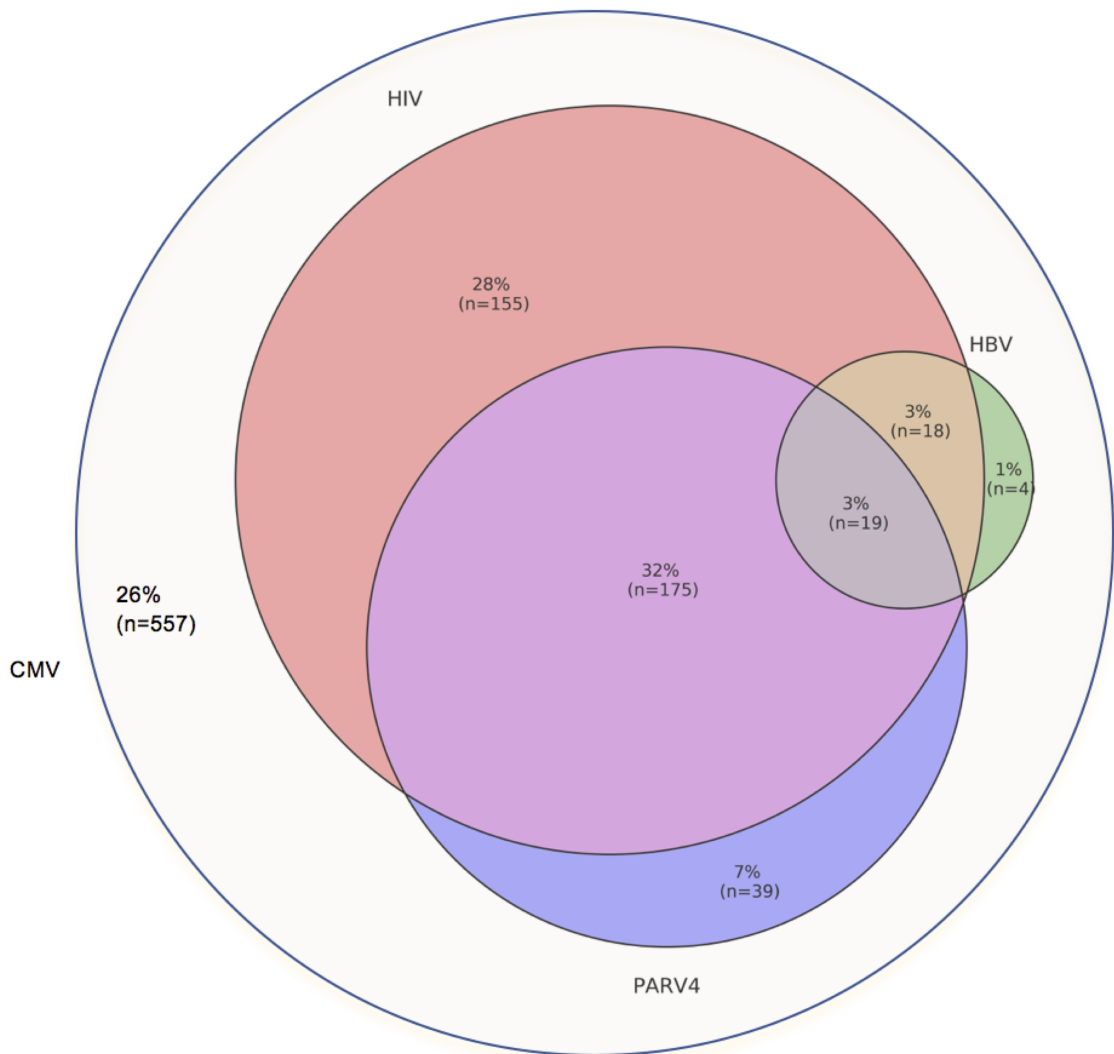


Figure 6.3: Venn diagram of CMV, HIV, PARV4 and HBV prevalence and co-infection in multicentre sub-Saharan African cohort. The Venn diagram was made using the matplotlib-venn package in Python. The circles are proportional to the parentages presented. The diagram represents: 280 HIV infected adults and 68 HIV infected children; 204 PARV4 IgG positive adults and 32 PARV4 IgG positive children; 43 HBsAg positive adults and 1 HBsAg positive child.

To assess what effect HBV infection has on HIV disease progression in adults, CD4+ T-cell counts were assessed in the context of HBV infection. Studies in the past have shown that HBV/HIV co-infection has an adverse effect on HIV related disease outcome (Bonacini et al. 2004; Chun et al. 2012; Matthews et al. 2014a). However, in this study HBsAg positivity had no significant impact on the CD4+ T-cell count or HIV-viral load in HIV infected adults.

6.3.3 High breadth and magnitude of PARV4-specific CD8+ T-cell responses in children and adults

IFN- γ ELISpot assays were used to characterise PARV4-specific CD8+ T-cell responses using a panel of overlapping peptides spanning the entire PARV4 proteome in 14 PARV4 IgG positive individuals (seven children from Kimberley, South Africa, and seven adults enrolled via the Thames Valley Cohort). We used a combination of ELISpot assays, peptide-HLA class I binding studies, and bioinformatics prediction to predict novel CD8+ T-cell epitopes in PARV4 (described in more detail in section 3.2.2). Among adults and children screened for ex vivo CD8+ T-cell responses, high magnitude IFN- γ ELISpot responses to peptides spanning all three PARV4 proteins were observed, particularly within NS and ARF proteins (Fig 6.4A,C). The magnitude of PARV4 specific CD8+ T-cell responses was higher in children compared to adults (median 525 vs. 465 SFC/106 PBMCs), however this did not reach statistical significance [$p=0.7$, Mann-Whitney U test] (Fig 6.4E). Similarly, the number of PARV4 specific CD8+ T-cell responses was higher in children compared to adults (median 6 vs. 3), but this difference was also not statistically significant [$p=0.12$, Mann-Whitney U test] (Fig 6.4F).

The individual 15mer OLPs in the reactive NS pools that were most targeted

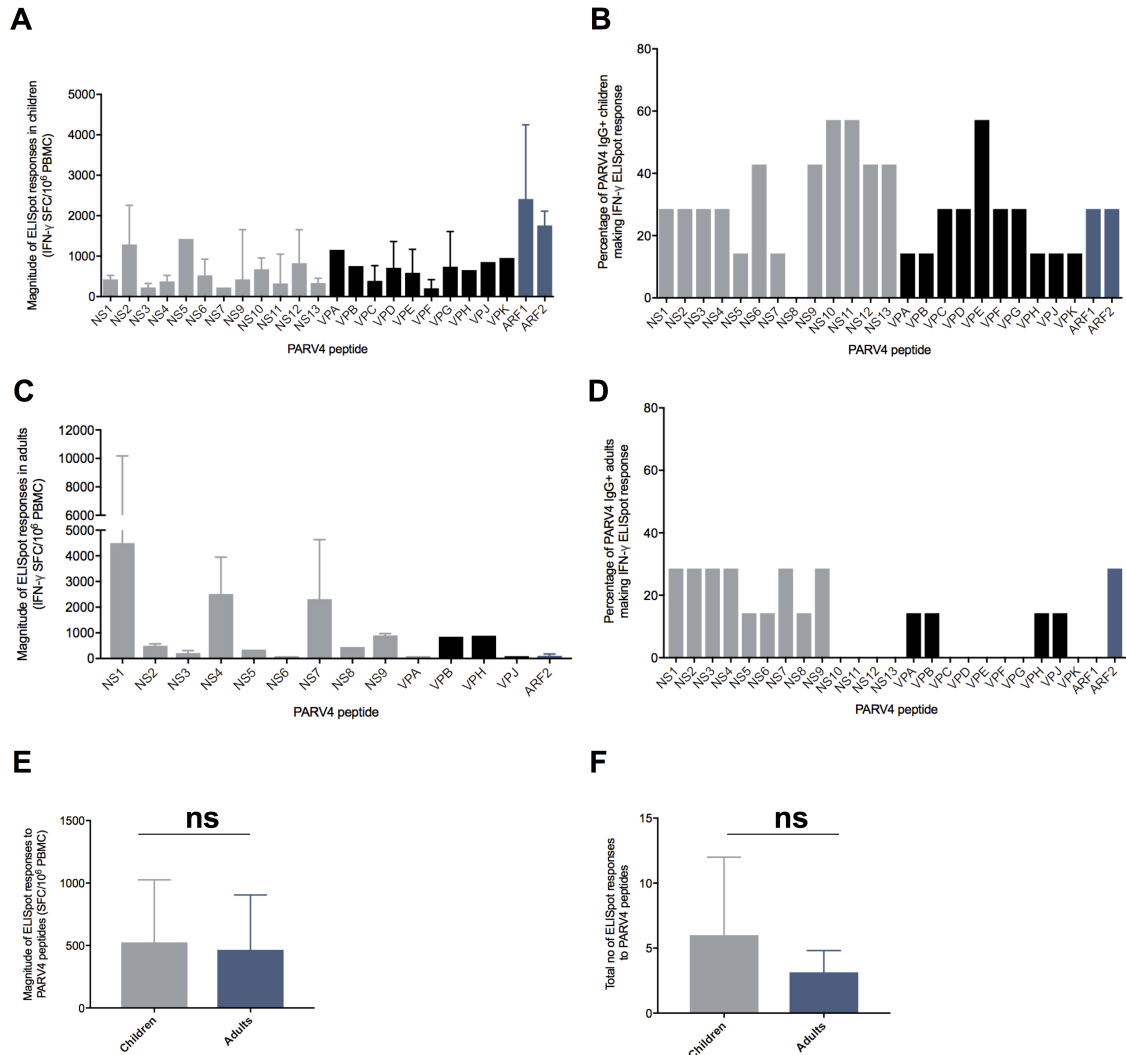


Figure 6.4: IFN- γ CD8+ T-cell responses to PARV4 peptides determined by ELISpot assay. (A) The magnitude of PARV4-specific CD8+ T-cell responses in children. (B) Percentage of PARV4 IgG positive children (n=7) who made a response to PARV4 peptides. (C) The magnitude of PARV4-specific CD8+ T-cell responses in adults. (D) Percentage of PARV4 IgG positive adults (n=7) who made a response to PARV4 peptides. (E) The plot shows the difference in magnitude of responses between children and adults. (F) The plot shows the difference in number of PARV4-specific CD8+ T-cell responses between adults and children. Boxes represent the median and whiskers represent the IQR. Responses to NS peptides are shown in grey, to VP peptides in black, and to ARF in blue. P-values by Mann-Whitney *U* test.

were tested using an ELISpot assay; we were unable to test the remaining reactive pools due to insufficient availability of PBMC. Table 6.4 shows the predicted optimal epitopes presented by the relevant HLA class I molecule. One predicted optimal epitope was tested using three possible peptide truncations found within OLPs 9.6 and 9.7: 8-mer TRITMFQF, 9-mer QTRITMFQF, and 10-mer LQTRITMFQF. This epitope was selected due to high magnitude responses directed at the OLPs corresponding to this epitope. These three truncations matched most closely with the binding motif for HLA-B*57:03 (namely A/S/T at position 2 and F/W/Y at the C-terminal position of the epitope) (Kloverpris et al. 2012). The epitope QF9 (QTRITMFQF) is the most likely HLA-B*57:03 restricted epitope as shown by the IFN- γ ELISpot responses to serial dilutions of epitope truncations (Fig 6.5). All the published PARV4 Genotype 3 sequences were downloaded from GenBank to assess the variability in the region where the predicted epitope sits. Analysis of 11 downloaded sequences revealed that the QF9 epitope sits in the NS protein between residue 442-450 (Fig 6.1). The downloaded sequences were from Ghana, South Africa and Cote d'Ivoire and showed no sequence polymorphism in the region where the epitope sits, or the flanking regions.

6.3.4 Impact of PARV4 infection on CMV- and HIV-specific CD8+ T-cell responses

Given the high co-infection rate of PARV4 with CMV and HIV in sub-Saharan Africa, I looked at the effect of PARV4 IgG positivity on CMV- and HIV-specific CD8+ T-cell responses. Overall, the PARV4 specific CD8+ T-cell responses were similar in magnitude and breadth to CMV-specific CD8+ T-cell responses (Fig 6.6A-B) [$p = 0.4$, $p=0.7$, respectively; Mann-Whitney U test].

OLP	OLP Sequence	HLA association	HLA Restrictor predicted epitope
2.5	AVTLPCRVKPSMFLQ	HLA-A03:01	VTLPCRVK
		HLA-A30:01	RVKPSMFL
		HLA-B58:01	VTLPCRVKPSM
		HLA-B57:03	VTLPCRVKPSM
		HLA-C06:02	RVKPSMFL
3.5/3.6	VFMAGYLVAPVWTL YLVAPVWTLWNIRE	HLA-B58:01	MAGYLVAPVW
		HLA-B57:03	MAGYLVAPVW
		HLA-C06:02	FMAGYLVAPV
		HLA-C18:01	VFMAGYLV
4.1/4.2	KLPLNDCYYAWTNID DCYYAWTNIDRFEAA	HLA-B42:01	LPLNDCYYAW
		HLA-A23:01	CYYAWTNIDRF
		HLA-A29:02	CYYAWTNIDRF
		HLA-B58:01	YAWTNIDRF
4.3	WTNIDRFEAAVLSVR	HLA-C17:01	NIDRFEAAVL
4.7	GPQGAIALPFTDAPP	HLA-B42:01	GPQGAIAL
4.8/4.9	IALPFTDAPRTPAA TDAPRTPAAEGVPP	HLA-B42:01	APRTPAA
6.2/6.3	NALQIAKREMVLAHP AKREMVLAHPLLSYL	HLA-A23:01	MVLAHPLL
		HLA-A29:02	MVLAHPLLSY
		HLA-B42:01	MVLAHPLLSYL
		HLA-B58:01	MVLAHPLLSY
6.4/6.5	VLAHPLLSYLTKNAS LLSYLTKNASAFEES	HLA-A23:01	SYLTKNASAF
		HLA-A29:02	YLTKNASAF
		HLA-B42:01	YLTKNASAF
		HLA-B58:01	LSYLTKNASAF
6.8	NKVAQLFSLNGYNPV	HLA-C06:02	SYLTKNASAF
		HLA-A29:02	QLFSLNGY
		HLA-B42:01	KVAQLFSL
		HLA-B58:01	
9.6/9.7	QHQEALQTRITMFQF LQTRITMFQFQRMVP	HLA-A30:01	QTRITMFQ
		HLA-B58:01	LQTRITMFQF
		HLA-B57:03	QTRITMFQF
		HLA-C06:02	TRITMFQF
		HLA-A03:01	RITMFQFQR
9.8	TMFQFQRMVDPGLAP	HLA-C06:02	FQFQRMVDPGL
10.2	VRSFFKLGEQELNMK	HLA-A23:01	SFFKLGEQEL
		HLA-B58:01	RSFFKLGEQEL

Table 6.4: HLA associations with the immunogenic 15mer OLPs in NS. The table shows the OLPs that were observed to be immunogenic in a confirmatory ELISpot assay. The HLA associations and the optimal epitopes predicted with HLA-Restrictor tool are shown.

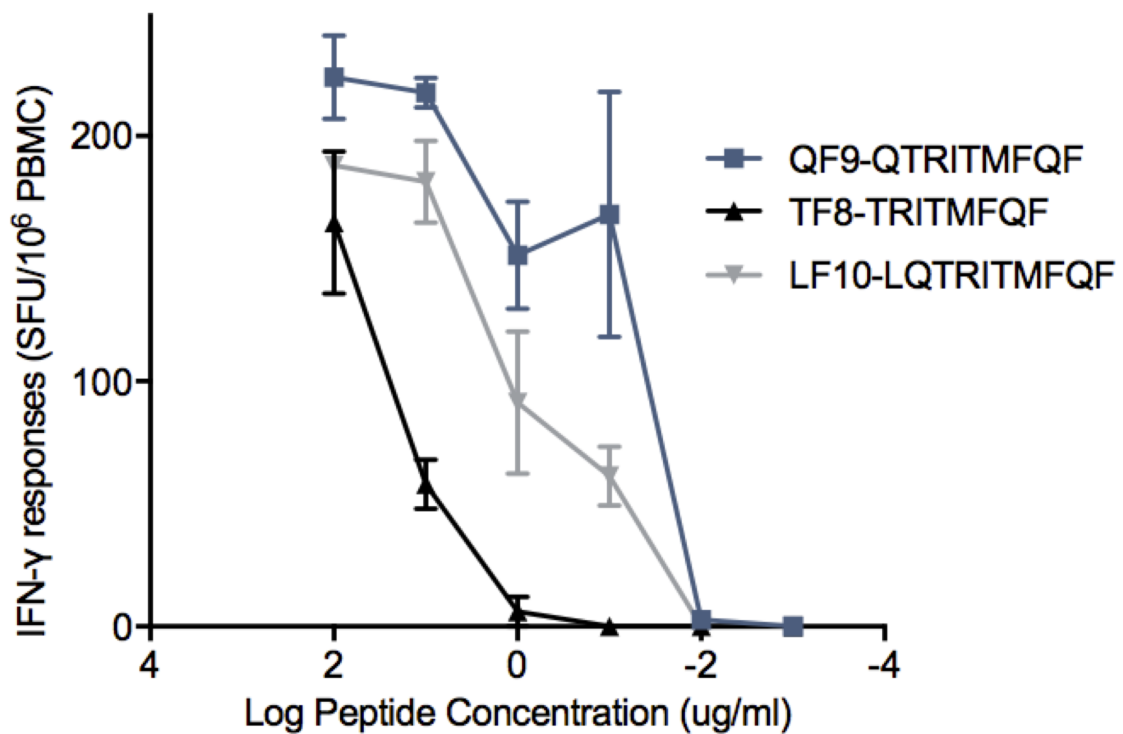


Figure 6.5: Validation of a novel HLA-B*5703-restricted epitope in PARV4 NS protein using peptide titration assay. PBMCs from PARV4 IgG-positive adult subject N087 (HIV-positive adult recruited via the Thames Valley cohort, HLA class I genotype HLA-A*0301/-A*3001/-B*5703/-B*5801/-C*0602/-C*1801) were screened by IFN- γ ELISpot for responses to peptide truncations derived from the PARV4 NS protein (sequences within OLPs 9.6 and 9.7) at different concentrations. Plots and error bars show the mean and SEM of assays performed in triplicate and confirm QF9 as the likely optimal epitope.

The magnitude of CMV-specific CD8⁺ T-cell responses was significantly higher in PARV4 IgG⁺ children (n=4) compared to PARV4 IgG⁻ children (n=16) (Fig 6.6C; median 800 vs 180 SFC/106 PBMC, p=0.01; Mann-Whitney *U* test). In adults, the magnitude of CMV specific CD8⁺ T-cell responses were not statistically different between PARV4 IgG⁺ group (n=6) and PARV4 IgG⁻ group (n=40).

The number of HIV-specific CD8⁺ T-cell responses were similar in the PARV IgG⁺ (n=10) and PARV IgG⁻ children (n=29) [p=0.5, Mann-Whitney *U* test]. Similarly, in adults, the HIV-specific CD8⁺ T-cell responses in PARV⁺ (n=17) and PARV4⁻ (n=17) were similar (p=0.7, Mann-Whitney *U* test).

6.3.5 Impact of HBV on HIV-specific CD8⁺ T-cell responses

CD8⁺ T-cell responses are well recognised as an important mediator of HIV disease control (Kiepiela et al. 2007; Payne et al. 2009; Goulder and Walker 2012). Considering that HBV/HIV co-infection leads to worse HIV disease outcome (Bonacini et al. 2004; Chun et al. 2012; Matthews et al. 2014a), it was hypothesised that there would be lower magnitude CD8⁺ responses in HIV/HBV co-infected patients. The impact of HBV co-infection on HIV-specific CD8⁺ T-cell responses was measured using IFN- γ ELISpot assays for 325 subjects from the Durban cohort. However, there was no significant difference in the breadth, magnitude, or protein-specificity of IFN- γ responses to HIV in the presence or absence of HBsAg.

6.3.6 HBV epitopes

The live interactive database “Hepitopes” containing all the published class I epitopes in HBV is available online [<http://www.expmedndm.ox.ac.uk/hepitopes>] (Lumley et al. 2016). Each HBV epitope entry contains the optimal and/or overlapping

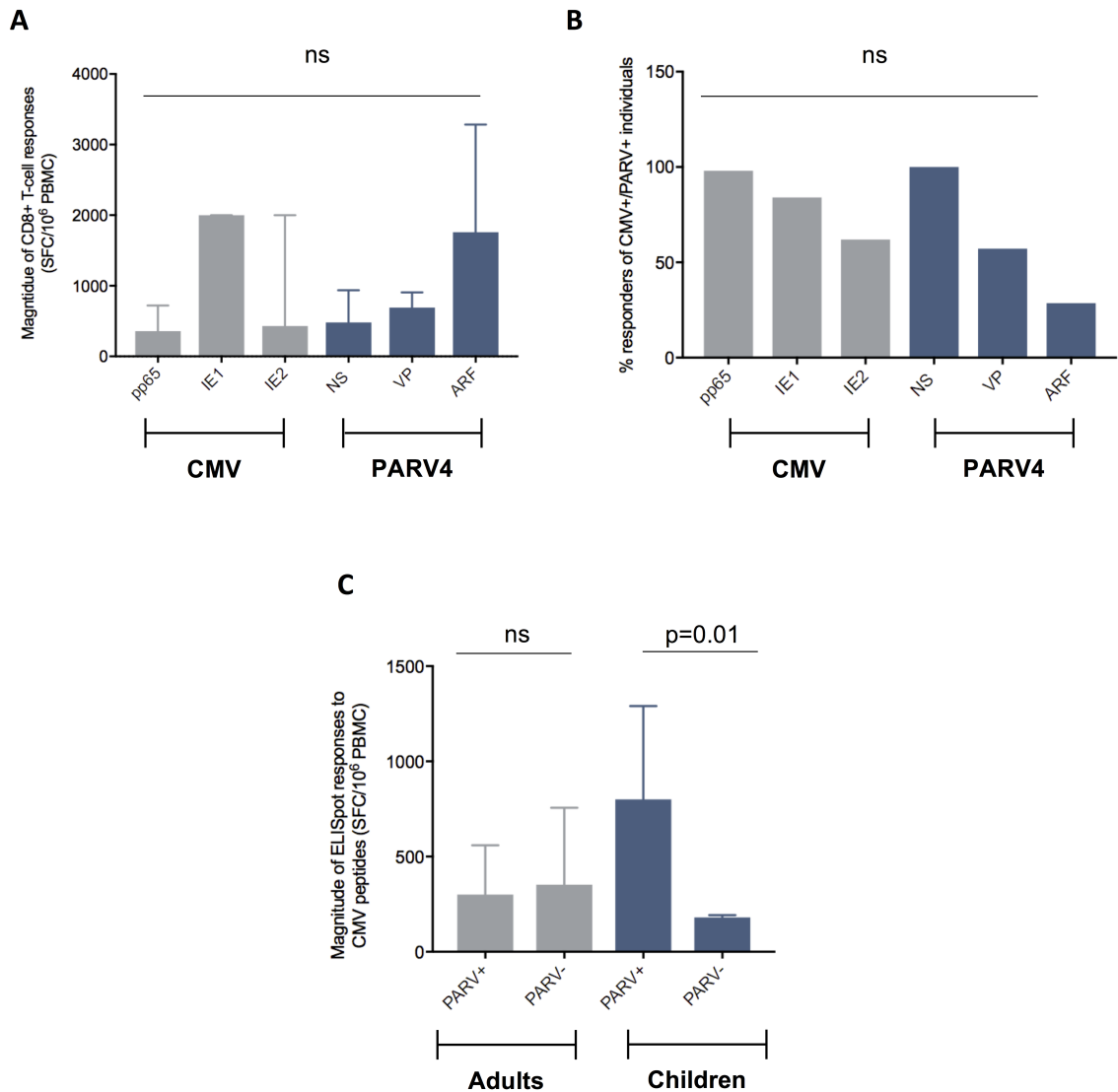


Figure 6.6: Impact of PARV4 IgG serostatus on CMV and PARV4-specific CD8+ T-cell responses. (A) The magnitude of CD8+ T-cell responses to CMV (pp65, IE1 and IE2) and PARV4 (NS, VP and ARF) proteins. (B) The percentage of PARV4 IgG+ individuals reacting to PARV4 proteins (NS, VP and ARF) and percentage of CMV+ individuals reacting to CMV proteins (pp65, IE1 and IE2). CMV responses are shown in black and PARV4 responses in blue. (C) The magnitude of CMV-specific responses in PARV4 IgG+ and PARV4 IgG- group. The responses in adults are shown in black and the responses in children are shown in blue. Boxes represent the median and whiskers represent the IQR. P-values by Mann-Whitney *U* test.

peptide sequence and defined HLA-class I restriction (Fig 6.7). Currently 199 distinct epitopes have been identified and the majority of epitopes are in HBV Pol protein (89/199; 44%). The most commonly published epitopes are HLA-A*02 restricted (n=87/199; 44%), but this likely reflects a bias in experimental work focusing on this common allele rather than a true biological enrichment of HLA-A*02 alleles in HBV.

Given the high magnitude HLA-B*4403 mediated T-cell response to CMV, and the enrichment of this allele in African populations, I sought evidence for a B*44 restricted response to HBV. The two identified HLA-B*44:03-restricted epitopes, VELLSFLPSDF and AELLAACF(A) (Sidney et al. 2003), both have glutamic acid at position 2, similar to the CMV-HLA-B*44:03 restricted NW8-epitope. The HBV HLA-B*44:03-restricted epitopes were shown to also be restricted by other HLA alleles in the B44 superfamily (HLA-B*1801, B*4002, B*4402, B*4501) (Sidney et al. 2003). This phenomenon of an epitope being targeted by multiple closely related alleles in a superfamily is rare in relation to other virus-specific CD8+ T-cell responses. The data shown in chapter 3 and previous studies have reported that the difference of even a few amino acids between closely-related HLA class I molecules can have a substantial impact on the CD8+ T-cell immunodominance hierarchy (Leslie et al. 2006; Carlson et al. 2012).

NS=not specified, NA=not available.

Columns to show: HBV protein Optimal amino acid sequence(s) HLA restriction(s) Database Authors Title Journal Publication year DOI

Show 10 entries Filter:

HBV protein	Optimal amino acid sequence(s)	HLA restriction(s)	Database	Authors	Title	Journal	Publication year
C	VELLSFLPSDF	B*1801, B*4002, B*4402, B*4403	Embase	Sidney J, Southwood S, Pasquetto V, Sette A.	Simultaneous Prediction of Binding Capacity for Multiple Molecules of the HLA B44 Supertype.	Journal of Immunology. 171(11): 5964-5974, 2003.	2003
Poi	AELLAACF(A)	B*1901, B*4002, B*4402, B*4403, B*4501	Embase	Sidney J, Southwood S, Pasquetto V, Sette A.	Simultaneous Prediction of Binding Capacity for Multiple Molecules of the HLA B44 Supertype.	Journal of Immunology. 171(11): 5964-5974, 2003.	2003

Showing 1 to 2 of 2 entries (filtered from 552 total entries) Previous 1 Next

Figure 6.7: Sample screenshot of the “Hepitope” database. The database can be searched or sorted by any of the column headers (outlined in black). The sample shows the result for all the HLA-B*44:03 restricted epitopes in the database. Multiple HLA alleles are displayed if the same epitope is restricted by more than one HLA allele.

6.4 Discussion

The PARV4 IgG seroprevalence in our African cohort was in line with previously reported African studies (Sharp et al. 2010b) but higher than in published studies of Western European populations. The prevalence of PARV4 was higher in adults and strongly associated with HIV infection in adults, but no association was observed with HBV. PARV4 serostatus was associated with HIV disease progression in terms of lower CD4⁺ T-cell counts and higher viral load, but this effect was seen in adults only. PARV4 infection had no impact on the HIV-specific CD8⁺ T-cell responses in adults or children.

The CD8⁺ T-cell responses to PARV4 spanned the entire viral proteome in children and adults. The responses were of particularly high magnitude in certain regions of NS and ARF proteins in children. Using a peptide titration assay, a novel PARV4 NS-specific HLA-B*57:03-restricted epitope was confirmed. The next experimental step would involve synthesising the tetramer to confirm and better characterise this PARV4 CD8⁺ T-cell response.

The PARV4-specific CD8⁺ T-cell responses were observed in the absence of detectable viraemia in children and adults. This supports the view that PARV4 may behave similarly to chronic herpes viruses, including CMV, in which a latent reservoir triggers episodic reactivation, maintaining T-cell responses in the long term (Adland et al. 2015a). Furthermore, the PARV4-specific CD8⁺ T-cell responses in subjects from our cohorts were similar in terms of magnitude and breadth to CMV-specific responses in a different subset of subjects from the same cohorts (Kimberley and TVC cohort). The magnitude of PARV4-specific CD8⁺ T-cell responses was slightly smaller in adults compared to children. This is in contrast

to what is seen in CMV, where the size of CD8+ T-cell responses is significantly bigger in adults, due to memory inflation with age (Klenerman and Oxenius 2016).

Interestingly, in PARV4 IgG+ children, the magnitude of CMV-specific CD8+ T-cell responses was higher compared to the PARV4 IgG- children (both groups were HIV positive). A similar increase in magnitude of CMV-specific CD8+ T-cell responses is usually expected in context of HIV infection (Gerna et al. 1998). The increase seen during HIV infection is attributed to immunosuppression and increase in immune activation (Paiardini and Müller-Trutwin 2013; Giorgi and Detels 1989), which results in suboptimal control of CMV leading to increase in magnitude of CMV-specific CD8+ T-cells. However, the increase in magnitude of CMV-specific CD8+ T-cell responses in the context of PARV4 infection has not been previously reported.

The work presented here adds to our understanding of PARV4-specific CD8+ T-cell responses in children and the impact of PARV4 infection on HIV, CMV and HBV, but more studies with greater sample numbers are required.

The overall prevalence of HBsAg in adults from our African cohorts was 7%, which is in line with previously published data on African populations (Matthews et al. 2014a). The prevalence of HBsAg in children was very low (0.4%), which was expected, as the introduction of prophylactic HBV vaccinations in South African infants since 1995 has significantly reduced the HBsAg prevalence in children (Amponsah-Dacosta et al. 2014). HBV was not associated with HIV disease progression in terms of CD4+ T-cell counts or HIV viral loads. HBsAg positivity had no impact on the HIV-specific CD8+ T-cell responses and we had no data to study the impact of HBV on CMV-specific CD8+ T-cell responses.

Although, HBV — like CMV, PARV4 and HIV — is a chronic viral infection, the size of HBV-specific CD8+ T-cell responses during chronic infection is significantly less. It is believed that during chronic HBV infection, there is a large intrahepatic influx of non-antigen specific CD8+ T-cells (Maini et al. 2000). Owing to the inflamed microenvironment, the ability to produce IL-2 is impaired in these cells (Das et al. 2008). IL-2 is required to drive CD8+ T-cells to proliferate in settings of impaired CD4 help (Zimmerli et al. 2005), such as the CD4-depleted liver environment in chronic HBV (Doherty et al. 1999). The impairment in IL-2 production manifested by non-antigen-specific CD8+ T-cells could therefore prevent the HBV-specific CD8+ T-cells from expanding and surviving (Das et al. 2008). It is, therefore important to study HBV-specific CD8+ T-cells in chronic HBV for successful development of therapeutic HBV vaccines.

In summary, the data presented here adds to our knowledge, however, the sample size is limited for some of the analysis. Furthermore, the attempt to combine analysis of individuals from different geographical locations could be misleading as the cohort demographics vary. Despite the limitations, the dataset presented here is of importance in providing insight into the extent and characteristics of CMV, HIV, HBV and CMV co-infections in certain African populations of children and adults, where these infections are co-endemic.

Chapter 7

Conclusions

The high prevalence of chronic viral infection and co-infection in Africa is associated with long-term morbidity and mortality. Pathogens such as CMV, HBV and PARV4 are co-endemic in this region. However, very little is known about the prevalence and immunology related to these chronic viruses in African populations. There is a need to advance or identify new therapeutic interventions/vaccines in instances where these viruses are detrimental to health.

In this thesis, I have focused primarily on characterising the CD8⁺ T-cell immune responses to CMV. By adolescence, CMV infection is almost universal in African individuals. I studied the high-frequency CMV specific CD8⁺ T-cell responses in African cohorts and identified novel epitopes that will increase our understanding of CMV-specific CD8⁺ T-cell response in African populations. The rationale behind this study was to develop a better insight into the immunological response to CMV in African cohorts, with the aim of furthering our understanding of the interplay between host and virus. Studying the immune response against CMV is also important because of its pathogenic potential in context of other viral infections/immunosuppression.

I found that CMV leads to oligoclonal expansions that can constitute up to 18% of the total CD8+ T-cell population, similar to what has been reported in earlier studies (Khan et al. 2002; Karrer et al. 2003). Similar to adults, the high magnitude CMV-specific CD8+ T-cell responses were also observed among children. These findings are consistent with previous studies of children (Appay et al. 2002; Komatsu et al. 2006; Miles et al. 2007, 2008).

Given that the CMV-specific CD8+ T-cell responses are of a high magnitude, it may impact the ability of the immune system to respond effectively to other pathogens. CMV disrupts the CD4/CD8 ratio and is therefore associated with “immune risk profile” (IRP) (Olsson et al. 2000) in older adults. “IRP” predicts higher mortality from infectious diseases (Looney et al. 1999; Pawelec 2014). The fact that a large proportion of the CD8+ T-cell population is focused on one virus in early childhood is somewhat surprising (Appay et al. 2002; Komatsu et al. 2006; Miles et al. 2007, 2008). On one hand, this can be detrimental, because the TCR repertoire is made up of approximately 25 million TCRs (Attaf et al. 2015), and if a single CMV response predominates, it may impair the ability to respond to non-CMV infections. For example, in the context of CMV/HIV co-infection, there is a worse HIV disease outcome (Kovacs et al. 1999) and CMV has also been implicated in malignant tumours (Michaelis et al. 2009; Pourgheysari et al. 2010; Ranganathan et al. 2012). On the other hand, CMV is thought to underlie rudimentary aspects of immunosenescence (Looney et al. 1999; Turner et al. 2014). Whilst this remains controversial (Pawelec 2014), one can speculate that the maturation of the immune system driven by CMV could help protect against early life infections in childhood (Furman et al. 2015). CMV can therefore be considered essential for immune maturation, in a similar way to gut microflora shaping the host mucosal immunity

(Prendergast et al. 2012). Whether CMV plays a beneficial and/or detrimental role, these studies indicate that there is a need to understand how CMV modulates the human immune system.

The TCR repertoire that shapes the clonal composition of CMV-specific CD8+ T-cell responses conferring immunodominance still remains elusive especially in African populations. Therefore, I dissected the TCR repertoire composition underlying one immunodominant epitope, namely NW8 restricted by a common African allele HLA-B*44:03, using high-throughput sequencing. The TCR repertoire analysis revealed a preferential use of a restricted set of V and J gene segments and a large number of shared (“public”) TCRs in multiple individuals. The observation that recognition of the virus by the host is restricted to a particular set of TCRs is a point of interest (Venturi et al. 2008b). CMV-specific clonal expansions have been implicated in post stem cell transplant (SCT) immune dysfunction (Suessmuth et al. 2015). CMV-specific public clonotypes shared amongst certain individuals could prove to be beneficial for monitoring SCT outcomes. Considering that public TCRs have high affinities for their ligands, they can also prove to be of a use in the development of vaccines and adoptive therapies (Miles et al. 2006). From the perspective of clinical application, with enough number of patients and follow up, identification of public TCRs might have both prognostic value and can aid in development of adoptive therapies (Dong et al. 2004; Price et al. 2009; van Bockel et al. 2011).

Given that nearly 90% of the African children acquire CMV by one year of age, CMV is one of the first immune challenges their immune systems are exposed to. Furthermore, the timing of CMV acquisition in African children coincides with the timing of vaccinations. Therefore, studying the ontogeny of CMV responses in

early childhood may provide us with an insight into how the immune system develops in response to natural infections and vaccinations during childhood. Studies looking at the changes of the CMV specific TCR repertoire have been very limited and those that have been undertaken (Hadrup et al. 2006; Day et al. 2007; Iancu et al. 2009; Klarenbeek et al. 2012) are usually in adults and pre-date the next generation sequencing era.

I studied the evolution of TCR repertoire in children and mothers over a span of 10 years using TCR- α and TCR- β chain next generation sequencing. I found that the TCR repertoires of children and mothers exhibit similarities, often dominated with public TCRs. The TCR usage over the course of CMV infection in both children and mothers displayed fluctuations. Considering that the adults had been infected with CMV for multiple decades, it was anticipated that their immune system would adapt to the virus and TCR usage would exhibit stability over time, however, a great degree of change in TCR usage was observed among adults. Contrary to previous findings, the TCR repertoires of children were highly skewed with limited TCR diversity. To my knowledge, this study was the first to delineate the CMV specific TCR-repertoire changes in the first 10 years of life. Next steps in TCR repertoire studies need to include high-throughput sequencing of both the TCR- α and TCR- β chains, simultaneously. To date, the available methods for combined TCR α/β sequencing are through single cell emulsion PCR where amplification products of both chains are fused with a chimeric primer through bridge PCR (Turchaninova et al. 2013). This approach is currently low-throughput but it has the potential for high-throughput application.

In this thesis, although the main focus was to further the understanding of CD8+ T-cell responses to CMV during early life, I describe my contribution to studies to

characterise PARV4 and HBV in African cohorts of children and adults that have high co-infection rates of HIV and CMV. There are multiple compelling reasons to further our understanding of characteristics and immune responses that control PARV4 and HBV infection: (i) pathogenesis that might be particularly relevant in high risk individuals, including children and those infected with chronic viral infection, such as CMV and HIV; (ii) the potential impact these viruses may have on the immune ontogeny, outcome of other viral infections and vaccinations.

I contributed to a study to characterise the prevalence of HBV, PARV4 and rate of co-infection with HIV/CMV (Matthews et al. 2014b, 2015b, 2015a, 2016; Lumley et al. 2016; Sharp et al. 2017). PARV4 was more prevalent and strongly associated with HIV infection in adults (55% in adults vs. 21%). An association between positive PARV4 IgG status and more advanced HIV disease has previously been reported in European studies but this was attributed to the potential confounding influence of co-infecting HCV in PARV4-positive individuals (Simmons et al. 2012). The HBV prevalence was determined in a large composite cohort of African women and children. As expected, the prevalence in children was low due to introduction of prophylactic vaccine since 1995, and in adults, the prevalence of 7% was comparable to previously reported studies (Matthews et al. 2014a). I studied the PARV4 specific CD8+ T-cell responses in children and adults and identified a novel PARV4 epitope which can be used in T-cell assays to characterise the PARV4-specific immune responses in African individuals, which are still largely undetermined. Overall, the data presented in chapter 6 will further our knowledge about the extent and characteristics of CMV, HIV, HBV and PARV4 co-infections in certain African populations of children and adults, where these infections are co-endemic. Looking forward, there are still a number of unresolved questions: (1) What is the pathogenic role of PARV4, if any? (2) The impact of PARV4 infection

on other pathogen disease outcomes and vaccinations, and (3) the role of PARV4-specific CD8+ T-cells or other immune responses in controlling PARV4 infection.

In summary, this thesis furthers the understanding of the CD8+ T-cell responses to chronic viral infections prevalent in Africa, where there is a paucity of data. The TCR studies revealed that CMV-specific responses in children and adults are not stable and undergo changes, indicating that CMV can continuously modulate the TCR repertoire.

Appendix

HLA-A alleles	PF (%)	HLA-B alleles	PF (%)	HLA-B alleles	PF (%)	HLA-C alleles	PF (%)
A*0101	5.7%	B*0702	12.5%	B*4403	18.8%	C*0102	0.9%
A*0103	0.5%	B*0705	2.7%	B*4415	0.4%	C*0202	4.3%
A*0201	16.0%	B*0801	7.8%	B*4501	12.2%	C*0210	13.7%
A*0202	4.7%	B*1302	2.4%	B*4901	2.7%	C*0217	0.9%
A*0205	10.4%	B*1303	0.8%	B*5001	0.4%	C*0302	4.7%
A*0214	0.9%	B*1401	3.9%	B*5101	1.2%	C*0303	1.4%
A*0301	6.6%	B*1402	2.7%	B*5201	1.2%	C*0304	9.5%
A*1101	0.5%	B*1501	1.6%	B*5301	8.6%	C*0401	26.5%
A*2301	20.8%	B*1503	14.9%	B*5601	0.4%	C*0407	0.5%
A*2402	6.1%	B*1510	11.4%	B*5702	1.2%	C*0501	0.9%
A*2601	2.4%	B*1516	2.4%	B*5703	6.7%	C*0602	25.1%
A*2901	2.8%	B*1518	0.4%	B*5801	12.2%	C*0701	19.0%
A*2902	11.8%	B*1801	4.3%	B*5802	15.3%	C*0702	14.7%
A*2911	1.4%	B*1803	1.2%	B*7301	0.4%	C*0704	2.8%
A*3001	14.2%	B*2703	0.4%	B*8101	8.2%	C*0802	6.6%
A*3002	16.0%	B*2705	1.2%	B*8201	0.4%	C*0804	5.2%
A*3004	3.8%	B*3501	5.1%	B*8802	0.4%	C*1202	0.9%
A*3101	1.9%	B*3502	0.4%			C*1203	3.3%
A*3201	6.1%	B*3701	0.4%			C*1402	1.9%
A*3301	3.8%	B*3802	0.8%			C*1403	1.4%
A*3303	3.8%	B*3901	0.4%			C*1502	0.5%
A*3402	9.0%	B*3910	3.9%			C*1505	2.4%
A*3601	2.4%	B*3924	0.8%			C*1601	13.3%
A*4301	1.9%	B*4001	0.8%			C*1602	0.5%
A*6601	6.6%	B*4002	0.8%			C*1700	1.4%
A*6602	1.4%	B*4006	0.4%			C*1701	18.0%
A*6801	4.2%	B*4016	0.4%			C*1801	10.4%
A*6802	14.6%	B*4101	1.2%				
A*6827	0.5%	B*4201	16.9%				
A*7401	11.8%	B*4202	2.0%				
A*8001	0.9%	B*4402	0.4%				

Table A1: HLA frequencies of cohort (Chapter 3)

A) PP65

Peptide No.	Sequence
1	MESRRRCPEMISVL
2	GRRCPMISVLGPIS
3	PEMISVLGPISGHVL
4	SVLGPISGHVLKAVF
5	PISGHVLKAVFSRGD
6	HVLKAVFSRGDTPVL
7	AVFSRGDTPVLPHET
8	RGDTPVLPHETRLLQ
9	PVLPHETRLLQTGIOH
10	HETRLLQTGIHVRVS
11	LLQTGIHVRVSQPSL
12	GIHVRVSQPSLILVS
13	RVSQPSLILVSQYTP
14	PSLILVSQYTPDSTP
15	LVSQYTPDSTPCHRG
16	YTPDSTPCHRGDNQL
17	STPCHRGDNQLVQVQH
18	HRGDNQLVQVQHTYFT
19	NQLVQVQHTYFTGSEV
20	VQHTYFTGSEVENVS
21	YFTGSEVENVSVNOH
22	SEVENVSVNVHNPTG
23	NVSVNVHNPTGRSIC
24	NVHNPTGRSICPSQE
25	PTGRSICPSQEPMSI
26	SICPSQEPMSIYVYA
27	SQEPMSIYVYALPLK
28	MSIYVYALPLKMLNI
29	VYALPLKMLNIPSIN
30	PLKMLNIPSINVHHY
31	LNIPSINVHHYPSAA
32	SINVHHYPSAAERKOH
33	HHYPSAAERKHRHLP
34	SAAERKHRHLPVADA
35	RKHRHLPVADAVIHA
36	HLPVADAVIHASGKQ
37	ADAVIHASGKQMWQA
38	IHASGKQMWQARLTV
39	GKQMWQARLTVSGLA
40	WQARLTVSGLAWTRQ
41	LTVSGLAWTRQQNQW
42	GLAWTRQQNQWKEPD
43	TRQQNQWKEPDVYYT
44	NQVVKEPDVYYTSAFV
45	EPDVYYTSAFVFPTK
46	YYTSAFVFPTKDVAL
47	AFVFPTKDVALRHW
48	PTKDVALRHWCAHE
49	VALRHWCAHELVCS
50	HWCAHELVCSMENT

51	AHELVCSMENTRATK
51	AHELVCSMENTRATK
52	VCSMENTRATKMQUI
53	ENTRATKMQUIGDQY
54	ATKMQUIGDQYVKVY
55	QVIGDQYVKVYLESF
56	DQYVKVYLESFCEDV
57	KVYLESFCEDVPSGK
58	ESFCEDVPSGKLFMOH
59	EDVPSGKLFMHVTLG
60	SGKLFMHVTLGSDVE
61	FMHVTLGSDVEEDLT
62	TLGSDVEEDLTMTRN
63	DVEEDLTMTRNPQPF
64	DLTMTRNPQPFMRPOH
65	TRNPQPFMRPHERNG
66	QPFMRPHERNGFTVL
67	RPHERNGFTVLCPKN
68	RNGFTVLCPKNMIK
69	TVLCPKNMIKPGKI
70	PKNMIKPGKISHIM
71	IIPKISHIMLDVA
72	GKISHIMLDVAFTSOH
73	HIMLDVAFTSHEHFG
74	DVAFTSHEHFGLLCP
75	TSHEHFGLLCPKSIP
76	HFGLLCPKSIPGLSI
77	LCPKSIPGLSISGNL
78	SIPGLSISGNLLMNG
79	LSISGNLLMNGQQIF
80	GNLLMNGQQIFLEVQ
81	MNGQQIFLEVQAIRE
82	QIFLEVQAIRETVEL
83	EVQAIRETVELRQYD
84	IRETVELRQYDPVAA
85	VELRQYDPVAALFFF
86	QYDPVAALFFFDIDL
87	VAALFFFDIDLLQR
88	FFFDIDLLQRGPQY
89	IDLLQRGPQYSEHP
90	LQRGPQYSEHPTFTS
91	PQYSEHPTFTSQYRI
92	EHPTFTSQYRIQGKL
93	FTSQYRIQGKLEYROH
94	YRIQGKLEYRHTWDR
95	GKLEYRHTWDRHDEG
96	YRHTWDRHDEGAAQG
97	WDRHDEGAAQGDDDV
98	DEGAAQGDDDVWTSG
99	AQGDDDVWTSGSDSD

100	DDVWTSGSDSDEELV
101	TSGSDSDEELVTTER
102	DSDEELVTTERKTTPR
103	ELVTTERKTTPRVTTGG
104	TERKTTPRVTTGGGAMA
105	TPRVTTGGGAMAGAST
106	TGGGAMAGASTSAGR
107	AMAGASTSAGRKRKS
108	ASTSAGRKRKSASSA
109	AGRKRKSASSATACT
110	RKSASSATACTSGVM
111	SSATACTSGVMTRGR
112	ACTSGVMTRGRLLKAE
113	GVMTRGRLLKAESTVA
114	RGRLLKAESTVAPEED
115	KAESTVAPEEDTDED
116	TVAPEEDTDESDNE
117	EEDTDESDNEIHNHP
118	DESDNEIHNPAVFT
119	DNEIHNPAVFTWPPW
120	HNPVFTWPPWQAGI
121	VFTWPPWQAGILARN
122	PPWQAGILARNLVP
123	AGILARNLVPVATV
124	ARNLVPVATVQGGQN
125	VPMVATVQGGQNLKYQ
126	ATVQGGQNLKYQEFFW
127	GQNLKYQEFFWDAND
128	KYQEFFWDANDIYRI
129	FFWDANDIYRIFAE
130	ANDIYRIFAELEGVW
131	YRIFAELEGVWQPAA
132	AELEGVWQPAAQPKR
133	GVWQPAAQPKRRRHR
134	PAAQPKRRRHRQOAL
135	PKRRRHRQOALPGPC
136	RHRQOALPGPCIAST
137	OALPGPCIASTPKKOH
138	LPGPCIASTPKKHRRG

B) IE1

Peptide No.	Sequence
139	MESSAKRKMDPDNDP
140	AKRKMDPDNPDEGPS
141	MDPDNPDEGPSSKVP
142	NPDEGPSSKVPRPET
143	GPSSKVPRPETPVTK
144	KVPRPETPVTKATTF
145	PETPVTKATTFLQTM
146	VTKATTFLQTMLRKE
147	TTFLQTMLRKEVNSQ
148	QTMLRKEVNSQLSLG
149	RKEVNSQLSLGDPLF
150	NSQLSLGDPLFPELA
151	SLGDPLFPELAEESL
152	PLFPELAEESLKTFE
153	ELAEESLKTFEQVTE
154	ESLKTFEQVTEDCNE
155	TFEQVTEDCNENPEK
156	VTEDCNENPEKDVLA
157	CNENPEKDVLAELVK
158	PEKDVLAELVKQIKV
159	VLAELVKQIKVRVDM
160	LVKQIKVRVDMVRHR
161	IKVRVDMVRHRIKEH
162	VDMVRHRIKEHMLKK
163	RHRIKEHMLKKYTQT
164	KEHMLKKYTQTEEFK
165	LKKYTQTEEFKTFGAF
166	TQTEEFKTFGAFNMMG
167	EKTFGAFNMMGGCLQ
168	GAFNMMGGCLQNALD
169	MMGGCLQNALDILDK
170	CLQNALDILDKVHEP
171	ALDILDKVHEPFEEM
172	LDKVHEPFEEMKCI
173	HEPFEEMKCIGLTMQ
174	EEMKCIGLTMQSMYE
175	CIGLTMQSMYENYIV
176	TMQSMYENYIVPEDK
177	MYENYIVPEDKREMW
178	YIVPEDKREMWMACI
179	EDKREMWMACIKELH
180	EMWMACIKELHDVSK
181	ACIKELHDVSKGAAN
182	ELHDVSKGAANKLGG
183	VSKGAANKLGGALQA
184	AANKLGGALQAKARA
185	LGGALQAKARAKKDE
186	LQAKARAKKDELRRK
187	ARAKKDELRRKMMYM
188	KDELRRKMMYMCYRN

189	RRKMMYMCYRNIEFF
190	MYMCYRNIEFFTKNS
191	YRNIEFFTKNSAFPK
192	EFFTKNSAFPKTTNG
193	KNSAFPKTTNGCSQA
194	FPKTTNGCSQAMAAL
195	TNGCSQAMAALQNLP
196	SQAMAALQNLPQCSP
197	AALQNLPQCSPDEIM
198	NLPQCSPDEIMAYAQ
199	CSPDEIMAYAQKIFK
200	EIMAYAQKIFKILDE
201	YAQKIFKILDEERDK
202	IFKILDEERDKVLTH
203	LDEERDKVLTHIDHI
204	RDKVLTHIDHIFMDI
205	LTHIDHIFMDILTTC
206	DHIFMDILTTCVETM
207	MDILTTCVETMCNEY
208	TTCVETMCNEYKVTS
209	ETMCNEYKVTSDACM
210	NEYKVTSDACMMTMY
211	VTSDACMMTMYGGIS
212	ACMMTMYGGISLLSE
213	TMYGGISLLSEFCRV
214	GISLLSEFCRVLCCY
215	LSEFCRVLCCYVLEE
216	CRVLCCYVLEETSVM
217	CCYVLEETSVMIAKR
218	LEETSVMIAKRPLIT
219	SVMLAKRPLITKPEV
220	AKRPLITKPEVISVM
221	LITKPEVISVMKRRRI
222	PEVISVMKRRRIEIC
223	SVMKRRRIEICMKVF
224	RRIEICMKVFAQYI
225	EICMKVFAQYILGAD
226	KVFAQYILGADPLRV
227	QYILGADPLRVCS
228	GADPLRVCSVDDL
229	LRVCSVDDLRAIA
230	SPSVDDLRAIAEESD
231	DDLRAIAEESDEEEA
232	AIAEESDEEEAIVAY
233	ESDEEEAIVAYTLAT
234	EEAIVAYTLATAGVS
235	VAYTLATAGVSSSDS
236	LATAGVSSSDSLVSP
237	GVSSSDSLVSPPEP
238	SDSLVSPPEPVPAT

239	VSPPEPVPATIPLS
240	ESPVPATIPLSVIV
241	PATIPLSVIVAENS
242	PLSSVIVAENSQEE
243	VIVAENSQEESEQS
244	ENSQEESEQSDEEE
245	QEESEQSDEEEEGA
246	EQSDEEEEGAQEER
247	EEEEEGAQEEREDTV
248	EGAQEEREDTVSVKS
249	EEREDTVSVKSEPV
250	DTVSVKSEPVSEIEE
251	VKSEPVSEIEEVAPE
252	PVSEIEEVAPEEEED
253	IEEVAPEEEEDGAE
254	APEEEEDGAEPTAS
255	EEDGAEPTASGGKS
256	AEEPTASGGKSTHPM
257	TASGGKSTHPMVRTS
258	GKSTHPMVRTSKADQ

C) IE2

Peptide No.	Sequence				
259	MESSAKRKMDPDNDP	312	VETRGATASSPSTGS	365	DAVRCRLGTMCNLAL
260	AKRKMDPDNPDEGPS	313	GATASSPSTGSGTTPR	366	CRLGTMCNLALSTPF
261	MDPDNPDEGPSSKVP	314	SSPSTGSGTTPRVTSPTSP	367	TMCNLALSTPFLMEH
262	NPDEGPSSKVPVPET	315	TGSGTTPRVTSPTHPL	368	LALSTPFLMEHTMPV
263	GPSSKVPVPETPVTK	316	TPRVTSPTHPLSQMN	369	TPFLMEHTMPVTHPP
264	KVPRPETPVTKATTF	317	TSPTHPLSQMNHPL	370	MEHTMPVTHPPEVAQ
265	PETPVTKATTFLQTM	318	HPLSQMNHPLPDPL	371	MPVTHPPEVAQRTAD
266	VTKATTFLQTMLRKE	319	QMNHPLPDPLGRPD	372	HPPEVAQRTADACNE
267	TTFLQTMLRKEVNSQ	320	PPLPDPLGRPDEDSS	373	VAQRTADACNEGVKA
268	QTMLRKEVNSQLSLG	321	DPLGRPDEDSSSSSS	374	TADACNEGVKAAWSL
269	RKEVNSQLSLGDPLF	322	RPDEDSSSSSSSSCS	375	CNEGVKAAWSLKLH
270	NSQLSLGDPLFPELA	323	DSSSSSSSSSCSSASD	376	VKAAWSLKLHQLH
271	SLGDPLFPELAEESL	324	SSSSSCSSASDSESE	377	WSLKLHQLHQLCPRS
272	PLFPELAEESLKTFE	325	SCSSASDSESESEEM	378	ELHQLHQLCPRSSDYR
273	ELAEESLKTFEQVTE	326	ASDSESESEEMKCSS	379	HQLCPRSSDYRNMI
274	ESLKTFEQVTEDCNE	327	ESESEEMKCSSGGGA	380	PRSSDYRNMIHAAT
275	TFEQVTEDCNENPEK	328	EEMKCSSGGGASVTS	381	DYRNMIHAATPVDL
276	VTEDCNENPEKDVLA	329	CSSGGGASVTSSSHG	382	MIHAATPVDLLGAL
277	CNENPEKDVLAELGD	330	GGASVTSSSHGRGGF	383	AATPVDLLGALNLCL
278	PEKDVLAELGDILAQ	331	VTSSHHGRGGFGGAA	384	VDLLGALNLCLPLMQ
279	VLAELGDILAQAVNH	332	HHGRGGFGGAASSSL	385	GALNLCLPLMQKFPK
280	LGDILAQAVNHAGID	333	GGFGGAASSLLSCG	386	LCLPLMQKFPKQVMV
281	LAQAVNHAGIDSSST	334	GAASSLLSCGHQSS	387	LMQKFPKQVMVRIFS
282	VNHAGIDSSSTGPTL	335	SSLLSCGHQSSGGAS	388	FPKQVMVRIFSTNQG
283	GIDSSSTGPTLTTHS	336	SCGHQSSGGASTGPR	389	VMVRIFSTNQGGMFL
284	SSTGPTLTTHSCSVS	337	QSSGGASTGPRKKKS	390	IFSTNQGGMFLPIYE
285	PTLTTHSCSVSSAPL	338	GASTGPRKKKSKRIS	391	NQGMFLPIYETAAK
286	THSCSVSSAPLNKPT	339	GPRKKKSKRISELDN	392	FMLPIYETAAKAYAV
287	SVSSAPLNKPTPTSV	340	KKSKRISELDNEKVR	393	IYETAAKAYAVGQFE
288	APLNKPTPTSVAVTN	341	RISELDNEKVRNIMK	394	AAKAYAVGQFEQPTE
289	KPTPTSVAVTNTPLP	342	LDNEKVRNIMKDKNT	395	YAVGQFEQPTEPPE
290	TSVAVTNTPLPGASA	343	KVRNIMKDKNTPFCT	396	QFEQPTEPPELDLT
291	VTNTPLPGASATPEL	344	IMKDKNTPFCTPNVQ	397	PTTEPPELDLTLSLA
292	PLPGASATPELSPRK	345	KNTPFCTPNVQTRRG	398	PPEDLDLTLSLAIEAA
293	ASATPELSPRKKPRK	346	FCTPNVQTRRGRVKI	399	LDLTLSLAIEAAIQDL
294	PELSPRKKPRKTTRP	347	NVQTRRGRVKIDEVS	400	SLAIEAAIQDLRNKS
295	PRKKPRKTTRPFKVI	348	RRGRVKIDEVSRMFR	401	LAIEAAIQDLRNKSQ
296	PRKTTRPFKVIKPP	349	VKIDEVSRMFRNTNR		
297	TRPFKVIKPPVPPA	350	EVSFRMFRNTNRSLEY		
298	KVIKPPVPPAPIML	351	MFRNTNRSLEYKNLP		
299	KPPVPPAPIMLPLIK	352	TNRSLEYKNLPFTIP		
300	PPAPIMLPLIKQEDI	353	LEYKNLPFTIPSMHQ		
301	IMLPLIKQEDIKPEP	354	NLPFTIPSMHQVLDE		
302	LIKQEDIKPEPDFTI	355	TIPSMHQVLDEAIKA		
303	EDIKPEPDFTIQYRN	356	MHQVLDEAIKACKTM		
304	PEPDFTIQYRNKIID	357	LDEAIKACKTMQVNN		
305	FTIQYRNKIIDTAGC	358	IKACKTMQVNNKGIQ		
306	YRNKIIDTAGCIVIS	359	KTMQVNNKGIQIYT		
307	IIDTAGCIVISDSEE	360	VNNKGIQIYTRNHE		
308	AGCIVISDSEEEQGE	361	GIQIYTRNHEVKSE		
309	VISDSEEEQGEVET	362	IYTRNHEVKSEVDAV		
310	SEEEQGEVETRGAT	363	NHEVKSEVDAVRCRL		
311	QGEVETRGATASSP	364	KSEVDAVRCRLGTM		

Table A2: Overlapping peptide sequences (Chapter 3)

Bibliography

- Abele R, Tampé R. Function of the transport complex TAP in cellular immune recognition. *Biochimica et Biophysica Acta - Biomembranes* 1999;1461(2):405–419.
- Adkins B. Heterogeneity in the CD4 T Cell Compartment and the Variability of Neonatal Immune Responsiveness. *Current Immunology Reviews* 2007;3(3):151–159.
- Adland E, Klenerman P, Goulder P, Matthews PC. Ongoing burden of disease and mortality from HIV/CMV coinfection in Africa in the antiretroviral therapy era. *Frontiers in Microbiology* 2015;6(1016).
- Adland E, Paioni P, Thobakgale C, Laker L, Mori L, Muenchhoff M, et al. Discordant Impact of HLA on Viral Replicative Capacity and Disease Progression in Pediatric and Adult HIV Infection. *PLoS Pathogens* 2015;11(6).
- Ahmed R, Oldstone MBa, Palese P. Protective immunity and susceptibility to infectious diseases: lessons from the 1918 influenza pandemic. *Nature Immunology* 2007;8(11):1188–93.
- Akbar AN, Fletcher JM. Memory T cell homeostasis and senescence during aging. *Current Opinion in Immunology* 2005;17(5):480–485.
- Akiyama Y, Maruyama K, Mochizuki T, Sasaki K, Takaue Y, Yamaguchi K. Identifi-

- cation of HLA-A24-restricted CTL epitope encoded by the matrix protein pp65 of human cytomegalovirus. *Immunology Letters* 2002;83(1):21–30.
- Ameres S, Mautner J, Schlott F, Neuenhahn M, Busch DH, Plachter B, et al. Presentation of an immunodominant Immediate-Early CD8+ T Cell epitope resists Human Cytomegalovirus immunoevasion. *PLoS Pathogens* 2013;9(5).
- Amponsah-Dacosta E, Lebelo R, Rakgole N, Burnett R, Selabe S, Mphahlele J. Evidence for a change in the epidemiology of hepatitis B virus infection after nearly two decades of universal Hepatitis B vaccination in South Africa. *Journal of Medical Virology* 2014;86:918–924.
- Andersen MH, Schrama D, thor Straten P, Becker JC. Cytotoxic T Cells. *Journal of Investigative Dermatology* 2006;126(1):32–41.
- Annels NE, Callan MFC, Tan L, Rickinson AB. Changing Patterns of Dominant TCR Usage with Maturation of an EBV-Specific Cytotoxic T Cell Response. *The Journal of Immunology* 2000;165(9):4831–4841.
- Antoine P, Orlislagers V, Huygens A, Lecomte S, Liesnard C, Donner C, et al. Functional exhaustion of CD4+ T lymphocytes during primary cytomegalovirus infection. *The Journal of Immunology* 2012;189(5):2665–2672.
- Appay V, Dunbar PR, Callan M, Klenerman P, Gillespie GMA, Papagno L, et al. Memory CD8+ T cells vary in differentiation phenotype in different persistent virus infections. *Nature Medicine* 2002;8(4):379–385.
- Appay V, Van Lier RAW, Sallusto F, Roederer M. Phenotype and function of human T lymphocyte subsets: Consensus and issues. *Cytometry Part A* 2008;73(11):975–983.

- Argaet VP, Schmidt CW, Burrows SR, Silins SL, Kurilla MG, Doolan DL, et al. Dominant selection of an invariant T cell antigen receptor in response to persistent infection by Epstein-Barr virus. *The Journal of experimental medicine* 1994;180:2335–2340.
- Arnon TI, Achdout H, Levi O, Markel G, Saleh N, Katz G, et al. Inhibition of the NKp30 activating receptor by pp65 of human cytomegalovirus. *Nature immunology* 2005;6(5):515–523.
- Arstila TP, Casrouge A, Baron V, Even J, Kanellopoulos J, Kourilsky P. A direct estimate of the human alphabeta T cell receptor diversity. *Science* 1999;286(5441):958–61.
- Attaf M, Huseby E, Sewell AK. $\alpha\beta$ T cell receptors as predictors of health and disease. *Cellular and Molecular Immunology* 2015;12(4):391–399.
- Babel N, Brestrich G, Gondek LP, Sattler A, Wlodarski MW, Poliak N, et al. Clonotype analysis of cytomegalovirus-specific cytotoxic T lymphocytes. *Journal of the American Society of Nephrology* 2009;20(2):344–52.
- Bachmann MF, Wolint P, Walton S, Schwarz K, Oxenius A. Differential role of IL-2R signaling for CD8+ T cell responses in acute and chronic viral infections. *European Journal of Immunology* 2007;37(6):1502–1512.
- Barnes E. Therapeutic vaccines in HBV : lessons from HCV. *Medical Microbiology and Immunology* 2015;204:79–86.
- Barry AP, Silvestri G, Safrit JT, Sumpter B, Kozyr N, McClure HM, et al. Depletion of CD8+ cells in sooty mangabey monkeys naturally infected with simian immunodeficiency virus reveals limited role for immune control of virus replication in a natural host species. *The Journal of Immunology* 2007;178(12):8002–8012.

- Bego M, Maciejewski J, Khaiboullina S, Pari G, Jeor SS. Characterization of an Antisense Transcript Spanning the UL81-82 Locus of Human Cytomegalovirus. *Journal of Virology* 2005;79(17):11022–11034.
- Belz GT, Stevenson PG, Doherty PC. Contemporary analysis of MHC-related immunodominance hierarchies in the CD8+ T cell response to influenza A viruses. *The Journal of Immunology* 2000;165(5):2404–2409.
- van de Berg PJ, Heutinck KM, Raabe R, Minnee RC, Young SL, van Donselaarvan der Pant K, et al. Human Cytomegalovirus Induces Systemic Immune Activation Characterized by a Type 1 Cytokine Signature. *The Journal of Infectious Diseases* 2010;202(5):690–699.
- Bevan MJ. In thymic selection, peptide diversity gives and takes away. *Immunity* 1997;7(2):175–178.
- van Bockel DJ, Price DA, Munier ML, Venturi V, Asher TE, Ladell K, et al. Persistent Survival of Prevalent Clonotypes within an Immunodominant HIV Gag-Specific CD8+ T Cell Response. *The Journal of Immunology* 2011;186(1):359–371.
- von Boehmer H, Teh HS, Kisielow P. The thymus selects the useful, neglects the useless and destroys the harmful. *Immunology Today* 1989;10(2):57–61.
- Bonacini M, Louie S, Bzowej N, Wohl A. Survival in patients with HIV infection and viral hepatitis B or C : a cohort study. *AIDS* 2004;18(15):2039–2045.
- Boni C, Fisticaro P, Valdatta C, Amadei B, Vincenzo PD, Giuberti T, et al. Characterization of hepatitis B virus (HBV)-specific T-cell dysfunction in chronic HBV infection. *Journal of Virology* 2007;81(8):4215–25.
- Boppana SB, Britt WJ. Antiviral Antibody-Responses and Intrauterine Transmis-

- sion After Primary Maternal Cytomegalovirus-Infection. *The Journal of Infectious Diseases* 1995;171(5):1115–1121.
- Bouso P, Wahn V, Douagi I, Horneff G, Pannetier C, Le Deist F, et al. Diversity, functionality, and stability of the T cell repertoire derived in vivo from a single human T cell precursor. *PNAS* 2000;97(1):274–8.
- Boyman O, Létourneau S, Krieg C, Sprent J. Homeostatic proliferation and survival of naïve and memory T cells. *European Journal of Immunology* 2009;39(8):2088–2094.
- Braendstrup P, Mortensen BK, Justesen S, Osterby T, Rasmussen M, Hansen AM, et al. Identification and HLA-tetramer-validation of human CD4+ and CD8+ T cell responses against HCMV proteins IE1 and IE2. *PLoS One* 2014;9(4):e94892.
- Britanova OV, Shugay M, Merzlyak EM, Staroverov DB, Putintseva EV, Turchaninova MA, et al. Dynamics of Individual T Cell Repertoires: From Cord Blood to Centenarians. *The Journal of Immunology* 2016;196(12):5005–5013.
- Britanova O, Putintseva E, Shugay M, Merzlyak E, Turchaninova M, Staroverov D, et al. Age-related decrease in TCR repertoire diversity measured with deep and normalized sequence profiling. *The Journal of Immunology* 2014;192(6):2689–2698.
- Britt WJ, Vugler L, Butfiloski EJ, Stephens EB. Cell surface expression of human cytomegalovirus (HCMV) gp55-116 (gB): use of HCMV-recombinant vaccinia virus-infected cells in analysis of the human neutralizing antibody response. *Journal of Virology* 1990;64(3):1079–1085.
- Broere F, Apasov SG, Sitkovsky MV, Eden WV. T cell subsets and T cell-mediated immunity. *Principles of Immunopharmacology* 2011;p. 15–27.

- Broers AEC, van Der Holt R, van Esser JW, Gratama JW, Henzen-Logmans S, Kuenen-Boumeester V, et al. Increased transplant-related morbidity and mortality in CMV-seropositive patients despite highly effective prevention of CMV disease after allogeneic T-cell-depleted stem cell transplantation. *Blood* 2000;95(7):2240–2245.
- Brown JA, Stevenson K, Kim HT, Cutler C, Ballen K, McDonough S, et al. Clearance of CMV viremia and survival after double umbilical cord blood transplantation in adults depends on reconstitution of thymopoiesis. *Blood* 2010;115(20):4111–4119.
- Brumme Z, Wang B, Nair K, Brumme C, De Pierres C, Reddy S, et al. Impact of select immunologic and virologic biomarkers on CD4 decline in chronic HIV-1 subtype C infection: Results from Sinikithemba longitudinal population-based cohort, Durban, South Africa. *Clinical Infectious Diseases* 2010;49(6):956–964.
- Bukowski JF, Woda BA, Habu S, Okumura K, Welsh RM. Natural killer cell depletion enhances virus synthesis and virus-induced hepatitis in vivo. *The Journal of Immunology* 1983;131(3):1531–8.
- Bukowski J, Warner J, Dennert G, Welsh RM. Adoptive transfer studies demonstrating the antiviral effect of natural killer cells in vivo. *The Journal of Experimental Medicine* 1985;161:40–52.
- Burrows SR, Khanna R, Burrows JM, Moss DJ. An alloresponse in humans is dominated by cytotoxic T lymphocytes (CTL) cross-reactive with a single Epstein-Barr virus CTL epitope: implications for graft-versus-host disease. *The Journal of Experimental Medicine* 1994;179(4):1155–61.
- Busca A, Kumar A. Innate immune responses in hepatitis B virus (HBV) infection. *Virology Journal* 2014;11(22).

- Callan M, Steven N, Krausa P, Wilson J, Moss P, Gillespie G, et al. Large clonal expansions of CD8+ T cells in acute infectious mononucleosis. *Nature Medicine* 1996;2(8):906–911.
- Campos-Lima PO, Levitsky V, Imreh MP, Gavioli R, Masucci MG. Epitope-dependent selection of highly restricted or diverse T cell receptor repertoires in response to persistent infection by Epstein-Barr virus. *The Journal of Experimental Medicine* 1997;186(1):83–9.
- Cannon MJ, Schmid DS, Hyde TB. Review of cytomegalovirus seroprevalence and demographic characteristics associated with infection. *Reviews in Medical Virology* 2010;20(4):202–213.
- Cano P, Klitz W, Mack SJ, Maiers M, Marsh SGE, Noreen H, et al. Common and well-documented HLA Alleles. *Human Immunology* 2007;68(5):392–417.
- Cao K, Hollenbach J, Shi X, Shi W, Chopek M, Fernández-Viña MA. Analysis of the frequencies of HLA-A, B, and C alleles and haplotypes in the five major ethnic groups of the United States reveals high levels of diversity in these loci and contrasting distribution patterns in these populations. *Human Immunology* 2001;62(9):1009–1030.
- Carlson JM, Listgarten J, Pfeifer N, Tan V, Kadie C, Walker BD, et al. Widespread impact of HLA restriction on immune control and escape pathways of HIV-1. *Journal of Virology* 2012;86(9):5230–5243.
- Castro-Malaspina H, Harris RE, Gajewski J, Ramsay N, Collins R, Dharan B, et al. Unrelated donor marrow transplantation for myelodysplastic syndromes: Outcome analysis in 510 transplants facilitated by the National Marrow Donor Program. *Blood* 2002;99(6):1943–1951.

- Chang MI, Liaw Yf. Hepatitis B flares in chronic hepatitis B : pathogenesis , natural course , and management. *Journal of Hepatology* 2014;61(6):1407–1417.
- Charini WA, Kuroda MJ, Schmitz JE, Beaudry KR, Lin W, Lifton Ma, et al. Clonally diverse CTL response to a dominant viral epitope recognizes potential epitope variants. *The Journal of immunology* 2001;167(9):4996–5003.
- Chen DH, Jiang H, Lee M, Liu F, Zhou ZH. Three-Dimensional Visualization of Tegument/Capsid Interactions in the Intact Human Cytomegalovirus. *Virology* 1999;260(1):10–16.
- Chen G, Yang X, Ko A, Sun X, Gao M, Zhang Y, et al. Sequence and Structural Analyses Reveal Distinct and Highly Diverse Human CD8+ TCR Repertoires to Immunodominant Viral Antigens. *Cell Reports* 2017;19(3):569–583.
- Chen My, Hung Cc, Lee Kl. Detection of human parvovirus 4 viremia in the follow-up blood samples from seropositive individuals suggests the existence of persistent viral replication or reactivation of latent viral infection. *Virology Journal* 2015;p. 4–8.
- Chen SF, Tu WW, Sharp MA, Tongson EC, He XS, Greenberg HB, et al. Antiviral CD8 T cells in the control of primary human cytomegalovirus infection in early childhood. *The Journal of Infectious Diseases* 2004;189(9):1619–1627.
- Cheung AKL, Abendroth A, Cunningham AL, Slobedman B. Viral gene expression during the establishment of human cytomegalovirus latent infection in myeloid progenitor cells. *Blood* 2006;108(12):3691–3699.
- Chevalier MS, Johnson DC. Human cytomegalovirus US3 chimeras containing US2 cytosolic residues acquire major histocompatibility class I and II protein degradation properties. *Journal of Virology* 2003;77(8):4731–4738.

- Chisari FV, Isogawa M, Wieland SF. Pathogenesis of hepatitis B virus infection. *Pathologie-biologie* 2010;58(4):258–66.
- Chothia C, Boswell DR, Lesk AM. The outline structure of the T-cell alpha beta receptor. *The EMBO Journal* 1988;7(12):3745–55.
- Chun HM, Roediger MP, Hullsiek KH, Thio CL, Agan BK, Bradley WP, et al. Hepatitis B virus coinfection negatively impacts HIV outcomes in HIV seroconverters. *The Journal of Infectious Disease* 2012;205:185–193.
- Compston L, Li C, Sarkodie F, Owusu-Ofori S, Opare-Sem O, Allain J. Prevalence of persistent and latent viruses in untreated patients infected with HIV-1 from Ghana, West Africa. *Journal of Medical Virology* 2009;81(11):1860–8.
- Compton T. Receptors and immune sensors: the complex entry path of human cytomegalovirus. *Trends in Cell Biology* 2004;14(1):5–8.
- Corbett NP, Blimkie D, Ho KC, Cai B, Sutherland DP, Kallos A, et al. Ontogeny of toll-like receptor mediated cytokine responses of human blood mononuclear cells. *PLoS ONE* 2010;5(11).
- Corey L, Wald A. Maternal and neonatal HSV infections. *New England Journal of Medicine* 2009;361(14):1376–1385.
- Costa AI, Koning D, Ladell K, McLaren JE, Grady BPX, Schellens IMM, et al. Complex T-Cell Receptor Repertoire Dynamics Underlie the CD8 + T-Cell Response to HIV-1. *Journal of Virology* 2015;89(1):110–119.
- Crough T, Khanna R. Immunobiology of human cytomegalovirus: from bench to bedside. *Clinical Microbiology Reviews* 2009;22(1):76–98.
- Das A, Hoare M, Davies N, Lopes AR, Dunn C, Kennedy PTF, et al. Functional

- skewing of the global CD8 T cell population in chronic hepatitis B virus infection. *The Journal of Experimental Medicine* 2008;205(9):2111–2124.
- David D, Levy O. Ontogeny of early life immunity. *Trends in Immunology* 2014;35(7):299–310.
- Davis MM, Bjorkman PJ. T-cell antigen receptor genes and T-cell recognition. *Nature* 1988;334(6181):395–402.
- Day E, Carmichael A, ten Berge I, Waller E, Sissons J, Wills M. Rapid CD8+ T Cell Repertoire Focusing and Selection of High-Affinity Clones into Memory Following Primary Infection with a Persistent Human Virus: Human Cytomegalovirus. *The Journal of Immunology* 2007;179(5):3203–3213.
- Deayton JR, Sabin CA, Johnson MA, Emery VC, Wilson P, Griffiths PD. Importance of cytomegalovirus viraemia in risk of disease progression and death in HIV-infected patients receiving highly active antiretroviral therapy. *The Lancet* 2004;363(9427):2116–2121.
- Dechanet J, Merville P, Berge F, Bone-Mane G, Taupin JL, Michel P, et al. Major expansion of $\gamma\delta$ T lymphocytes following cytomegalovirus infection in kidney allograft recipients. *The Journal of Infectious Diseases* 1999;179(1):1–8.
- Déchanet J, Merville P, Lim A, Retière C, Pitard V, Lafarge X, et al. Implication of $\gamma\delta$ T cells in the human immune response to cytomegalovirus. *The Journal of Clinical Investigation* 1999;103(10):1437–1449.
- Desiderio SV, Yancopoulos GD, Paskind M, Thomas E, Boss MA, Landau N, et al. Insertion of N regions into heavy-chain genes is correlated with expression of terminal deoxytransferase in B cells. *Nature* 1984;311(5988):752–755.

- Detels R, Leach C, Hennessey K, Liu Z, Visscher B, Cherry J, et al. Persistent cytomegalovirus infection of semen increases risk of AIDS. *The Journal of Infectious Diseases* 1994;169(4):766–768.
- Doherty DG, Norris S, Madrigal-Estebas L, Mcentee G, Traynor O, Hegarty JE, et al. The human liver contains multiple populations of NK cells, T cells, and CD3 + CD56+ natural T cells with distinct cytotoxic activities and Th1, Th2, and Th0 cytokine secretion patterns. *The Journal of Immunology* 1999;163(4):2314–21.
- Doherty PC, Zinkernagel RM. A biological role for the major histocompatibility antigens. *The Lancet* 1975;1(7922):1406–9.
- Dolan A, Cunningham C, Hector RD, Hassan-Walker AF, Lee L, Addison C, et al. Genetic content of wild-type human cytomegalovirus. *Journal of General Virology* 2004;85(5):1301–1312.
- Dollard SC, Staras SAS, Amin MM, Schmid DS, Cannon MJ. National prevalence estimates for cytomegalovirus IgM and IgG avidity and association between high IgM antibody titer and low IgG avidity. *Clinical and Vaccine Immunology* 2011;18(11):1895–1899.
- Donald PR, Marais BJ, Barry CE. Age and the epidemiology and pathogenesis of tuberculosis. *The Lancet* 2010;375(9729):1852–1854.
- Dong T, Stewart-Jones G, Chen N, Easterbrook P, Xu X, Papagno L, et al. HIV-specific cytotoxic T cells from long-term survivors select a unique T cell receptor. *The Journal of experimental medicine* 2004;200(12):1547–57.
- Drexler JF, Reber U, Muth D, Herzog P, Annan A, Ebach F, et al. Human parvovirus 4 in nasal and fecal specimens from children, Ghana. *Emerging Infectious Diseases* 2012;18(10):1650–1653.

- Dyall R, Messaoudi I, Janetzki S, Nikolic-Zugic J. MHC Polymorphism Can Enrich the T Cell Repertoire of the Species by Shifts in Intrathymic Selection. *The Journal of Immunology* 2000;164(4):1695–1698.
- Einsele H, Roosnek E, Rufer N, Sinzger C, Riegler S, Löffler J, et al. Infusion of cytomegalovirus (CMV) specific T cells for the treatment of CMV infection not responding to antiviral chemotherapy. *Blood* 2002;99(11):3916–3922.
- Elkington R, Walker S, Crough T, Menzies M, Tellam J, Bharadwaj M, et al. Ex vivo profiling of CD8⁺-T-cell responses to human cytomegalovirus reveals broad and multispecific reactivities in healthy virus carriers. *Journal of Virology* 2003;77(9):5226–40.
- Enders J, Weller T, Robbins F. Cultivation of the Lansing strain of poliomyelitis virus in cultures of various human embryonic tissues. *Science* 1949;109:85–87.
- Erice A, Tierney C, Hirsch M, Caliendo AM, Weinberg A, Kendall MA, et al. Cytomegalovirus (CMV) and human immunodeficiency virus (HIV) burden, CMV end-organ disease, and survival in subjects with advanced HIV infection (AIDS Clinical Trials Group Protocol 360). *Clinical Infectious Diseases* 2003;37(4):567–78.
- 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.
- Feau S, Arens R, Togher S, Schoenberger SP. Autocrine IL-2 is required for secondary population expansion of CD8⁽⁺⁾ memory T cells. *Nature Immunology* 2011;12(9):908–913.
- Ferrari C, Penna A, Bertoletti A, Valli A, Antoni AD, Cavalli A, et al. Cellular immune

- response to hepatitis B virus-encoded antigens in acute and chronic hepatitis B virus infection. *The Journal of Immunology* 1990;145:342–3449.
- Fielding K, Koba A, Grant AD, Charalambous S, Day J, Spak C, et al. Cytomegalovirus viremia as a risk factor for mortality prior to antiretroviral therapy among HIV-infected gold miners in South Africa. *PLoS ONE* 2011;6(10):e25571.
- Fietze E, Prosch S, Reinke P, Stein J, Docke WD, Staffa G, et al. Cytomegalovirus infection in transplant recipients. *Transplantation* 1994;58(6):675–680.
- Fowler KB, Boppana SB. Congenital cytomegalovirus (CMV) infection and hearing deficit. *Journal of Clinical Virology* 2006;35(2):226–231.
- Fryer F, Lucas S, Paldley D, Sally B. Parvoviruses PARV4/5 in Hepatitis C Virus-infected Persons. *New England Journal of Medicine* 2007;356(1):175–176.
- Fugmann SD, Lee AI, Shockett PE, Villey IJ, Schatz DG. The RAG proteins and V(D)J recombination: complexes, ends, and transposition. *Annual Review of Immunology* 2000;18:495–527.
- Furman D, Jojic V, Sharma S, Shen-Orr S, Angel CJL, Onengut-Gumuscu S, et al. Cytomegalovirus infection improves immune responses to influenza. *Science Translational Medicine* 2015;7(281):1–22.
- Gamadia LE, Remmerswaal EBM, Weel JF, Bemelman F, van Lier R, ten Berge I. Primary immune responses to human CMV : a critical role for IFN- γ -producing CD4⁺ T cells in protection against CMV disease. *Blood* 2003;101(7):2686–2692.
- Garderet L, Dulphy N, Douay C, Chalumeau N, Schaeffer V, Zilber MT, et al. The umbilical cord blood alphabeta T-cell repertoire: characteristics of a polyclonal and naive but completely formed repertoire. *Blood* 1998;91(1):340–346.

- Gavin MA, Gilbert MJ, Riddell SR, Greenberg PD, Bevan MJ. Alkali hydrolysis of recombinant proteins allows for the rapid identification of class I MHC-restricted CTL epitopes. *The Journal of immunology* 1993;151(8):3971–3980.
- Gellert M. V(D)J Recombination: RAG Proteins, Repair Factors, and Regulation. *Annual Review of Biochemistry* 2002;71(1):101–132.
- Germain RN. T-cell development and the CD4CD8 lineage decision. *Nature Reviews Immunology* 2002;2(5):309–322.
- Gerna G, Percivalle E, Baldanti F, Sarasini A, Zavattoni M, Furione M, et al. Diagnostic significance and clinical impact of quantitative assays for diagnosis of human cytomegalovirus infection/disease in immunocompromised patients. *New Microbiologica* 1998;21(3):293–308.
- Gerna G, Sarasini A, Patrone M, Percivalle E, Fiorina L, Campanini G, et al. Human cytomegalovirus serum neutralizing antibodies block virus infection of endothelial/epithelial cells, but not fibroblasts, early during primary infection. *Journal of General Virology* 2008;89(4):853–865.
- Gianella S, Scheffler K, Mehta SR, Little SJ, Freitas L, Morris SR, et al. Seminal shedding of CMV and HIV transmission among men who have sex with men. *International Journal of Environmental Research and Public Health* 2015;12(7):7585–7592.
- Gilbert MJ, Riddell SR, Plachter B, Greenberg PD. Cytomegalovirus selectively blocks antigen processing and presentation of its immediateearly gene product. *Nature* 1996;383(6602):720–722.
- Gilfillan S, Benoist C, Mathis D. Mice Lacking Terminal Deoxynucleotidyl Transferase: Adult Mice with a Fetal Antigen Receptor Repertoire. *Immunological Reviews* 1995;148(1):201–219.

- Gillespie GMA, Wills MR, Appay V, O'Callaghan C, Murphy M, Smith N, et al. Functional Heterogeneity and High Frequencies of Cytomegalovirus-Specific CD8⁺ T Lymphocytes in Healthy Seropositive Donors. *Journal of Virology* 2000;74(17):8140–8150.
- Giorgi J, Detels R. T-cell subset alterations in HIV-infected homosexual men: NIAID Multicenter AIDS cohort. *Clinical Immunology and Immunopathology* 1989;52(1):10–8.
- Godfrey D, Kennedy J, Suda T, Zlotnik A. A developmental pathway involving four phenotypically and functionally distinct subsets of CD3-CD4-CD8- triple-negative adult mouse thymocytes defined by CD44 and CD25 expression. *The Journal of Immunology* 1993;150:4244–4252.
- Goldmacher VS, Bartle LM, Skaletskaya A, Dionne CA, Kedersha NL, Vater CA, et al. A cytomegalovirus-encoded mitochondria-localized inhibitor of apoptosis structurally unrelated to Bcl-2. *PNAS* 1999;96(22):12536–41.
- Goldrath AW, Bevan MJ. Selecting and maintaining a diverse T-cell repertoire. *Nature* 1999;402(6759):255–262.
- Goodrum F, Reeves MB, Sinclair J, High K, Shenk T. Human cytomegalovirus sequences expressed in latently infected individuals promote a latent infection in vitro. *Immunobiology* 2011;110(3):937–945.
- Goulder P, Walker B. HIV and HLA Class I: an evolving relationship. *Immunity* 2012;37(3):426–440.
- Griffiths PD. CMV as a cofactor enhancing progression of AIDS. *Journal of Clinical Virology* 2006;35(4):489–492.

- Hadrup SR, Strindhall J, Køllgaard T, Seremet T, Johansson B, Pawelec G, et al. Longitudinal studies of clonally expanded CD8 T cells reveal a repertoire shrinkage predicting mortality and an increased number of dysfunctional cytomegalovirus-specific T cells in the very elderly. *The Journal of Immunology* 2006;176(4):2645–2653.
- Hansen SG, Vieville C, Whizin N, Coyne-johnson L, Siess DJ, Drummond DD, et al. Effector-memory T cell responses are associated with protection of rhesus monkeys from mucosal SIV challenge. *Nature Medicine* 2009;15(3):293–299.
- Hawkins C, Christian B, Ye J, Nagu T, Aris E, Chalamilla G, et al. Prevalence of hepatitis B co-infection and response to antiretroviral therapy among HIV-infected patients in Tanzania. *AIDS* 2013;27(6):919–927.
- Heath WR, Carbone FR. Cross-presentation in viral immunity and self-tolerance. *Nature Reviews Immunology* 2001;1(2):126–134.
- Hegde NR, Tomazin RA, Wisner TW, Dunn C, Boname JM, Lewinsohn DM, et al. Inhibition of HLA-DR Assembly, Transport, and Loading by Human Cytomegalovirus Glycoprotein US3: a Novel Mechanism for Evading Major Histocompatibility Complex Class II Antigen Presentation. *Journal of Virology* 2002;76(21):10929–10941.
- Hertoghs KML, Moerland PD, Stijn AV, Remmerswaal EBM, Yong SL, Berg PJE-JVD, et al. Molecular profiling of cytomegalovirus- induced human CD8 + T cell differentiation. *Journal of Clinical Investigation* 2010;120(11):4077–4090.
- Hiom K, Gellert M. Assembly of a 12/23 Paired Signal Complex: A Critical Control Point in V(D)J Recombination. *Molecular Cell* 1998;1(7):1011–1019.
- Hozumi N, Tonegawa S. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *PNAS* 1976;73(10):3628–3632.

- Huang CY, Sleckman BP, Kanagawa O. Revision of T cell receptor {alpha} chain genes is required for normal T lymphocyte development. PNAS 2005;102(40):14356–61.
- Humphreys IR, Loewendorf A, De Trez C, Schneider K, Benedict CA, Munks MW, et al. OX40 Costimulation Promotes Persistence of Cytomegalovirus-Specific CD8 T Cells: A CD4-Dependent Mechanism. The Journal of Immunology 2007;179:2195–2202.
- Huygens A, Lecomte S, Tackoen M, Olislagers V, Delmarcelle Y, Burny W, et al. Functional Exhaustion Limits CD4+ and CD8+ T-Cell Responses to Congenital Cytomegalovirus Infection. The Journal of Infectious Diseases 2015;212(3):484–494.
- Huygens A, Dauby N, Vermijlen D, Marchant A. Immunity to cytomegalovirus in early life. Frontiers in Immunology 2014;5.
- Iancu EM, Corthesy P, Baumgaertner P, Devevre E, Voelter V, Romero P, et al. Clonotype Selection and Composition of Human CD8 T Cells Specific for Persistent Herpes Viruses Varies with Differentiation but Is Stable Over Time. The Journal of Immunology 2009;183(1):319–331.
- Jacobson MA, Maecker HT, Orr PL, D'Amico R, Natta MV, Li XD, et al. Results of a Cytomegalovirus (CMV) Specific CD8+/Interferon-gamma Cytokine Flow Cytometry Assay Correlate with Clinical Evidence of Protective Immunity in Patients with AIDS with CMV Retinitis. The Journal of Infectious Diseases 2004;189:1362–1373.
- Jenkins C, Garcia W, Godwin MJ, Spencer JV, Stern JL, Abendroth A, et al. Immunomodulatory properties of a viral homolog of human interleukin-10 ex-

- pressed by human cytomegalovirus during the latent phase of infection. *Journal of Virology* 2008;82(7):3736–50.
- Jesionek A, Kiolemenoglou B. Ueber einen Befund von protozoenartigen Gebilden in den Organen eines hereditär-luetischen Foetus. *Muenchner Med Wochenschr* 1904;51(19051907).
- Jones EY. MHC class I and class II structures. *Current Opinion in Immunology* 1997;9(1):75–79.
- Jones LL, Colf LA, Stone JD, Garcia KC, Kranz DM. Distinct CDR3 conformations in TCRs determine the level of cross-reactivity for diverse antigens, but not the docking orientation. *The Journal of Immunology* 2008;181(9):6255–6264.
- Jones MS, Kapoor A, Lukashov VV, Simmonds P, Hecht F, Delwart E. New DNA viruses identified in patients with acute viral infection syndrome. *Journal of Virology* 2005;79(13):8230–6.
- Jonjić S, Pavić I, Polić B, Crnković I, Lucin P, Koszinowski UH. Antibodies are not essential for the resolution of primary cytomegalovirus infection but limit dissemination of recurrent virus. *The Journal of Experimental Medicine* 1994;179(5):1713–7.
- Jung M, Spengler U, Schraut W, Hoffmann R, Zachoval R, Eisenburg J, et al. Hepatitis B virus antigen-specific T-cell activation in patients with acute and chronic hepatitis B. *Journal of Hepatology* 1991;13(3):310–7.
- Kaech SM, Wherry EJ, Ahmed R. Effector and Memory T-Cell Differentiation: Implications for Vaccine Development. *Nature Reviews Immunology* 2002;2(4):251–262.

- Kalejta RF. Tegument Proteins of Human Cytomegalovirus. *Microbiology and Molecular Biology Reviews* 2008;72(2):249–265.
- Kappler J, Kubo R, Haskins K, Hannum C, Marrack P, Pigeon M, et al. The major histocompatibility complex-restricted antigen receptor on T cells in mouse and man: Identification of constant and variable peptides. *Cell* 1983;35(1):295–302.
- Karrer U, Sierro S, Wagner M, Oxenius A, Hengel H, Koszinowski UH, et al. Memory inflation: continuous accumulation of antiviral CD8⁺ T cells over time. *The Journal of Immunology* 2003;170(4):2022–2029.
- Karrer U, Wagner M, Sierro S, Oxenius A, Hengel H, Dumrese T, et al. Expansion of Protective CD8⁺ T-Cell Responses Driven by Recombinant Cytomegaloviruses. *Journal of Virology* 2004;78(5):2255–2264.
- Kern F, Surel IP, Faulhaber N, Frömmel C, Schneider-Mergener J, Schönemann C, et al. Target structures of the CD8(+)-T-cell response to human cytomegalovirus: the 72-kilodalton major immediate-early protein revisited. *Journal of Virology* 1999;73(10):8179–84.
- Khan N, Hislop A, Gudgeon N, Cobbold M, Khanna R, Nayak L, et al. Herpesvirus-specific CD8 T cell immunity in old age: Cytomegalovirus impairs the response to a coresident EBV infection. *The Journal of Immunology* 2004;173(12):7481–7489.
- Khan N, Shariff N, Cobbold M, Bruton R, Ainsworth J, Sinclair A, et al. Cytomegalovirus Seropositivity Drives the CD8 T Cell Repertoire Toward Greater Clonality in Healthy Elderly Individuals. *The Journal of Immunology* 2002;169(4):1984–1992.
- Khan N, Best D, Bruton R, Nayak L, Rickinson AB, Moss PA. T cell recognition

- patterns of immunodominant cytomegalovirus antigens in primary and persistent infection. *The Journal of Immunology* 2007;178(7):4455–4465.
- Kiepiela P, Leslie AJ, Honeyborne I, Ramduth D, Thobakgale C, Chetty S, et al. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. *Nature* 2004;432(7018):769–775.
- Kiepiela P, Ngumbela K, Thobakgale C, Ramduth D, Honeyborne I, Moodley E, et al. CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. *Nature Medicine* 2007;13(1):46–53.
- Kim J, Kim AR, Shin EC. Cytomegalovirus Infection and Memory T Cell Inflation. *Immune Network* 2015;15(4):186–90.
- Kisselev AF, Akopian TN, Woo KM, Goldberg AL. The Sizes of Peptides Generated from Protein by Mammalian 26 and 20S Proteasomes. *The Journal of Biological Chemistry* 1999;274(6):3363–3371.
- Klarenbeek PL, Remmerswaal EBM, ten Berge IJM, Doorenspleet ME, van Schaik BDC, Esveldt REE, et al. Deep Sequencing of Antiviral T-Cell Responses to HCMV and EBV in Humans Reveals a Stable Repertoire That Is Maintained for Many Years. *PLoS Pathogens* 2012;8(9):e1002889.
- Klenerman P, Hill A. T cells and viral persistence: lessons from diverse infections. *Nature immunology* 2005;6(9):873–879.
- Klenerman P, Oxenius A. T cell responses to cytomegalovirus. *Nature Reviews Immunology* 2016;16(6):367–377.
- Kline JN, Hunninghake GM, He B, Monick MM, Hunninghake GW. Synergistic Activation of The Human Cytomegalovirus Major Immediate Early Promoter by

- Prostaglandin E 2 and Cytokines. *Experimental Lung Research* 1998;24(1):3–14.
- Kloverpris H, McGregor R, McLaren J, Ladell K, Harndahl M, Stryhn A, et al. CD8+ T-cell receptor bias and immunodominance in HIV-1 infection. *The Journal of Immunology* 2015;194(11):5329–5345.
- Knipe DM, Howley PM. *Fields Virology*. Lippincott Williams and Wilkins; 2007.
- Kollmann TR, Crabtree J, Rein-Weston A, Blimkie D, Thommai F, Wang XY, et al. Neonatal Innate TLR-Mediated Responses Are Distinct from Those of Adults. *The Journal of Immunology* 2009;183(11):7150–7160.
- Kollmann TR, Levy O, Montgomery R, Goriely S. Innate immune sensing by Toll-like receptors in newborns and the elderly. *Immunity* 2012;37(5):771–783.
- Komatsu H, Inui A, Sogo T, Fujisawa T, Nagasaka H, Nonoyama S, et al. Large scale analysis of pediatric antiviral CD8+ T cell populations reveals sustained, functional and mature responses. *Immunity & Ageing* 2006;3(11).
- Kondo E, Akatsuka Y, Kuzushima K, Tsujimura K, Asakura S, Tajima K, et al. Identification of novel CTL epitopes of CMV-pp65 presented by a variety of HLA alleles. *Blood* 2004;103(2):630–638.
- Kotenko SV, Sacconi S, Izotova LS, Mirochnitchenko OV, Pestka S. Human cytomegalovirus harbors its own unique IL-10 homolog (cmvIL-10). *PNAS* 2000;97(4):1695–1700.
- Kotturi MF, Scott I, Wolfe T, Peters B, Sidney J, Cheroutre H, et al. Naive precursor frequencies and MHC binding rather than the degree of epitope diversity shape CD8+ T cell immunodominance. *The Journal of Immunology* 2008;181(3):2124–33.

- Kovacs A, Schluchter M, Easley K, Demmler G, Shearer W, Russa PL, et al. Cytomegalovirus Infection and HIV-1 Disease Progression in Infants Born To HIV-1 Infected Women Cytomegalovirus Infection and HIV-1 Disease Progression. *The New England Journal of Medicine* 1999;341(2):77–84.
- Krauss JL, Potempa J, Lambris JD, Hajishengallis G. Complementary Tolls in the periodontium: How periodontal bacteria modify complement and Toll-like receptor responses to prevail in the host. *Periodontology* 2000 2010;52(1):141–162.
- Kutza A, Muhl E, Hackstein H, Kirchner H, Bein G. High Incidence of Active Cytomegalovirus Infection Among Septic Patients. *Clinical Infectious Diseases* 1998;26(5):1076–1082.
- La Gruta NL, Kedzierska K, Pang K, Webby R, Davenport M, Chen W, et al. A virus-specific CD8+ T cell immunodominance hierarchy determined by antigen dose and precursor frequencies. *PNAS* 2006;103(4):1–6.
- La Gruta NL, Rothwell WT, Cukalac T, Swan NG, Valkenburg SA, Kedzierska K, et al. Primary CTL response magnitude in mice is determined by the extent of naive T cell recruitment and subsequent clonal expansion. *The Journal of Clinical Investigation* 2010;120(6):1885–1894.
- Ladell K, Hashimoto M, Iglesias MC, Wilmann PG, McLaren JE, Gras S, et al. A Molecular Basis for the Control of Preimmune Escape Variants by HIV-Specific CD8+ T Cells. *Immunity* 2013;38(3):425–436.
- Ladep NG, Agaba PA, Agbaji O, Muazu A, Ugoagwu P, Imade G, et al. Rates and impact of hepatitis on human immunodeficiency virus infection in a large African cohort. *World Journal of Gastroenterology* 2013;19(10):1602–1610.
- Larsen EM, Klopperpris H, Stryhn A, Koofhethile CK, Sims S, Ndung'U T, et al.

- HLArestrictor-a tool for patient-specific predictions of HLA restriction elements and optimal epitopes within peptides. *Immunogenetics* 2011;63:43–55.
- Lavoie M, Sharp CP, Pépin J, Pennington C, Foupouapouognigni Y, Pybus OG, et al. Human Parvovirus 4 Infection, Cameroon. *Emerging Infectious Diseases* 2012;18(4):680–683.
- Lefranc MP, Pommié C, Ruiz M, Giuducelli V, Foulquier E, Truong L, et al. IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Developmental and Comparative Immunology* 2003;27(1):55–77.
- Lehner PJ, Wang ECY, Moss PAH, Williams S, Platt K, Friedman SM, et al. Human HLA-A0201-restricted Cytotoxic T Lymphocyte Recognition of Influenza A Is Dominated by T Cells Bearing the $V\beta 17$ Gene Segment. *The Journal of Experimental Medicine* 1995;181(1):79–91.
- Leisner C, Loeth N, Lamberth K, Justesen S, Sylvester-Hvid C, Schmidt EG, et al. One-pot, mix-and-read peptide-MHC tetramers. *PLoS ONE* 2008;3(2):e1678.
- Leslie A, Price DA, Mkhize P, Bishop K, Rathod A, Day C, et al. Differential selection pressure exerted on HIV by CTL targeting identical epitopes but restricted by distinct HLA alleles from the same HLA supertype. *The Journal of Immunology* 2006;177(7):4699–4708.
- Levitsky V, de Campos-Lima PO, Frisan T, Masucci MG. The clonal composition of a peptide-specific oligoclonal CTL repertoire selected in response to persistent EBV infection is stable over time. *The Journal of Immunology* 1998;161(2):594–601.
- Levy O. Innate immunity of the newborn: basic mechanisms and clinical correlates. *Nature Reviews Immunology* 2007;7(5):379–390.

- Li CR, Greenberg PD, Gilbert MJ, Goodrich JM, Riddell SR. Recovery of HLA-restricted cytomegalovirus (CMV)-specific T-cell responses after allogeneic bone marrow transplant: correlation with CMV disease and effect of ganciclovir prophylaxis. *Blood* 1994;83(7):1971–9.
- Li H, Ye C, Ji G, Han J. Determinants of public T cell responses. *Cell Research* 2012;22(1):33–42.
- Ljunggren HG, Karre K. In search of the missing self:MHC molecules and NK cell recognition. *Immunology Today* 1990;11(7):237–244.
- Loewendorf AI, Arens R, Purton JF, Surh CD, Benedict CA. Dissecting the Requirements for Maintenance of the CMV-Specific Memory T-Cell Pool. *Viral Immunology* 2011;24(4):351–355.
- Longmate J, York J, Rosa CL, Krishnan R, Zhang M, Senitzer D, et al. Population coverage by HLA class-I restricted cytotoxic T-lymphocyte epitopes. *Immunogenetics* 2001;52(3-4):165–173.
- Looney RJ, Falsey A, Campbell D, Torres A, Kolassa J, Brower C, et al. Role of Cytomegalovirus in the T Cell Changes Seen in Elderly Individuals. *Clinical Immunology* 1999;90(2):213–219.
- Lou S, Xu B, Huang Q, Zhi N, Cheng F, Wong S, et al. Molecular characterization of the newly identified human Parvovirus 4 in the family Parvoviridae. *Virology* 2012;422(1):59–69.
- Lumley S, Noble H, Hadley MJ, Callow L, Malik A, Chua YY, et al. Hepitopes : A live interactive database of HLA class I epitopes in hepatitis B virus. *Wellcome Open research* 2016;1(9).

- Lurcharchaiwong W, Chieochansin T, Payungporn S, Theamboonlers A, Poovorawan Y. Parvovirus 4 (PARV4) in serum of intravenous drug users and blood donors. *Infection* 2008;4(5):488–491.
- Macdonald WA, Purcell AW, Mifsud NA, Ely LK, Williams DS, Chang L, et al. A naturally selected dimorphism within the HLA-B44 supertype alters Class I structure, peptide repertoire, and T cell recognition. *The Journal of Experimental Medicine* 2003;198(5):679–691.
- Maini MK, Boni C, Lee CK, Larrubia JR, Reignat S, Ogg GS, et al. The role of virus-specific CD8+ cells in liver damage and viral control during persistent hepatitis B virus infection. *The Journal of Experimental Medicine* 2000;191(8):1269–1280.
- Maini MK, Boni C, Ogg GS, King AS, Reignat S, Lee CK, et al. Direct ex vivo analysis of Hepatitis B virusspecific CD8+ T cells associated with the control of infection. *Gastroenterology* 1999;117(6):1386–1396.
- Malu S, Malshetty V, Francis D, Cortes P. Role of non-homologous end joining in V(D)J recombination. *Immunologic Research* 2012;54:233–246.
- Manning A, Willey SJ, Bell JE, Simmonds P. Comparison of tissue distribution, persistence, and molecular epidemiology of parvovirus B19 and novel human parvoviruses PARV4 and human bocavirus. *The Journal of infectious diseases* 2007;195(9):1345–1352.
- Marc J, Moon J. The role of naïve T cell precursor frequency and recruitment in dictating immune response magnitude. *The Journal of Immunology* 2012;188(9):4135–4140.
- Marchant A, Appay V, van der Sande M, Dulphy N, Liesnard C, Kidd M, et al. Mature CD8(+) T lymphocyte response to viral infection during fetal life. *The journal of Clinical Investigation* 2003;111(11):1747–1755.

- Marchant A, Goldman M. T cell-mediated immune responses in human newborns: Ready to learn? *Clinical and Experimental Immunology* 2005;141(1):10–18.
- Marchant A, Kollmann TR. Understanding the ontogeny of the immune system to promote immune-mediated health for life. *Frontiers in Immunology* 2015;6(77).
- Market E, Papavasiliou FN. V(D)J recombination and the evolution of the adaptive immune system. *PLoS Biology* 2003;1(1):24–27.
- Marsh SGE, Parham P, Barber LD. *The HLA Facts Book*. Elsevier; 1999.
- Marzo AL, Klonowski KD, Bon AL, Borrow P, Tough F, Lefrançois L. Initial T cell frequency dictates memory CD8+ T cell lineage commitment. *Nature Immunology* 2005;6(8):793–799.
- Mason D. A very high level of crossreactivity is an essential feature of the T- cell receptor. *Immunology Today* 1998;19(9):395–404.
- Matthews PC, Adland E, Listgarten J, Leslie AJ, Mkhwanazi N, Carlson JM, et al. HLA-A*7401 mediated control of HIV viremia is independent of its linkage disequilibrium with HLA-B*5703. *The Journal of Immunology* 2011;186(10):211–220.
- Matthews PC, Beloukas A, Malik A, Carlson JM, Jooste P, Ogwu A, et al. Prevalence and characteristics of hepatitis B virus (HBV) coinfection among HIV-Positive women in South Africa and Botswana. *PLoS ONE* 2015;10(7):1–11.
- Matthews PC, Carlson JM, Beloukas A, Malik A, Jooste P, Ogwu A, et al. HLA-A is a Predictor of Hepatitis B e Antigen Status in HIV-Positive African Adults. *Journal of Infectious Diseases* 2016;213(8):1248–1252.
- Matthews PC, Geretti AM, Goulder PJR, Klenerman P. Epidemiology and impact of HIV coinfection with Hepatitis B and Hepatitis C viruses in Sub-Saharan Africa. *Journal of Clinical Virology* 2014;61(1):20–33.

- Matthews PC, Listgarten J, Carlson JM, Payne R, Huang KHG, Frater J, et al. Co-operative additive effects between HLA alleles in control of HIV-1. *PLoS ONE* 2012;7(10).
- Matthews PC, Malik A, Simmons R, Sharp C, Simmonds P, Klenerman P. PARV4: An Emerging Tetraparvovirus. *PLoS Pathogens* 2014;10(5).
- Matthews PC, Sharp CP, Malik A, Gregory WF, Adland E, Jooste P, et al. Human Parvovirus 4 Infection among Mothers and Children in South Africa. *Emerging Infectious Diseases* 2015;21(4):713–715.
- May J, Drexler JF, Reber U, Sarpong N, Adjei O, Panning M, et al. Human parvovirus 4 viremia in young children, Ghana. *Emerging Infectious Diseases* 2012;18(10):1690–1692.
- McGeoch DJ, Cook S, Dolan A, Jamieson FE, Telford EA. Molecular phylogeny and evolutionary timescale for the family of mammalian herpesviruses. *Journal of Molecular Biology* 1995;247(3):443–58.
- McGoldrick SM, Bleakley ME, Guerrero A, Turtle CJ, Yamamoto TN, Pereira SE, et al. Cytomegalovirus-specific T cells are primed early after cord blood transplant but fail to control virus in vivo. *Blood* 2013;121(14):2796–2803.
- Mendelson M, Monard S, Sissons P, Sinclair J. Detection of endogenous human cytomegalovirus in CD34+ bone marrow progenitors. *Journal of General Virology* 1996;77(12):3099–3102.
- Messaoudi I, Guevara-Patino JA, Dyall R, LeMaout J, Nikolich-Žugich J. Direct link between mhc polymorphism, T cell avidity, and diversity in immune defense. *Science* 2002;298(5599):1797–1800.

- Messaoudi I, LeMaoult J, Guevara-Patino JA, Metzner BM, Nikolich-Žugich J. Age-related CD8 T cell clonal expansions constrict CD8 T cell repertoire and have the potential to impair immune defense. *The Journal of Experimental Medicine* 2004;200(10):1347–1358.
- Meyer-Olson D, Brady KW, Bartman MT, O’Sullivan KM, Simons BC, Conrad JA, et al. Fluctuations of functionally distinct CD8+ T-cell clonotypes demonstrate flexibility of the HIV-specific TCR repertoire. *Blood* 2006;107(6):2373–2383.
- Meyer-Olson D, Shoukry NH, Brady KW, Kim H, Olson DP, Hartman K, et al. Limited T cell receptor diversity of HCV-specific T cell responses is associated with CTL escape. *The Journal of Experimental Medicine* 2004;200(3):307–319.
- Michaelis M, Doerr HW, Cinatl J. The story of human cytomegalovirus and cancer: increasing evidence and open questions. *Neoplasia* 2009;11(1):1–9.
- Miconnet I. Probing the T-cell receptor repertoire with deep sequencing. *Current Opinion in HIV and AIDS* 2012;7(1):64–70.
- Miles DJC, van der Sande M, Jeffries D, Kaye S, Ismaili J, Ojuola O, et al. Cytomegalovirus infection in Gambian infants leads to profound CD8 T-cell differentiation. *Journal of Virology* 2007;81(11):5766–76.
- Miles DJC, Sande MVD, Jeffries D, Kaye S, Ojuola O, Sanneh M, et al. Maintenance of large subpopulations of differentiated CD8 T-cells two years after cytomegalovirus infection in Gambian infants. *PLoS ONE* 2008;3(8).
- Miles JJ, Silins SL, Burrows SR. Engineered T cell receptors and their potential in molecular medicine. *Current Medical Chemistry* 2006;13(23):2725–2736.
- Miles JJ, Bulek AM, Cole DK, Gostick E, Schauenburg A, Dolton G, et al. Ge-

- netic and structural basis for selection of a ubiquitous T cell receptor deployed in Epstein-Barr virus infection. *PLoS Pathogens* 2010;6(11).
- Miller-Kittrell M, Sparer TE. Feeling manipulated: cytomegalovirus immune manipulation. *Virology Journal* 2009;6(4).
- Minder W. Die Aetiologie der Cytomegalia infantum. *Schweiz Med Wochenschr* 1953;83:1180–1182.
- Mocarski ES. Immunomodulation by cytomegaloviruses: Manipulative strategies beyond evasion. *Trends in Microbiology* 2002;10(7):332–339.
- Mogensen TH. Pathogen recognition and inflammatory signaling in innate immune defenses. *Clinical Microbiology Reviews* 2009;22(2):240–273.
- Monteiro J, Hingorani R, Choi IH, Silver J, Pergolizzi R, Gregersen PK. Oligoclonality in the human CD8+ T cell repertoire in normal subjects and monozygotic twins: implications for studies of infectious and autoimmune diseases. *Molecular Medicine* 1995;1(6):614–24.
- Moon JJ, Jenkins MK. The human T-cell repertoire grows up. *Immunology and Cell Biology* 2015;93(7):601–602.
- Morley JK, Batliwalla FM, Hingorani R, Gregersen PK. Oligoclonal CD8+ T cells are preferentially expanded in the CD57+ subset. *The Journal of Immunology* 1995;154(11):6182–6190.
- Moss P, Moots R, Rosenberg W, Rowland-Jones S, Bodmer H, McMichael A, et al. Extensive conservation of alpha and beta chains of the human T-cell antigen receptor recognizing HLA-A2 and influenza A matrix peptide. *PNAS* 1991;88(20):8987–90.

- Moss P, Khan N. CD8+ T-cell immunity to cytomegalovirus. *Human Immunology* 2004;65(5):456–464.
- Muenchhoff M, Goulder PJR. Sex differences in pediatric infectious diseases. *The Journal of Infectious Diseases* 2014;209(3).
- Murasko DM, Weiner P, Kaye D. Decline in mitogen induced proliferation of lymphocytes with increasing age. *Clinical and Experimental Immunology* 1987;70(2):440–8.
- Nair-Gupta P, Blander JM. An updated view of the intracellular mechanisms regulating cross-presentation. *Frontiers in Immunology* 2013;4:1–9.
- Naylor K, Li G, Vallejo AN, Lee WW, Koetz K, Bryl E, et al. The influence of age on T cell generation and TCR diversity. *The Journal of Immunology* 2005;174(11):7446–7452.
- Niehues T, Horneff G, Knipp S, Adams O, Wahn V. Treatment-resistant expansion of CD8+CD28-cells in pediatric HIV infection. *Pediatric Research* 2000;47(3):418–21.
- Nikolich-Zugich J. Ageing and life-long maintenance of T-cell subsets in the face of latent persistent infections. *Nature Reviews Immunology* 2008;8(7):512–522.
- Nitta T, Murata S, Sasaki K, Fujii H, Ripen AM, Ishimaru N, et al. Thymoproteasome Shapes Immunocompetent Repertoire of CD8+ T Cells. *Immunity* 2010;32(1):29–40.
- Ocama P, Opio KC, Kagimu M, Seremba E, Wabinga H, Colebunders R. Hepatitis B virus and HIV infection among patients with primary hepatocellular carcinoma in Kampala, Uganda. *African Health Sciences* 2011;11(1):8–11.

- Odeberg J, Plachter B, Brandén L, Söderberg-Nauclér C. Human cytomegalovirus protein pp65 mediates accumulation of HLA-DR in lysosomes and destruction of the HLA-DR alpha-chain. *Blood* 2003;101(12):4870–7.
- van Oers N, Von BH, Weiss A. The pre-T cell receptor (TCR) complex is functionally coupled to the TCR-zeta Subunit subunit. *Journal of Experimental Medicine* 1995;182:1585–1590.
- Oettinger MA, Schatz DG, Gorka C, Baltimore D. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science* 1990;248(4962):1517–1523.
- O’Hara GA, Welten SPM, Klenerman P, Arens R. Memory T cell inflation: Understanding cause and effect. *Trends in Immunology* 2012;33(2):84–90.
- Olaru A, Petrie HT, Livák F. Beyond the 12/23 rule of VDJ recombination independent of the Rag proteins. *The Journal of Immunology* 2005;174(10):6220–6.
- Olsson J, Wikby A, Johansson B, Löfgren S, Nilsson BO, Ferguson FG. Age-related change in peripheral blood T-lymphocyte subpopulations and cytomegalovirus infection in the very old: The Swedish longitudinal OCTO immune study. *Mechanisms of Ageing and Development* 2000;121(1-3):187–201.
- Paiardini M, Müller-Trutwin M. HIV-associated chronic immune activation. *Immunological reviews* 2013;254(1):78–101.
- Panning M, Kobbe R, Vollbach S, Drexler JF, Adjei S, Adjei O, et al. Novel human parvovirus 4 genotype 3 in infants, Ghana. *Emerging Infectious Diseases* 2010;16(7):1143–1146.
- Pantaleo G, Demarest JF, Schacker T, Vaccarezza M, Cohen OJ, Daucher M, et al. The qualitative nature of the primary immune response to HIV infection is a

- prognosticator of disease progression independent of the initial level of plasma viremia. *PNAS* 1997;94(1):254–8.
- Pantaleo G, Demarest JF, Soudeyns H, Graziosi C, Denis F, Adelsberger JW, et al. Major expansion of CD8⁺ T cells with a predominant V beta usage during the primary immune response to HIV. *Nature* 1994;370(6489):463–467.
- Parkin J, Cohen B. An overview of the immune system. *The Lancet* 2001;357(9270):1777–1789.
- Pass RF, Stagno S, Britt WJ, Alford CA. Specific cell-mediated immunity and the natural history of congenital infection with cytomegalovirus. *Journal of Infectious Diseases* 1983;148(6):953–961.
- Pawelec G. Immunosenescence: Role of cytomegalovirus. *Experimental Gerontology* 2014;54:1–5. <http://dx.doi.org/10.1016/j.exger.2013.11.010>.
- Payne R, Muenchhoff M, Mann J, Roberts HE, Matthews P, Adland E, et al. Impact of HLA-driven HIV adaptation on virulence in populations of high HIV seroprevalence. *PNAS* 2014;111(50):E5393–E5400.
- Payne RP, Kløverpris H, Sacha JB, Brumme Z, Brumme C, Buus S, et al. Efficacious early antiviral activity of HIV Gag- and Pol-specific HLA-B*2705-restricted CD8⁺ T cells. *Journal of Virology* 2010;84(20):10543–10557.
- Payne RP, Matthews PC, Prado JG, Goulder PJR. HLA-mediated control of HIV and HIV adaptation to HLA. *Advances in Parasitology* 2009;68:1–20.
- Penfold ME, Dairaghi DJ, Duke GM, Saederup N, Mocarski ES, Kemble GW, et al. Cytomegalovirus encodes a potent alpha chemokine. *PNAS* 1999;96(17):9839–9844.

- Pourgheysari B, Bruton R, Parry H, Billingham L, Fegan C, Murray J, et al. The number of cytomegalovirus-specific CD4⁺ T cells is markedly expanded in patients with B-cell chronic lymphocytic leukemia and determines the total CD4⁺ T-cell repertoire. *Blood* 2010;116(16):2968–2974.
- Prendergast A, Mphatswe W, Tudor-Williams G, Rakgotho M, Pillay V, Thobakgale C, et al. Early virological suppression with three-class antiretroviral therapy in HIV-infected African infants. *AIDS* 2008;22(11):1333–1343.
- Prendergast AJ, Klenerman P, Goulder PJR. The impact of differential antiviral immunity in children and adults. *Nature Reviews Immunology* 2012;12(9):636–648.
- Price DA, Asher TE, Wilson NA, Nason MC, Brenchley JM, Metzler IS, et al. Public clonotype usage identifies protective Gag-specific CD8⁺ T cell responses in SIV infection. *The Journal of Experimental Medicine* 2009;206(4):923–936.
- Price DA, West SM, Betts MR, Ruff LE, Brenchley JM, Ambrozak DR, et al. T cell receptor recognition motifs govern immune escape patterns in acute SIV infection. *Immunity* 2004;21(6):793–803.
- Prösch S, Wendt CEC, Reinke P, Priemer C, Oppert M, Krüger DH, et al. A Novel Link between Stress and Human Cytomegalovirus (HCMV) Infection: Sympathetic Hyperactivity Stimulates HCMV Activation. *Virology* 2000;272(2):357–365.
- Pui JC, Allman D, Xu L, DeRocco S, Karnell FG, Bakkour S, et al. Notch1 expression in early lymphopoiesis influences B versus T lineage determination. *Immunity* 1999;11(3):299–308.
- Putintseva EV, Britanova OV, Staroverov DB, Merzlyak EM, Turchaninova MA,

- Shugay M, et al. Mother and child T cell receptor repertoires: deep profiling study. *Frontiers in Immunology* 2013;4(463).
- Quinnan GV, Kirmani N, Rook AH, Manischewitz JF, Jackson L, Moreschi G, et al. Cytotoxic t cells in cytomegalovirus infection: HLA-restricted T-lymphocyte and non-T-lymphocyte cytotoxic responses correlate with recovery from cytomegalovirus infection in bone-marrow-transplant recipients. *The New England Journal of Medicine* 1982;307(1):7–13.
- Ranganathan P, Clark PA, Kuo JS, Salamat MS, Kalejta RF. Significant association of multiple human cytomegalovirus genomic Loci with glioblastoma multiforme samples. *Journal of Virology* 2012;86(2):854–64.
- Raulet DH. The Structure, Function and Molecular Genetics of the δ/γ T Cell Receptor. *Annual Review of Immunology* 1989;7:75–207.
- Reddehase MJ, Weiland F, Munch K, Jonjic S, Luske A, Koszinowski UH. Interstitial murine cytomegalovirus pneumonia after irradiation: characterization of cells that limit viral replication during established infection of the lungs. *Journal of Virology* 1985;55(2):264–273.
- Redeker A, Welten SPM, Arens R. Viral inoculum dose impacts memory T-cell inflation. *European Journal of Immunology* 2014;44(4):1046–1057.
- Reeves E, Edwards CJ, Elliott T, James E. Naturally Occurring ERAP1 Haplotypes Encode Functionally Distinct Alleles with Fine Substrate Specificity. *The Journal of Immunology* 2013;191(1):35–43.
- Reeves MB, Lehner PJ, Sissons JGP, Sinclair JH. An in vitro model for the regulation of human cytomegalovirus latency and reactivation in dendritic cells by chromatin remodelling. *Journal of General Virology* 2005;86(11):2949–2954.

- Reits E, Griekspoor A, Neijssen J, Groothuis T, Jalink K, Van Veelen P, et al. Peptide Diffusion, Protection, and Degradation in Nuclear and Cytoplasmic Compartments before Antigen Presentation by MHC Class I. *Immunity* 2003;18(1):97–108.
- Repasky JAE, Corbett E, Boboila C, Schatz DG. Mutational Analysis of Terminal Deoxynucleotidyltransferase-Mediated N-Nucleotide Addition in V(D)J Recombination. *The Journal of Immunology* 2004;172(9):5478–5488.
- Reusser BP, Riddell SR, Meyers JD, Greenberg PD. Cytotoxic T-Lymphocyte response to cytomegalovirus after human allogeneic bone marrow transplantation: pattern of recovery and correlation with cytomegalovirus infection and disease. *Blood* 2015;78(5):1373–1381.
- Reusser P, Cathomas G, Attenhofer R, Tamm M, Thiel G. Cytomegalovirus (CMV)-specific T cell immunity after renal transplantation mediates protection from CMV disease by limiting the systemic virus load. *The Journal of Infectious Diseases* 1999;180(2):247–53.
- Revello MG, Zavattoni M, Sarasini A, Percivalle E, Simoncini L, Gerna G. Human cytomegalovirus in blood of immunocompetent persons during primary infection: prognostic implications for pregnancy. *The Journal of Infectious Diseases* 1998;177(5):1170–1175.
- Ribbert H. Veber protozoenartig Zellen in der Niere eines syphilitischen Negeborenen und in der Parotis von Kindern. *Zentralbl All pathol* 1904;15:945–948.
- Riddell S, Greenberg P. T cell therapy of human CMV and EBV infection in immunocompromised hosts. *Reviews in Medical Virology* 1997;7(3):181–192.
- Riddell SR, Watanabe KS, Goodrich JM, Li CR, Agha ME, Greenberg PD. Restora-

- tion of viral immunity in immunodeficient humans by the adoptive transfer of T cell clones. *Science* 1992;257(5067):238–41.
- Rist M, Smith C, Bell MJ, Burrows SR, Khanna R. Cross-recognition of HLA DR4 alloantigen by virus-specific CD8⁺ T cells : a new paradigm for self-/ nonself-recognition. *Blood* 2009;114(11):2244–2253.
- Roberts ET, Haan MN, Dowd JB, Aiello AE. Cytomegalovirus antibody levels, inflammation, and mortality among elderly latinos over 9 years of follow-up. *American Journal of Epidemiology* 2010;172(4):363–371.
- Robey E, Fowlkes BJ. Selective events in T cell development. *Annual Review of Immunology* 1994;12:675–705.
- Robins H, Srivastava S, Compregner PV, Turtle C, Andriesen J, Riddell S, et al. Overlap and effective size of the human CD8⁺ T-cell receptor repertoire. *Science Translational Medicine* 2010;2(47).
- Rock KL, Gramm C, Rothstein L, Clark K, Stein R, Dick L, et al. Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules. *Cell* 1994;78(5):761–771.
- Rosa CL, Diamond DJ. The immune response to human CMV. *Future Virology* 2012;7(3):279–293.
- Rowe W, Hartley J, Waterman S, Turner H, Huebner R. Cytopathogenic agent resembling human salivary gland virus recovered from tissue cultures of human adenoids. *Experimental Biology and Medicine* 1956;92(2):418–424.
- Rudd B, Venturi V, Davenport M, Nikolich-Zugich J. Evolution of the antigen-specific CD8⁺ TCR repertoire across the lifespan: evidence for clonal homog-

- enization of the old TCR repertoire. *Journal of Immunology* 2011;186(4):2056–2064.
- Sallusto F, Geginat J, Lanzavecchia A. Central memory and effector memory T cell subsets: function, generation, and maintenance. *Annual Review of Immunology* 2004;22(1):745–763.
- Sallusto F, Lenig D, Förster R, Lipp M, Lanzavecchia A. Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. *Nature* 1999;401(6754):708–712.
- Schmidt J, Neumann-Haefelin C, Altay T, Gostick E, Price DA, Lohmann V, et al. Immunodominance of HLA-A2-restricted hepatitis C virus-specific CD8+ T cell responses is linked to naive-precursor frequency. *Journal of Virology* 2011;85(10):5232–6.
- Schneider B, Fryer JF, Reber U, Fischer Hp, Tolba RH, Baylis SA, et al. Persistence of novel human Parvovirus PARV4 in liver tissue of adults. *Journal of Medical Virology* 2008;80(2):345–351.
- Schubert U, Antón LC, Gibbs J, Norbury CC, Yewdell JW, Bennink JR. Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature* 2000;404(6779):770–774.
- Sciammas R, Kodukula P, Tang Q, Hendricks RL, Bluestone JA. T cell receptor-gamma/delta cells protect mice from herpes simplex virus type 1-induced lethal encephalitis. *The Journal of Experimental Medicine* 1997;185(11):1969–75.
- Servant-Delmas A, Laperche S, Lionnet F, Sharp C, Simmonds P, Lefrère J. Human parvovirus 4 infection in low-and high-risk French individuals. *Transfusion* 2014;54(3):744–5.

- Sester M, Sester U, Gärtner B, Girndt M, Meyerhans A, Köhler H. Dominance of Virus-Specific CD8 T Cells in Human Primary Cytomegalovirus Infection. *Journal of the American Society of Nephrology* 2002;13(10):2577–2584.
- Sester U, Gärtner BC, Wilkens H, Schwaab B, Wössner R, Kindermann I, et al. Differences in CMV-specific T-cell levels and long-term susceptibility to CMV infection after kidney, heart and lung transplantation. *American Journal of Transplantation* 2005;5(6):1483–1489.
- Sewell AK. Why must T cells be cross-reactive? *Nature Reviews Immunology* 2012;12(9):669–677.
- Shapiro R, Hughes M, Ogwu A, Kitch D, Lockman S, Moffat C, et al. Antiretroviral Regimens in Pregnancy and Breast-Feeding in Botswana. *New England Journal of Medicine* 2010;362(24):2282–2294.
- Sharp CP, Gregory WF, Hattingh L, Malik A, Adland E, Daniels S, et al. PARV4 prevalence, phylogeny, immunology and coinfection with HIV, HBV and HCV in a multicentre African cohort. *Wellcome Open Research* 2017;2:26. <https://wellcomeopenresearch.org/articles/2-26/v1>.
- Sharp CP, Lail A, Donfield S, Gomperts ED, Simmonds P. Virologic and clinical features of primary infection with human parvovirus 4 in subjects with hemophilia: Frequent transmission by virally inactivated clotting factor concentrates. *Transfusion* 2012;52(7):1482–1489.
- Sharp CP, Lail A, Donfield S, Simmons R, Leen C, Delwart E, et al. High frequencies of exposure to the novel human parvovirus, PARV4 in haemophiliacs and injecting drug users detected by a serological assay for PARV4 antibodies. *Journal of Infectious Diseases* 2010;200(7):1119–1125.

- Sharp CP, Vermeulen M, Nebie Y, Djoko CF, Lebreton M, Tamoufe U, et al. Epidemiology of human parvovirus 4 infection in sub-Saharan Africa. *Emerging Infectious Diseases* 2010;16(10):1605–1607.
- Shen DF, Doukhan L, Kalams S, Delwart E. High-resolution analysis of T-cell receptor β -chain repertoires using DNA heteroduplex tracking: Generally stable, clonal CD8 + expansions in all healthy young adults. *Journal of Immunological Methods* 1998;215(1-2):113–121.
- Sidney J, Southwood S, Paschetto V, Sette A. Simultaneous prediction of binding capacity for multiple molecules of the HLA B44 supertype. *The Journal of Immunology* 2003;171(11):5964–74.
- Siegrist CA. Neonatal and early life vaccinology. *Vaccine* 2001;19(25-26):3331–3346.
- Silins BSL, Cross SM, Elliott SL, Pye SJ, Burrows S, Burrows JM, et al. Development of Epstein-BarrVirus-specific Memory T Cell Receptor Clonotypes in Acute Infectious Mononucleosis. *The Journal of Experimental Medicine* 1996;184:1815–1824.
- Simmonds P, Douglas J, Bestetti G, Longhi E, Antinori S, Parravicini C, et al. A third genotype of the human parvovirus PARV4 in sub-Saharan Africa. *Journal of General Virology* 2008;89:2299–2302.
- Simmonds P, Manning A, Kenneil R, Carnie FW, Bell JE. Parenteral transmission of the novel human parvovirus PARV4. *Emerging Infectious Diseases* 2007;13(9):1386–1388.
- Simmons R, Sharp C, Levine J, Bowness P, Simmonds P, Cox A, et al. Evolution of CD8+ T cell responses after acute PARV4 infection. *Journal of Virology* 2013;87(6):3087–96.

- Simmons R, Sharp C, McClure CP, Rohrbach J, Kovari H, Frangou E, et al. Parvovirus 4 infection and clinical outcome in high-risk populations. *Journal of Infectious Diseases* 2012;205(12):1816–1820.
- Simmons R, Sharp C, Sims S, Klooverpris H, Goulder P, Simmonds P, et al. High frequency, sustained T cell responses to PARV4 suggest viral persistence in vivo. *Journal of Infectious Diseases* 2011;203(10):1378–1387.
- Simon AK, Hollander GA, McMichael A. Evolution of the immune system in humans from infancy to old age. *Proc Biol Sci* 2015;282(1821).
- Sinclair J. Human cytomegalovirus: Latency and reactivation in the myeloid lineage. *Journal of Clinical Virology* 2008;41(3):180–185.
- Sinclair J. Chromatin structure regulates human cytomegalovirus gene expression during latency, reactivation and lytic infection. *Biochimica et Biophysica Acta* 2010;1799(3-4):286–295.
- Sinclair J, Sissons P. Latency and reactivation of human cytomegalovirus. *Journal of General Virology* 2006;87(7):1763–1779.
- Singer A, Adoro S, Park Jh. Lineage fate and intense debate: myths, models and mechanisms of CD4/CD8 lineage choice. *Nature Review Immunology* 2008;8(10):788–801.
- Skaletskaya A, Bartle LM, Chittenden T, McCormick AL, Mocarski ES, Goldmacher VS. A cytomegalovirus-encoded inhibitor of apoptosis that suppresses caspase-8 activation. *PNAS* 2001;98(14):7829–7834.
- Smith M. Propagation in tissue cultures of a cytopathogenic virus from human salivary gland virus (SVG) disease. *Proc Soc Exp Biol Med* 1956;92:424–430.

- Snyder CM, Cho KS, Morrison EL, Dommelen S, Shellam GR, Hill AB. Memory Inflation During Chronic Viral Infection is Maintained by Continuous Production of Short-Lived Functional T Cells. *Immunity* 2008;29(4):650–659.
- Söderberg-Nauclér C, Fish KN, Nelson JA. Reactivation of latent human cytomegalovirus by allogeneic stimulation of blood cells from healthy donors. *Cell* 1997;91(1):119–126.
- Spiliotis ET, Osorio M, Zúñiga MC, Edidin M. Selective export of MHC Class I molecules from the ER after their dissociation from TAP. *Immunity* 2000;13(6):841–51.
- Spits H. Development of $\alpha\beta$ T cells in the human thymus. *Nature Reviews Immunology* 2002;2(10):760–772.
- Stabinski L, Reynolds SJ, Ocama P, Laeyendecker O, Boaz I, Ndyanabo A, et al. High prevalence of liver fibrosis associated with HIV infection: a cross-sectional study in rural Rakai, Uganda. *Antiviral therapy* 2011;16(3):405–411.
- Staras S, Dollard S, Radford K, Flanders W, Pass R, Cannon M. Seroprevalence of cytomegalovirus infection in the United States, 1988-1994. *Clinical Infectious Diseases* 2006;43:1143–51.
- Stetson DB, Medzhitov R. Type I Interferons in Host Defense. *Immunity* 2006;25(3):373–381.
- Stinski MF. Sequence of protein synthesis in cells infected by human cytomegalovirus: early and late virus-induced polypeptides. *Journal of Virology* 1978;26(3):686–701.
- Suessmuth Y, Mukherjee R, Watkins B, Koura DT, Finstermeier K, Desmarais C,

- et al. CMV reactivation drives posttransplant T-cell reconstitution and results in defects in the underlying TCR β repertoire. *Blood* 2015;125(25):3835–3850.
- Sylwester A, Mitchell B, Edgar J, Taormina C, Pelte C, Ruchti F, et al. Broadly targeted human cytomegalovirus-specific CD4⁺ and CD8⁺ T cells dominate the memory compartments of exposed subjects. *The Journal of Experimental Medicine* 2005;202(5):673–685.
- Tan ACL, La Gruta NL, Zeng W, Jackson DC. Precursor frequency and competition dictate the HLA-A2-restricted CD8⁺ T cell responses to Influenza A infection and vaccination in HLA-A2. *The Journal of Immunology* 2011;187(4):1895–1902.
- Tembo J, Kabwe M, Chilukutu L, Chilufya M, Mwaanza N, Chabala C, et al. Prevalence and risk factors for betaherpesvirus DNAemia in children >3 weeks and <2 years of age admitted to a large referral hospital in sub-Saharan Africa. *Clinical Infectious Diseases* 2015;60(3):423–431.
- Thimme R, Wieland S, Steiger C, Ghrayeb J, Reimann KA, Purcell RH, et al. CD8⁺ T cells mediate viral clearance and disease pathogenesis during acute hepatitis B virus Infection. *Journal of Virology* 2003;77(1):68–76.
- Tomasec P, Braud V, Rickards C, Powell M, McSharry B, Gadola S, et al. Surface Expression of HLA-E, an Inhibitor of Natural Killer Cells, Enhanced by Human Cytomegalovirus gpUL40. *Science* 2000;287(5455):1031–1033.
- Tomasec P, Wang ECY, Davison AJ, Vojtesek B, Armstrong M, Griffin C, et al. Downregulation of natural killer cellactivating ligand CD155 by human cytomegalovirus UL141. *Nature Immunology* 2010;6(2):181–188.
- Tomazin R, Boname J, Hegde NR, Lewinsohn DM, Altschuler Y, Jones TR, et al. Cytomegalovirus US2 destroys two components of the MHC class II pathway, preventing recognition by CD4⁺ T cells. *Nature Medicine* 1999;5(9):1039–1043.

- Tomtishen J. Human cytomegalovirus tegument proteins (pp65, pp71, pp150, pp28). *Virology Journal* 2012;9(22).
- Tonegawa S. Somatic generation of antibody diversity. *Nature* 1983;302(5909):575–581.
- Topham NJ, Hewitt EW. Natural killer cell cytotoxicity: How do they pull the trigger? *Immunology* 2009;128(1):7–15.
- Torti N, Walton SM, Murphy KM, Oxenius A. Batf3 transcription factor-dependent DC subsets in murine CMV infection: Differential impact on T-cell priming and memory inflation. *European Journal of Immunology* 2011;41(9):2612–2618.
- Tortorella D, Gewurz BE, Furman MH, Schust DJ, Ploegh HL. Viral Subversion of the Immune System. *Annual Review of Immunology* 2000;18:861–926.
- Trapani JA, Smyth MJ. Functional significance of the perforin/granzyme cell death pathway. *Nature Reviews Immunology* 2002;2(10):735–747.
- Trepo C, Chan H, Lok A. Hepatitis B virus infection. *The Lancet* 2014;384(9959):2053–63.
- Tu W, Chen S, Sharp M, Dekker C, Manganello AM, Tongson EC, et al. Persistent and selective deficiency of CD4⁺ T cell immunity to cytomegalovirus in immunocompetent young children. *The Journal of Immunology* 2004;172(5):3260–3267.
- Turchaninova MA, Britanova OV, Bolotin DA, Shugay M, Putintseva EV, Staroverov DB, et al. Pairing of T-cell receptor chains via emulsion PCR. *European Journal of Immunology* 2013;43(9):2507–2515.
- Turner JE, Campbell JP, Edwards KM, Howarth LJ, Pawelec G, Aldred S, et al. Rudimentary signs of immunosenescence in Cytomegalovirus-seropositive healthy young adults. *Age* 2014;36(1):287–297.

- Turner S, La Gruta N, Kedzierska K, Thomas P, Doherty P. Functional implications of T cell receptor diversity. *Current Opinion in Immunology* 2009;21(3):286–290.
- Varela-Rohena A, Molloy PE, Dunn SM, Li Y, Suhoski MM, Carroll RG, et al. Control of HIV-1 immune escape by CD8 T cells expressing enhanced T-cell receptor. *Nature Medicine* 2008;14(12):1390–1395.
- Varnum SM, Streblow DN, Monroe ME, Smith P, Auberry KJ, Pasa-Tolic L, et al. Identification of Proteins in Human Cytomegalovirus (HCMV) Particles: the HCMV Proteome. *Journal of Virology* 2004;78(20):10960–10966.
- Venema H, van den Berg A, van Zanten C, van Son WJ, van der Giessen M, The TH. Natural killer cell responses in renal transplant patients with cytomegalovirus infection. *Journal of Medical Virology* 1994;42(2):188–92.
- Venturi V, Quigley M, Greenaway H, Ng P, Ende Z, McIntosh T, et al. A mechanism for TCR sharing between T cell subsets and individuals revealed by pyrosequencing. *The Journal of Immunology* 2011;186(7):4285–4294.
- Venturi V, Kedzierska K, Price DA, Doherty PC, Douek DC, Turner SJ, et al. Sharing of T cell receptors in antigen-specific responses is driven by convergent recombination. *PNAS* 2006;103(49):18691–6.
- Venturi V, Kedzierska K, Tanaka MM, Turner SJ, Doherty PC, Davenport MP. Method for assessing the similarity between subsets of the T cell receptor repertoire. *Journal of Immunological Methods* 2008;329(1-2):67–80.
- Venturi V, Price Da, Douek DC, Davenport MP. The molecular basis for public T-cell responses? *Nature Reviews Immunology* 2008;8(3):231–238.
- VonGlahn W, Pappenheimer A. Intranuclear inclusions in visceral disease. *The American Journal of Pathology* 1925;1(5):445–465.

- Wallace JM, Hannah J. Cytomegalovirus pneumonitis in patients with AIDS. *Chest* 1987;92(2):198–203.
- Waller ECP, McKinney N, Hicks R, Carmichael aJ, Sissons JGP, Wills MR. Differential costimulation through CD137 (41BB) restores proliferation of human virus-specific effector memory (CD28 CD45RA HI) CD8+ T cells. *Blood* 2007;110(13):4360–4366.
- Walter EA, Greenberg PD, Gilbert MJ, Finch RJ, Watanabe KS, Thomas ED, et al. Reconstitution of cellular immunity against cytomegalovirus in recipients of allogeneic bone marrow by transfer of T-cell clones from the donor. *New England Journal of Medicine* 1995;333(16):1038–1044.
- Wang EC, Moss P, Frodsham P, Lehner P, Bell J, Borysiewicz LK. CD8^{high}CD57⁺ T lymphocytes in normal, healthy individuals are oligoclonal and respond to human cytomegalovirus. *The Journal of Immunology* 1995;155(10):5046–5056.
- Waring P, Müllbacher A. Cell death induced by the Fas/Fas ligand pathway and its role in pathology. *Immunology and Cell Biology* 1999;77(4):312–317.
- Wedderburn L, Maini MK, Patel A, Beverley PC, Woo P. Molecular fingerprinting reveals non-overlapping T cell oligoclonality between an inflamed site and peripheral blood. *International immunology* 1999;11(4):535–43.
- Wedderburn L, Patel A, Varsani H, Woo P. The developing human immune system: T-cell receptor repertoire of children and young adults shows a wide discrepancy in the frequency of persistent oligoclonal T-cell expansions. *Immunology* 2001;102(3):301–309.
- Weekes MP, Wills MR, Mynard KIM, Carmichael AJ, Sissons JGP. The memory cytotoxic T-lymphocyte (CTL) response to human cytomegalovirus infection con-

- tains individual peptide-specific CTL clones that have undergone extensive expansion in vivo. *Journal of Virology* 1999;73(3):2099–2108.
- Welgama U, Wickramasinghe C, Perera J. Varicella-zoster virus infection in the Infectious Diseases Hospital, Sri Lanka. *Ceylon Medical Journal* 2003;48(4):119–21.
- Weller T. Cytomegalovirus: the difficult years. *Journal of Infectious Disease* 1970;122:532–53.
- Weller T, Macauley J, Craig J, Wirth P. Isolation of intranuclear inclusion producing agents from infants with illnesses resembling cytomegalic inclusion disease. *Proc Soc Exp Biol Med* 1957;94(1):4–12.
- Willinger T, Freeman T, Hasegawa H, McMichael AJ, Callan MFC. Molecular Signatures Distinguish Human Central Memory from Effector Memory CD8 T Cell Subsets. *The Journal of Immunology* 2005;175(9):5895–5903.
- Wills MR, Ashiru O, Reeves MB, Okecha G, Trowsdale J, Tomasec P, et al. Human Cytomegalovirus Encodes an MHC Class I-Like Molecule (UL142) That Functions to Inhibit NK Cell Lysis. *The Journal of Immunology* 2005;175(11):7457–7465.
- Wills MR, Okecha G, Weekes MP, Gandhi MK, Sissons PJG, Carmichael AJ. Identification of naive or antigen-experienced human CD8(+) T cells by expression of costimulation and chemokine receptors: analysis of the human cytomegalovirus-specific CD8(+) T cell response. *The Journal of Immunology* 2002;168(11):5455–5464.
- Wyatt J, Saxton J, Lee R, Pinkerton H. Generalized cytomegalic inclusion disease. *The Journal of Pediatrics* 1950;36:271–294.

- Wynn KK, Fulton Z, Cooper L, Silins SL, Gras S, Archbold JK, et al. Impact of clonal competition for peptide-MHC complexes on the CD8 + T-cell repertoire selection in a persistent viral infection. *Blood* 2008;111(8):4283–4292.
- Yang PL, Althage A, Chung J, Maier H, Wieland S, Isogawa M, et al. Immune effectors required for hepatitis B virus clearance. *PNAS* 2010;107(2):798–802.
- Yang SJ, Hung Cc, Chang Sy, Lee Kl, Chen My. Immunoglobulin G and M antibodies to human parvovirus 4 (PARV4) are frequently detected in patients with HIV-1 infection. *Journal of Clinical Virology* 2011;51:64–67.
- Yang X, Gao M, Chen G, Pierce BG, Lu J, Weng NP, et al. Structural basis for clonal diversity of the public T cell response to a dominant human cytomegalovirus epitope. *Journal of Biological Chemistry* 2015;290(48):29106–29119.
- Yeager AS, Grumet CF, Hafleigh EB, Arvin AM, Bradley JS, Prober CG. Prevention of transfusion-acquired cytomegalovirus infections in newborn infants. *The Journal of Pediatrics* 1981;98(2):281–287.
- Yewdell JW, Reits E, Neefjes J. Making sense of mass destruction: quantitating MHC class I antigen presentation. *Nature Reviews Immunology* 2003;3(12):952–961.
- Zhang X, Zhivaki D, Lo-Man R. Unique aspects of the perinatal immune system. *Nature Reviews Immunology* 2017;17(8):495–507.
- Zimmerli SC, Harari A, Cellerai C, Vallelian F, Bart Pa, Pantaleo G. HIV-1-specific IFN- γ /IL-2-secreting CD8 T cells support CD4-independent proliferation of HIV-1-specific CD8 T cells. *PNAS* 2005;102(20):7239–44.