Investigating cross-clade immune responses in HIV-1 subtype C-infected individuals from South Africa: Implications for HIV Vaccine Design

By

Lycias Zembe

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy (PhD) in the Department of Clinical Laboratory Sciences, Division of Medical Virology, University of Cape Town

February 2012

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Firstly, I offer my sincere gratitude to my supervisor Dr Wendy Burgers and co-supervisor Prof Carolyn Williamson who supported me throughout the project and thesis write up. I am so grateful for their continued encouragement and unfailing effort and all the support in every way.

I would like to acknowledge and Professor Clive Gray, initially our research collaborator and now my mentor who without fail kindly provided immunology reagents, advice and insight into the work reported in this thesis.

I would like to express my deepest appreciation to Kathi and Richard Bouma, Phillipa and David Rickards who have been there all the time with every form of support. Thank you so much may God bless you all. I would also like to thank Barbara Naussbum and Jane Griffiths for their invaluable support throughout the thesis writing up period.

Many thanks the Geographical Information System (GIS) Unit, University of Cape Town particularly Nicholas Lindenberg for his help with the mapping software. Rubi and Agano, thank you so much for flow cytometry panel optimization. To Cobus, Kathryn, Nyasha, Tinashe, Sibo, Marcel, Nthabie and Ramla, you are all awesome. Thanks a lot for all your support.

Thanks to Dr Catherine Riou for her insight on flow cytometry data analysis.

My sincere appreciation goes to South African AIDS Vaccine Initiative (SAAVI), Poliomyelitis Research Foundation (PRF) and the Beit Student Support Fund for their financial support, International AIDS Vaccine Initiative (IAVI) for funding the project and The Bill and Melinda Gates Foundation for Global Travel Awards that gave me the opportunity to travel to Keystone Symposioum international conferences on HIV/AIDS.

To God be the Glory!!!!

LIST OF ACRONYMS AND ABBREVIATIONS

ABI	- Applied Biosystems
ANC	-Ancenstral/Ancestor
APC Cy7	-Allophycocyanin Cyanine 7
APOBEC3G	-Apolipoprotein B mRNA-editing, enzyme-catalytic, polypeptide-like 3G
Bp	- Base pair
β-TrCP2	-Beta-transducin-repeat-containing protein
bNAb	-Broadly Neutralizing Antibody
BST-2	-Bone marrow stromal cell antigen-2
C_{CH}	- Chinese subtype C sequence
C_{Du422}	- South African subtype C Du422 sequence
CD	-Compact disc
cDNA	- Complementary deoxyribonucleic acid
CD14	-Cluster of Differentiation marker 14
CD19	-Cluster of Differentiation marker 19
CD4	- Cluster of differentiation 4
CD8	- Cluster of differentiation 8
CEF	- Cytomegalovirus, Epstein-Bar virus and Influenza virus
CI	- Confidence interval
CPT	- cell preparatory tube
CTL	- Cytotoxic T-lymphocyte
Con	-Consensus
COT	-Centre-of-tree
CRF	-Circulating recombinant form
dH_2O	- distilled water
DMSO	- Dimethylsulphoxide
ddNTP	- Dideoxynucleotide triphosphate
DECP	- Diethyl pyrocarbonate
DNA	- Deoxyribonucleic acid
DNases	- Deoxyribonucleases

dNTP	- Deoxynucleotide triphosphate
EDTA	- Ethylenediaminetetraacetic acid
ELISpot	- Enzyme Linked ImmunoSpot
Env	- Envelope
Epicover	-Epitope coverage
FCS	- Foetal Calf Serum
FITC	-Fluorescein isothiocyanate
Gag	- Group specific antigen
HIV-1	- Human immunodeficiency virus type 1
HLA	- Human leukocyte antigen
HRP	- Horseradish Peroxidase
IAVI	- International AIDS Vaccine Initiative
ICS	-Intracellular cytokine staining
IFN-γ	- Interferon gamma
IL-2	-Interleukin-2
IIDMM	- Institute of Infectious Diseases and Molecular Medicine
IN	- Integrase
kD	- KiloDalton
LTR	- Long terminal repeat
mAb	-Monoclonal antobody
MEGA	- Molecular Evolutionary Genetics Analysis
MFI	-Median Florescence Intensity
MHC	- Major Histocompatibility Complex
MIP-1β	-Macrophage Inflammatory Protein One-Beta
mM	- Milimolar
ml	- Millilitre
MRCA	-Most Recent Common Ancestor
NICD	- National Institute for Communicable Diseases
NJ	- Neighbor-Joining
Pac Blue	-Pacific Blue
PAUP	- Phylogenetic Analysis Using Parsimony
PBMC	- Peripheral Blood mononuclear cells

PBS	- Phosphate Buffer Saline
PCR	- Polymerase chain reaction
PE	-Phycoerythrin
PE Cy7	-Phycoerythrin Cyanine 7
PerCP Cy5.5	-Peridin Chlorophyll Protein Cyanine 5.5
PHA	- Phytohaemagglutinin
PIC	-Preintegration complex
Pmol	- Picomole
Pol	- Polymerase
Posicover	-Positional epitope coverage
PR	-Protective ratio
PTE	-Potential T-cell epitope
P-TEFb	- Positive-acting transcription elongation factor b
p-value	-Probability value
QC	- Quality control
RCF-*g*	- Relative Centrifugal Force *g*
Rev	- Regulator of Virion
RH10	-RPMI with 10% human AB serum
RNA	- Ribonucleic acid
RNases	- Ribonucleases
RPM	- Revolution per minute
RPMI	- Roswell Park Memorial Institute media
RT	- Room temperature (15^0C-25^0C) or Reverse Transcriptase
R10	- RPMI media with 10% FCS
R20	- RPMI media with 20% FCS
SAAVI	- South African AIDS Vaccine Initiative
SEB	-Staphylococcus enterotoxin B
SFU	-Spot Forming Units
TCR	-T-Cell Receptor
TBE	- Tris Boric acid EDTA buffer
TNF-α	-Tumor Necrosis Factor-alpha
TR	- Reverse Transcriptase

UV	- Ultraviolet
URF	-Unique recombinant form
Vif	- Viral infectivity factor
Vpr	-Viral protein R
Vpu	- Viral protein u
µg	- Microgram
µl	- Microlitre
^{0}C	- Degrees Celsius

LIST OF FIGURES

Figures in thesis chapters

Figures from thesis appendices

LIST OF TABLES

Tables from thesis chapters

Tables from thesis appendices

ABSTRACT

Background

An effective HIV vaccine remains the main hope for controlling the HIV epidemic and is a global health priority. The genetic diversity of the virus across the globe is a major impediment to developing an effective vaccine. Whether a universal vaccine is possible still remain elusive. Therefore, there is need to fully characterise clusters of commonly targeted regions across the different HIV-1 clades. Centralized sequences have been suggested as vaccine immunogens and peptide reagents for assessing vaccine-induced responses, but their cross-reactivity has not been fully assessed in larger cohorts of subtype C-infection and in regions of differing HIV epidemics. In addition, the functional profile of HIV-specific T-cells recognizing variant epitopes has not been fully characterized. Whether cross-reactivity observed by IFN-γ production in an ELISpot assay can be observed at physiological concentrations of the peptides and for other functions of HIV-specific T-cells is an important question that remains to be answered.

Methods

The cross-reactivity of HIV-specific T-cells was assessed using clade-specific peptide reagents forming part of current candidate vaccine inserts based on the HIV-1 Gag protein from clades C_{Du422}, C_{CH}, A, B and D in 40 clade C-infected study participants using the IFN-γ ELISpot assay. To test the reactivity of group M consensus peptide reagents, 66 individuals, 44 of whom were ARV naïve, were assessed for HIV-specific T cell responses to group M Gag and Nef peptides. A selection of these individuals was screened for HIV-specific T-cell responses to clade C_{Du422} Gag peptides. Cross-reactivity of peptide variants was assessed at physiologically relevant peptide concentrations by functional avidity studies using peptide dilution IFN-γ ELISpot assays. Additionally, the cytokine profile, cytotoxic potential and proliferative capacity of cross-reactive peptide variants was characterised using multiparameter flow cytometry.

Results

The magnitude and breadth of HIV-specific T-cell responses were similar between the two clade-C peptide reagents in a clade C-infected population. However, the magnitude and breadth of responses to peptides based on clades A, B and D were significantly lower compared to the clade C peptides. Clusters of commonly targeted regions cross-reactive

across the four clades investigated resided predominantly in conserved regions. Interestingly, there were Gag regions that were exclusively recognized in the different clades that had significantly lower entropy scores for the reactive variants than their non-reactive counterparts, suggesting that the variability in targeted regions could have been shaped by host immune pressure. For consensus group M peptides, the magnitude and breadth of Gag responses were significantly higher than that of Nef in clade C-infected individuals. In addition, consensus group M Gag peptides had significantly lower magnitude and breadth of HIV-specific T-cell responses compared to clade C peptide reagents, suggesting loss of responses by centralised reagents despite their central nature to group M viruses. On the contrary, the magnitude and breadth of responses to consensus group M Gag peptides were comparable to that of clade-mismatched peptides, namely clades A, B and D, suggesting that these reagents can be used interchangeably. Peptide dilution assays showed that amino acid mismatches have discordant effects on functional avidity and that some peptides are cross-reactive at physiological concentrations. Similarly, discordant effects (differences in functional avidity, cytokine and cytotoxic profiles and proliferative capacity) of amino acid mismatches on cytokine and cytotoxic potential profiles as well as proliferative capacity were observed.

Conclusion

People infected by a particular HIV clade can recognize HIV peptides based on other clades. However, the magnitude and breadth of responses are greater for the matched clades compared to mismatched clades, suggesting that there may be an advantage of using vaccines based on matched over unmatched clades. Group M based consensus sequences can be recognized in HIV-infected individuals, but with a lower frequency, magnitude and breadth of responses compared to clade-matched peptides, suggesting a limitation of these peptides both as reagents and vaccine immunogens. However, the frequency, magnitude and breadth T-cell reponses to consensus group M peptides were comparable to clade-mismatched responses, suggesting that these reagents may be used interchangeably. Furthermore, amino acid variations across corresponding viral regions have discordant effects on the functional avidity, cytokine profile, cytotoxic potential and proliferative capacity; implying that qualitative measures of cross-reactivity beyond IFN-γ frequencies need to be considered. These data may aid in the development of reagents for the assessment of vaccine-induced responses as well as in HIV vaccine immunogen design.

CHAPTER 1: LITERATURE REVIEW

Table of Contents

Worldwide, there were an estimated 34 million people living with HIV at the end of 2010, with 17% of these having been infected since 2001 (UNAIDS Epidemic Update, 2011). Sub-saharan Africa remains the worst affected by the HIV/AIDS epidemic, with 68% of all HIV infections residing in this region (UNAIDS Epidemic Update, 2011). The degree of the burden is clearly apparent from the fact that Sub-Saharan African accounts for only 12% of the global population yet nearly 70% of all new HIV infections in 2010 were from this region. The epidemic is more severe in South Africa, which has the highest number of HIV infections than any other country in the world, estimated at 5.7 million out of a population of 48 million (UNAIDS Epidemic Update, 2011).

To some extent, behavior change in certain countries and greater access to antiretroviral treatment (ART) has recently stabilised or reduced the number of new HIV infections and HIV/AIDS-related deaths worldwide, respectively (UNAIDS Epidemic Update, 2011). However, a safe and effective HIV vaccine is currently a critically important global health priority to halt new infections (there were still a staggering 2.7 million new infections in 2010; UNAIDS Epidemic Update, 2011), as well as for long-term control of the epidemic. The development of such a vaccine faces many challenges, the most important being the huge genetic diversity and plasticity the virus exhibits, our incomplete knowledge of what constitutes a durable and effective immune response against the virus, and how to generate such a response with a vaccine. Therefore, a greater understanding of the biology of the virus and its interaction with the immune system, its diversity and distribution worldwide, and immune correlates of protection in natural HIV infection are of major importance in the design of an effective HIV vaccine. This chapter provides an overview of the key concepts in HIV biology, diversity, immunity and vaccine design pertinent to the studies described in the thesis.

1.1. The genetic organization and gene products of HIV

The HIV genome is approximately 9.8 kilobase pairs long and is composed of nine genes that code for 15 proteins; three structural or enzyme polyproteins whose products are essential components of the viral particle (Gag, further processed into p24, p17 and p6, Env, further processed into gp120 and gp41, and Pol, which is further processed into

four enzymes, namely integrase, reverse transcriptase, protease and RNase-H), two regulatory proteins (Tat and Rev) and four accessory proteins (Nef, Vif, Vpr and Vpu, Figure 1.1, http://hiv-web.lanl.gov). The genome has two long terminal repeat units (LTR), at the unique 5' (U5) and unique 3' (U3) ends that have control elements required for transcription of the viral genome (Frenza *et al*., 1998). Previous studies have shown that almost all of the expressed HIV proteins are immunogenic (Addo *et al*., 2003). In terms of variability, these proteins can vary by as much as 5-25% within a single HIV clade and by 10-40% between different HIV-1 clades (Kuiken *et al*., 2010).

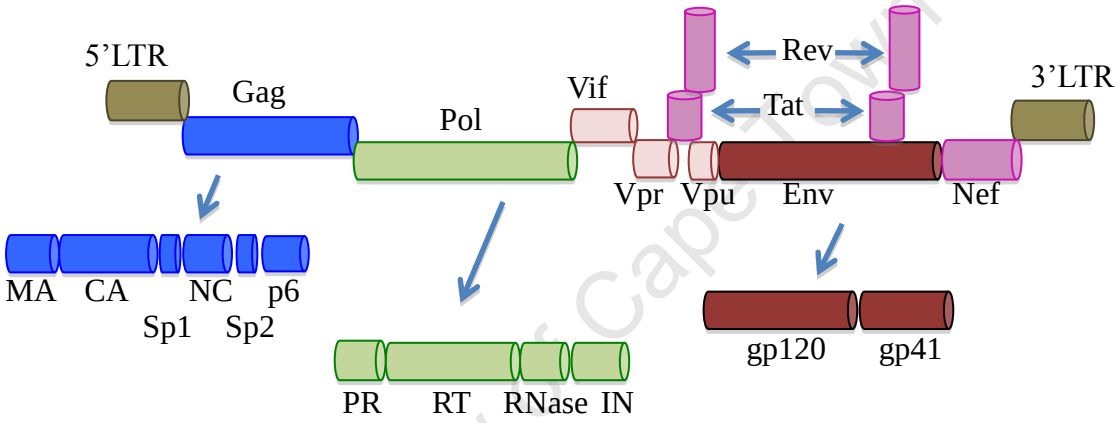


Figure 1.1. A diagrammatic representation of the genetic organization of HIV genome. The proviral genome is composed of nine genes that are flanked by two long terminal repeats (LTR, 5' and 3'). The nine open reading frames code for at least 15 proteins. The *gag* gene is translated into the structural Gag precursor ($Pr55^{Gag}$) that is cleaved into matrix, p17 (MA), capsid, CA (p24), nucleocapsid, NC (p7) and p6 (Sp1 and Sp2 are spacer proteins). The *pol* gene is translated by a -1 ribosomal frameshift as a Gag-Pol precursor polypeptide that yields the four enzymes Reverse Transcriptase (RT), Integrase (IN), Protease (PR) and Ribonuclease H (RNaseH) upon cleavage. The *env* gene encodes the anchor structural precursor Env, gp160, which is cleaved into the surface (gp120) and trans-membrane (gp41) glycoproteins. The Gag-Pol mRNAs are spliced to encode the regulatory (Rev, Tat) and accessory (Nef, Vpr, Vpu and Vif) proteins. The diagram was drawn based on information from the Los Alamos HIV sequence database (http://hiv-web.lanl.gov).

1.1.1. Structural proteins

Gag

Gag is one of the main HIV-1 proteins that this thesis focuses on. It is a 55 kD protein that forms the building block of the virus particle core and is composed of three main

polypeptides, namely the matrix protein (MA), capsid protein (CA), nucleocapsid protein (NC) and three smaller peptides, the spacer proteins 1 and 2 (SP1 and SP2) and protein 6 (p6, Briggs *et al*., 2004; Henderson *et al*., 1992; Mervis *et al*., 1988), that are produced by the activity of the viral protease enzyme (Swanstrom and Wills, 1997; Vogt, 1997). These proteins function together to coordinate membrane binding and Gag-Gag lattice interaction in immature virions (Ganser-Pornillos *et al*., 2008; Scarlata *et al*., 2003). The MA that is derived from the myristylated N-terminal of p55 encompasses p17, and forms the scaffold between the core and the outer envelope of the capsid. This region of Gag also contains a nuclear localization domain that facilitates the nuclear transport of the viral genome and allows HIV to infect non-dividing cells (Fassati, 2006; Qi *et al*., 2008).

The CA encompasses the p24 region of Gag and forms the conical core of viral particles (Gottlinger *et al*., 2001), and is important for proper folding of the HIV virion particle (Yoo *et al*., 1997) and also following entry into a new target cell (Gottlinger *et al*., 2001). It is the region of Gag that interacts with the cellular protein Cyclophilin A (CyPA) and mediates its incorporation into HIV virions (Dorfman and Gottlinger, 1996; Lopez *et al*., 2011). Cyclophilin A is a cytosolic human protein involved in protein folding (Zydosky *et al*., 1992), and previous studies have demonstrated that decreased levels of virion-associated CyPA result in reduced virion infectivity (Thali *et al*., 1994) and decreased HIV replication rates in T-cell lines that lack CyPA (Braaten and Luban, 2001). The CA protein is also involved in the incorporation of the Gag-Pol precursor, and therefore enzymes into the capsid (Bukovsky and Gottlinger, 1996; Huang and Martin, 1997; Srinivasakumar *et al*., 1995; Wills and Craven, 1991).

The NC, also referred to as the p7 protein, Sp1, Sp2 and p6 encompass the p15 region of the Gag protein. The nucleocapsid protein is important for transmission and pathogenesis of the virus. It is responsible for specifically recognizing the packaging signal of the genome, and introducing two copies of the viral genome into progeny virions (Heath *et al*., 2003; Dussupt *et al*., 2011). This protein is highly basic and has a non-specific nucleic acid binding ability that is utilized during annealing reactions in the virus life cycle (Berkowitz *et al*., 1996). The p6 protein constitutes one of the most variable regions of Gag, situated at its C-terminal domain. It has many important functions in the

viral life cycle. These include binding to Tsg101 and ALIX, cellular factors important for viral budding from infected cells (Doherty *et al*., 2005). In addition, p6 is a substrate for phosphorylation, ubiquitination and sumoylation, thereby mediating the incorporation of the Vpr protein into the viral particles (Paxton *et al*., 1993; von Schwedler *et al*., 2003; Votteler *et al*., 2011). Spacer protein 1 (Sp1) is located between the CA and NC, and spacer protein 2 (Sp2) is located between NC and p6 (Bryant *et al*., 2011). Sp1 plays a major role in Gag multimerization, higher order organization of the Gag lattice and virus particle assembly (Datta *et al*., 2011; Liang *et al*., 2002). Sp2 plays similar roles and cleavage between p6 and Sp2 has been shown to be critical for the formation of fully infectious particles (Yu *et al*., 1995). Furthermore, a recent study showed that cleavage of the Sp2-p6 site is necessary for genome integration by the infecting virus (Coren *et al*., 2007). Overall, these cleavage processes are important for the maturation of the virion (Vogt, 1996; Wiegers *et al*., 1998).

The p24 region is relatively conserved compared to the p17 and p15 regions of Gag, and targeting of epitopes within the p24 region of the HIV Gag protein has been shown to be associated with viral control (Zuniga *et al*., 2006). Overall, CD8+ T-cell escape mutations in the Gag protein, particularly p24 epitopes, have been associated with a fitness cost to the virus (Ammaranond *et al*., 2011; Martinez-Picado *et al*., 2006; Troyer *et al*., 2009), suggesting that Gag is a good target for inclusion in a vaccine.

Pol

The *pol* gene encodes four different enzymes that are involved in different processes of HIV replication. The enzymes are produced as a Gag-Pol precursor protein, that is further processed by the viral protease enzyme into the individual proteins. Reverse transcriptase (RT) is one of the main enzymes encoded by *pol* and is unique to retroviruses; reverse transcribing the single-stranded HIV RNA (ssRNA) to DNA (Kuiken *et al*., 2000). Another enzyme coded for by the *pol* gene is RNase-H, which breaks down the retroviral genome after infection of a cell; integrase, which integrates the DNA copy of the HIV genome into the host DNA (Craigie, 2001; Wang *et al*., 2001); and protease, which processes multiple polyproteins of HIV so that they can become part of fully-functional HIV proteins and viral particles (Short *et al*., 2000).

Env

The envelope protein is cleaved into two main proteins, a non-covalent complex of the external gp120 and a transmembrane glycoprotein, gp41. These proteins are synthesized as a precursor protein gp160 that matures through various steps, including disulphide bond formation (Coffin *et al*., 1997), extensive glycosylation, cleavage by membrane-associated proteases such as furin (Oliva *et al*., 2003), transport to the cell surfaces as a non-covalent membrane complex (gp120) and finally incorporation into the budding virus (Adamson *et al*., 2007). The mature gp120-gp41 proteins interact together by means of non-covalent interactions, with gp41 transversing through the lipid bilayer of membranes and gp120 on the cell surface. Gp120 is composed of five conserved domains, C1-C5, and five variable domains, V1-V5, and gp41 is composed of seven transmembrane domains. The gp120 glycoprotein, through its V3 loop, is responsible for cell tropism for either T-lymphocytes or primary macrophages (Landau *et al*., 1988), and contains the binding site for the CD4 receptor, interacts with CXCR4 and CCR5 chemokine receptors (Hwang *et al*., 1991) and DC-SIGN on dendritic cells, which facilitates mucosal transmission through transportation of HIV to lymphoid tissue (Deng *et al*., 1996; Geijtenbeek et al., 2000). It is also the primary target for neutralizing antibodies (Feng *et al*., 1996). The gp41 protein contains an N-terminal fusogenic domain, binds to the MA protein and mediates viral and cellular membrane fusion (Camerini and Seed, 1990).

1.1.2. Regulatory proteins

These proteins regulate the transcriptional and posttranscriptional steps involved in viral gene expression, and have all been shown to be important for viral propagation.

Tat

HIV Tat (trans-activator) protein acts as a transcriptional regulator of viral gene expression that function early in viral infection. Tat binds to the trans-activating responsive sequence (TAR) RNA element located at the 5’ end of the HIV RNA genome and initiates viral transcription and elongation from the LTR promoter by binding to P-TEFb (positive-acting transcription elongation factor b), a general human elongation

factor and cofactor that prevents polymerase arrest resulting in the production of full-length HIV-1 transcripts (He *et al*., 2011). This accelerates the production of HIV viral particles (Roy *et al*., 1990) and up-regulates the expression of all viral genes, promoting the elongation phase of HIV-1 transcription and allowing full-length transcripts to be produced (Feinberg *et al*., 1991; Kao *et al*., 1987). Absence of Tat results in premature termination of HIV-1 transcription generating short transcripts (Jones, 1997).

Rev

HIV Rev is produced from fully spliced mRNA and functions early in HIV infection. Rev binds to the rev responsive element (RRE; Malim *et al*., 1989) and regulates unspliced and incompletely spliced viral RNA nuclear export to the cytoplasm (Hope, 1999). Rev has also been shown to up-regulate the expression of structural proteins (Gag, Pol and Env) while down regulating itself and Tat (Douglas *et al*., 1997), through down-regulation of mRNA splicing (since it is produced from fully spliced mRNA). HIV Rev also facilitates the transition from the early to late phase of HIV gene expression (Kim *et al*., 1989).

1.1.3. Accessory or auxiliary proteins

These proteins are not necessary for HIV propagation in tissue culture but they have been conserved in different lentiviruses and experimental observation in *in vitro* studies suggests that they are important *in vivo* as they have been found to influence viral replication and the course of disease progression (reviewed in Kirchhoff *et al*., 2010). Overall, accessory proteins enable viral persistence, replication, dissemination and transmission through preventing immune-mediated or intrinsic antiviral resistance by modifying the infected host cell environment (Malim and Emerman, 2008).

Nef

Nef is a small (27kD) multifunctional protein that is predominantly cytoplasmic but through myristol residues and basic amino acids at the N-terminal of the protein, also associates with the cell membrane (Welker *et al*., 1998; Yu and Felsted, 1992). Studies have shown that this association of Nef with the membrane is critical for most of its

major functions (Alexander *et al*., 2003; Walk *et al*., 2001). Studies of HIV-1 long-term non-progressors have identified viruses with defective Nef, suggesting that this protein is essential for the efficient spread of the virus and disease progression (Brambilla *et al*., 1999; Deacon *et al*., 1995; Rhodes *et al*., 2000; Salvi *et al*., 1998; Tolstrup *et al*., 2006). Nef down-regulates MHC class I molecules (Schwartz *et al*., 1996; Toussaint *et al*., 2008), and this function of Nef protects infected cells from recognition and killing by cytotoxic T-lymphocytes (Schwartz *et al*., 1996). Nef has the ability to interfere with many cellular pathways, with the ultimate consequence of favoring virus persistence in infected individuals. Some of these functions include down-regulation of a host of receptors from the cell surface, including CD4, CD28 and MHC class II-associated invariant chain (reviewed in Lindwasser *et al*., 2007; Kirchhoff *et al*., 2008). In addition, Nef has been shown to down-regulate CCR5 and CXCR4, which is thought to protect cells from super-infection (Michel *et al*., 2005; 2006). Structurally, Nef has an overall conserved central region that is highly immunogenic and that is flanked by variable regions that are less immunogenic (Addo *et al*., 2003).

Vif

Virion infectivity factor (Vif) is a cytoplasmic protein that functions during viral particle assembly and inhibits the incorporation of cellular antiviral agents into viral particles. For example, Vif interacts with the cellular restriction factor APOBEC3G, leading to degradation of the protein, mediated through polyubiquitination of APOBEC3G, thus directing it to degradation by proteases (Opi *et al*., 2007). APOBEC3G is an intracellular antiviral protein that results in defective virions by deaminating the minus-strand of the viral genome, introducing numerous G to A mutations (Fan *et al*., 2010; Yu *et al*., 2004).

Vpr

Viral protein R (Vpr) accelerates the production of HIV proteins by promoting cell-cycle arrest at the G2 phase of cell cycle through activation of p34cdc2/cyclin B complex, a regulator of the cell cycle important for cell entry into mitosis (Ho *et al*., 1995; Jowett *et al*., 1995), thereby inhibiting cell division by mitosis (Planelles *et al*., 1995). The incorporation of Vpr into virions is mediated through interaction with the p6 C-terminal

of p55 Gag (Bachand *et al*., 1999). Vpr also facilitates the formation of the pre-integration complex that is required for the HIV genome to be integrated into the cellular genome (Miller & Sarver, 1997; Popov *et al*., 1998). Vpr can also allow infection of non-dividing tissue macrophages, notably terminally differentiated macrophages (Vodicka *et al*., 1998) through interaction with the nuclear transport pathway (Eckstein et al., 2001; Vodicka *et al*., 1998). This interaction leads to nuclear importation of HIV-1 pre-integration complex (PIC) and the presence of Vpr in the cytoplasm for incorporation into virions, thus enhancing viral spread in these cells (Sherman *et al*., 2003). Non-dividing cells constitute important reservoirs of the virus in infected individuals (Hockett *et al*., 1999; Igarashi *et al*., 2001).

Vpu

Vpu (viral protein U) is a transmembrane protein unique to certain lentiviruses namely HIV-1, its chimpanzee and gorilla precursors (Gao *et al*., 1999), and SIVgsm, SIVmus, SIVmon and SIVden (Barklow *et al*., 2003; Courgnand *et al*., 2003; Dazza *et al*., 2005). Vpu functions in promoting CD4 degradation in the endoplasmic reticulum (ER) as well CD4 down-regulation from the cell surface (Magadan and Bonifacino, 2012), and helps virion release from the plasma membrane of HIV-1 infected cells (Deora & Ratner, 2001). Down-regulation and degradation of CD4 reduces the likelihood of a cell becoming super-infected. Vpu counteracts the cellular restriction factor Tetherin (BST-2), which prevents budding of viral particles, by binding it and directing it to proteasomal degradation by β-TrCP2 (beta-transducin-repeat-containing protein), a human ubiquitin ligase that degrades proteins, (Mangeat *et al*., 2009; Dube *et al*., 2010).

1.2. HIV genetic diversity and distribution

One of the major characteristics of HIV relevant to vaccine immunogen design is its extensive and constantly changing genetic diversity, that results in different variants that can be categorized into distinct genetic subtypes as well as recombinant forms. These subtypes and recombinant forms are unevenly distributed throughout the world (Hemelaar *et al*., 2006, 2011). HIV transmission, pathogenesis and the current

immunogens being designed to be included in vaccines can be further understood in the context of the genetic diversity the virus exhibits and generates during infection.

1.2.1. HIV classification

HIV-1 is classified into four main groups, which are groups M (the main group), O (the outlier group), N (non-M nor O group, Ayouba *et al*., 2000; Gurtler *et al*., 1994; Rayfield *et al*., 1996; Simon *et al*., 1996) and the recently identified group P (Plantier *et al*., 2009; Vallari *et al*., 2011). It is HIV-1 group M that is responsible for the majority of HIV infections throughout the world, and it is further subdivided into ten phylogenetic subtypes designated by the letters from A-K (Figure 1.2). These subtypes are approximately genetically equidistant from each other, except for subtypes B and D, that are more closely related to each other than to other subtypes (Gao *et al*., 1996). The subtypes are classified based on phylogenetic clustering of sub-genomic regions such as *gag*, *nef*, *pol* and *env* of individual genes from the same isolates. In addition, advances in sequencing technology have enabled the use of multiple sub-genomic regions and near full-length genome sequence analysis for subtyping and classification of HIV-1 variants (Louwagie *et al*., 1993; Triques *et al*., 2000). These studies highlight the degree to which recombination contributes to HIV-1 subtypes diversity, and highlight the limitations of using a single method or gene for classification, since some viruses are the result of a mosaic of genes from different viruses (Gao *et al*., 1998).

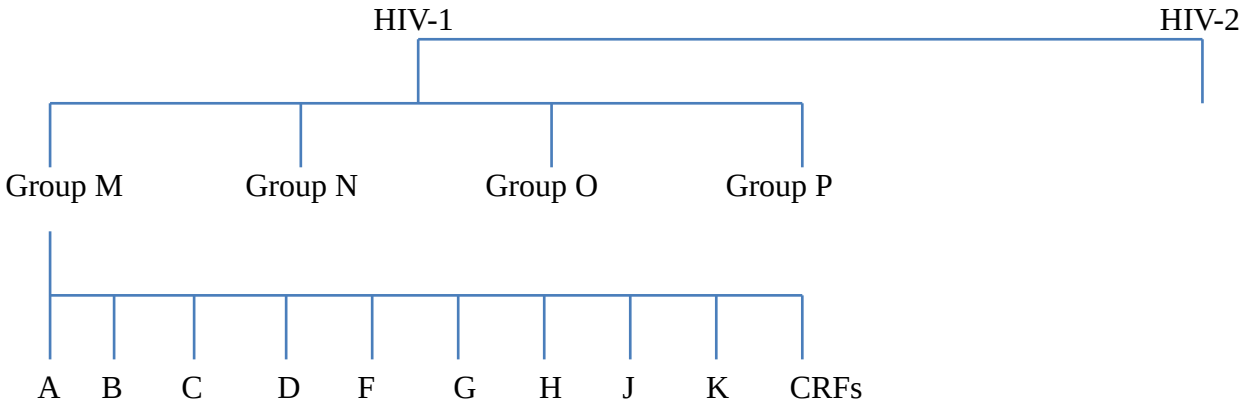


Figure 1.2. The different levels of HIV classification. More than 90% of HIV-1 infections worldwide belong to HIV-1 group M. Group O is found in west-central Africa, group N was discovered in 1998 in Cameroon and is exceptionally rare, and Group P was discovered in 2009 and has been described in only two individuals from Cameroon.

The inter-subtype genetic distance within group M is 15% for the *gag* gene and as high as 25% for the *env* gene (Janssens *et al*., 1994; Robertson *et al*., 2000). Sub-subtypes have been identified, which are formed when viruses from the same subtypes recombine, and appear to be more genetically related to each other when compared to other subtypes (Robertson *et al*., 2000). Examples of sub-subtypes are A1 and A2 and F1 to F2 from subtypes A and F, respectively (Gao *et al*., 2001; Triques *et al*., 2000).

There are also viral isolates that are a product of genetic recombination between two distinct subtypes (Burke *et al*., 1997) and are thought to originate from people multiply infected with viruses of two or more subtypes. Those recombinant forms that can be identified by full-length genome sequencing in at least three epidemiologically unlinked individuals have been designated circulating recombinant forms (CRFs; McCutchan, 2006; Peeters, 2001; Robertson *et al*., 2000). At least 49 CRFs have been identified thus far (http://www.hiv.lanl.gov/), and they are identified by numbers in ascending order according to discovery, followed by letters of the parental subtypes. If only a single sequence is available, they are designated unique recombinant forms (Robertson *et al*., 2000; Ragupathy *et al*., 2011). These recombinants are common in regions with

multiclade epidemics, such as countries in West central Africa. CRF02_AG and CRF01_AE are amongst the commonly found CRFs.

This extensive heterogeneity of HIV-1 is driven by different mechanisms that include the high error rate of reverse transcriptase enzyme during transcription (Op de Coul *et al*., 2001; Roberts *et al*., 1988), rapid turnover of HIV-1 *in vivo* (Ho *et al*., 1995), host selective immune pressure (Michael, 1999) and recombination events during replication (Temin, 1993).

The high error prone reverse transcriptase enzyme that results in the incorporation of incorrect nucleotides during transcription results in rapid HIV evolution (Op de Coul *et al*., 2001; Roberts *et al*., 1988). A huge amount of genetic variation is generated that is fueled by the rapid replication rates of the virus in CD4+ T lymphocytes (Ho *et al*., 1995). There are studies that have shown that rapid HIV evolution correlates with rapid HIV disease progression (Mullins and Jensen, 2006).
Another mechanism that generates the much genetic variation in HIV is the high rate of recombination. This can be on average three events per genome per round of replication (Hu and Temin, 1990a and 1990b). Coupled to the rapid replication rate of the virus, a huge amount of genetic variation is generated between and within HIV subtypes (Schroeder *et al*., 2005; Charpentier *et al*., 2006).

HIV has been shown to evolve under immune selection pressure, with a vast amount of data generated suggesting that certain regions of HIV targeted by CD8+ T cells mutate in an effort to escape recognition by the immune system (Price *et al*., 1997; Borrow *et al*., 1997). However, as some of these mutations may be in regions that have structural roles for the virus, this can result in a fitness cost that results in a reduced replicative capacity (Gesprasert *et al*., 2010; Martinez-Picado *et al*., 2006; Novitsky *et al*., 2009; Troyer *et al*., 2009). For example, the extensively studied dominant Gag TW10 escape mutation, T242N, has been shown to reduce viral replicative capacity (Martinez-Picado *et al*., 2006). This issue is further discussed in section 1.4.

1.2.2. Geographic distribution of HIV-1 subtypes

Except in Africa, where almost all HIV subtypes are found, other regions in the world show a clear, distinct pattern of HIV subtype distribution (Hemelaar *et al*., 2006; 2011; Osmanov *et al*., 2002). The migration of people from one place to another or a prevalent route of transmission favoring a certain subtype in that particular population could have resulted in this distribution of HIV subtypes (Buonaguro *et al*., 1995; McCutchan *et al*., 1992a and 1992b; Myers, 1994; Naderi *et al*., 2006; Ou *et al*., 1992). However, this pattern is being affected by increases in global travel (Perrin *et al*., 2003).

Recent studies have identified HIV-1 subtypes A, B and C as the most prevalent HIV-1 subtypes globally, with subtype C accounting for almost 50% of infections worldwide (Hemelaar *et al*., 2011). Subtype A and B account for approximately 12% and 11% of global infections, respectively. HIV-1 subtype A viruses are the predominant subtypes in countries in west, central and eastern Africa (Kenya, Uganda, Tanzania, and Rwanda), and also make-up the eastern European epidemic (Bobkov *et al*., 2004, Figure 1.3). HIV-1 subtype B is the main subtype circulating in western and central Europe, north and south America, and Australia, and is also common in other countries in South East Asia, northern Africa and the Middle East. In South Africa and Russia, subtype B is mainly found among men who have sex with other men (Le Vu *et al*., 2010). HIV-1 subtype C viruses are predominant in those countries with >80% of all global HIV-1 infections, such as southern Africa and India (Hemelaar *et al*., 2006; 2011; Osmanov *et al*., 2002). HIV-1 subtypes G and D each account for 5% and 2% of global HIV-1 infections respectively, and are commonly found in central Africa. HIV-1 subtypes F, H J and K together account for less than 1% of global HIV-1 infections.

CRFs now account for at least 20% of HIV-1 infections worldwide, and this fact has increased their recognition as being relevant in the global HIV-1 pandemic (Hemelaar *et al*., 2011; Osmanov *et al*., 2002). CRF01_AE is the predominant local form in South East Asia (Motomura *et al*., 2000; Piyasirisilp *et al*., 2000), while CRF02_AG is predominantly found in West and West Central Africa (McCutchan *et al*., 1999; Peeters *et al*., 2000), as shown in Figure 1.3. These two recombinant forms account for 13% of global infections with HIV-1 recombinants. In regions where these recombinants

circulate, some of them dominate the epidemic; for example, in Cameroon, CRF02_AG is the predominant virus accounting for more infections than its parental subtypes (subtypes A and G) found in the that country (McCutchan *et al*., 2000; Carr *et al*., 1998; Kijak *et al*., 2004). Overall, although there was a notable increase in the contribution of CRFs to the global HIV-1 epidemic, their global and regional distribution is broadly stable (Hemelaar *et al*., 2011). Unique recombinant forms account for at least 30% of infections in regions where they circulate (Peeters *et al*., 2000; as reviewed in Tebit and Arts, 2011). Since recombination and mutation continue to occur, it is likely that new genetic subtypes and CRFs of HIV-1 will arise in the future. Furthermore, due to global travel, the current subtypes and CRFs will continue to spread to new areas as the global epidemic continues.

In addition to the genetic diversity of HIV-1, the task of designing an effective vaccine is made difficult by the diverse human leukocyte antigen (HLA) genetic backgrounds of individuals in different populations. An important consideration in understanding cellular immune responses to HIV-1 infection is the role played by the HLA molecules, through which CD8+ and CD4+ T-cells recognize viral epitopes.

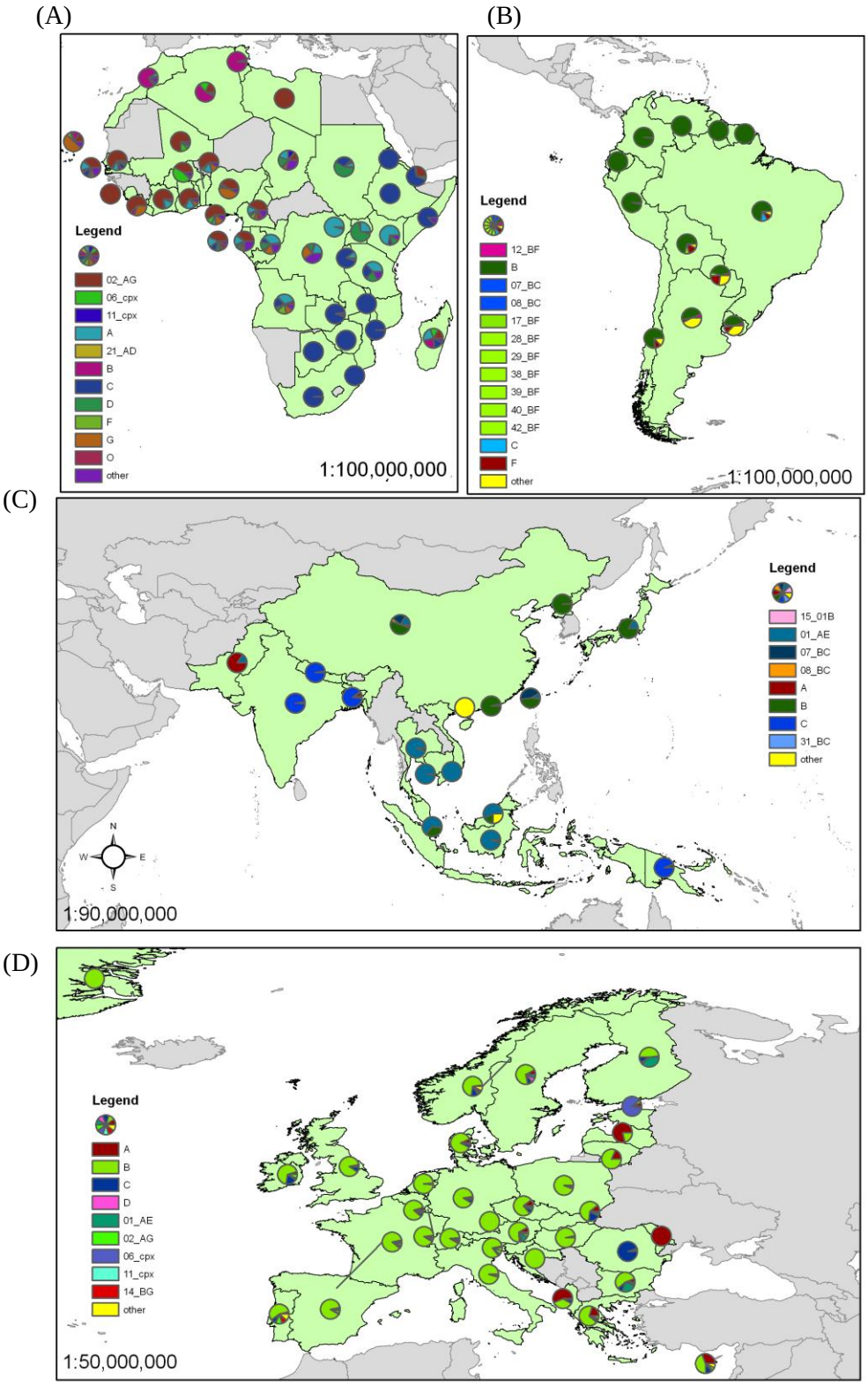


Figure 1.3. Global distribution of HIV-1 subtypes and circulating recombinant forms. (A) Africa. **(B)** South America. **(C)** Asia. **(D)** Europe. Each pie represents the contribution of the particular subtype or recombinant form to the total sequences in the database from the specific country from different regions, with the different subtypes and recombinants being color-coded. The table of names and geographical locations of countries were obtained from the Natural Earth Data website (www.naturalearthdata.com). Sequence data and information on classification of countries into regions were obtained from the HIV geography site of the Los Alamos Sequence database (www.hiv.lanl.gov), as of 31 December 2011. The tables were joined and modified and map drawn using the Arc Geographical Information System (GIS) software version 9.3.1 (Arc GIS, ESRI ArcMapTM, CA).

1.3. HLA genetic diversity, distribution and association with disease progression

HLA genetic diversity and distribution

The HLA genes can be divided into three main loci, namely HLA class I, II and III (Shiina *et al*., 2009). HLA class I molecules are involved in the presentation of immunogenic peptides to CD8+ T-cells, while HLA class II molecules present peptides to CD4+ T-cells (Gebe *et al*., 2002). HLA class I alleles involved in peptide presentation are classified into HLA-A, -B, -C, while HLA class II are classified into HLA-DP, -DQ and –DR. HLA class III molecules are unrelated to peptide presentation. The type of peptide that binds to a particular HLA molecule depends on the HLA anchor residues and the sequence of that particular peptide.

The genes that code for HLA molecules are highly polymorphic, with extensive allelic variation within the HLA class I loci (Lazaro *et al*., 2011a, 2011b and 2011c; Pacheco *et al*., 2010). At least 3249 HLA class I and 1198 HLA class II alleles have been described to date (http://www.ebi.ac.uk/imgt/hla/stats.html). Different alleles show different degrees of diversification, with HLA B molecules being more diverse when compared to other HLA class I alleles (Tian *et al*., 2003). There have been 965 HLA-A, 1 543 HLA-B and 625 HLA-C molecules identified thus far (http://www.ebi.ac.uk/imgt/hla/stats.html). These HLA alleles can differ by as little as one nucleotide substitution within a genomic sequence of 3300 base pairs (Robinson *et al*., 2003a and 2003b; 2009), which may result in different peptide recognition patterns.

HLA allele variation differs between populations, with some populations exhibiting greater diversity compared to others. For example, Brazilians were shown to have a high variation within the HLA-B locus compared to Mexican, Cuban, South African Zulus, Omani and Singapore Chinese, with the Mexicans exhibiting the least diversity of them all (Williams *et al*., 2001). Different populations and ethnicities express different HLA alleles (reviewed in Stephens *et al*., 2005), with certain alleles being dominant or even specific to certain populations. For example, HLA-A*02:02; A*34:02; A*36:01; A*74:01; B*15:03; B*42:01; B*57:03 and B*58:02 were shown to be found only in multiple African populations (Cao *et al*., 2004). Also, an allele that is common to a particular population may be rare in another population. Some alleles are limited to particular ethnic groups, while others are widely spread among ethnically distinct populations. For example, HLA-A*30:02, HLA-B*57:03 and Cw*18:01 are found in high frequencies in the Zambian population, and there is a high frequency of B*39:10; B*42:01,

B*81:01, Cw*81:01, B*15:03 and B*15:10 in Zulus and Xhosas when compared to other populations in South Africa (Honeyborne *et al*., 2006). Even in individuals of the same ethnicity, particular HLA alleles expressed differ among individuals (Du Toit *et al*., 1990). For example, HLA-A*03, B*49 and B*57 have relatively higher frequencies in Ethiopians than in other African populations (Ferrari *et al*., 2004). Overall, HLA-A2 is the most prevalent allele worldwide (Browning and Krausa, 1996). This allele has also been associated with reduced risk of HIV-1 acquisition in two Kenyan cohorts (McDonald *et al*., 2000), and so might be of interest in development of T-cell based HIV vaccines. Altogether, these data showing the differential expression of different HLA alleles in ethnically diverse populations highlight challenges in designing a T-cell-based HIV vaccine immunogen with high global population coverage.

HLA and disease progression

There are strong data to suggest that not all HLA alleles have a similar impact on HIV-1 control. Genome wide association studies (GWAS), an approach involving the identification of genetic variations in human genomes that are associated with any disease outcome, have generated data to suggest that variation attributed to HLA types is associated with differential HIV disease outcome. Several GWAS performed on populations of European ancestry showed that the B*5701 allele is strongly associated with lower viral set point (the viral load at approximately a year after infection) and slower CD4+ T-cell decline (Fellay *et al*., 2007; Limou *et al*., 2009; Dalmasso *et al*., 2008; Fellay *et al*., 2009). In addition, similar analyses in African-Americans identified B*57:03 to be associated with slow disease progression (Pelak *et al*., 2010). These studies are lacking in African populations. Complementing the GWAS studies mentioned above, certain HLA types have been associated with better diseases outcome, including HLA B*57 (Altfeld *et al*., 2003), B*27 (Stephens *et al*., 2005), B*58:01 (Gao *et al*., 2001) and B63 (Frahm *et al*., 2005). On the other hand, HLA B*35, B*58:02 and Cw*04 have been shown to negatively affect the outcome of disease (Walker and Korber, 2001). Studies have also shown that variations at the B*35Px group have different associations with disease outcome. HLA B*35:03 is associated with better disease outcome (Table 1.1), whilst B*35:01 does not significantly differ from the population average in terms of disease outcome (Kiepiela *et al*., 2004; Kawashima *et al*., 2009). Similarly, these data extend to macaque models of SIV infection. For example, association between MHC class I genotypes *Mamu-A*01*, *Mamu-A*02*, *Mamu-B*08* and *Mamu-B*17* with enhanced SIV control has been demonstrated (Goulder and Watkins, 2004; Muhl *et al*., 2002; Mothe *et al*., 2003; Yant *et al*., 2006; Loffredo *et al*., 2007),

while *Mamu-B*01* and *Mamu-A*08* have been observed as alleles associated with rapid disease progression (Goulder and Watkins, 2004). The mechanism underlying the ability of certain HLA types to be associated with better disease outcome was not fully understood. One possible mechanism is that CD8+ T-cells restricted by these protective HLA alleles are able to suppress regulatory T cells, which themselves would have suppressed the proliferation of these HIV-specific T-cells (Elahi *et al*., 2011). This in part provides an explanation of why HIV-specific T-cells restricted by these protective HLA alleles have a higher proliferative capacity compared to those restricted by non-protective alleles. However, the ability of these HLA alleles to perform this function, where other non-protective alleles do not, is not fully understood, and needs further research. It is likely that the nature of regions being targeted by these HLA alleles and the timing of responses are among the possible explanations. Therefore, specificity of the epitopes targeted by these CD8+ T-cells may also play an important role in their protective nature, since studies have shown that these alleles can present conserved viral epitopes, for example HLA-B*27 has been shown to restrict a conserved epitope in Gag that is associated with viral control (Goulder *et al*., 1997), since mutation of this region in HIV Gag may result in a reduced replication capacity of the virus.

A separate issue with HLA and disease progression is homozygosity. Individuals homozygous for HLA class I alleles appear to have limited breadth of the CD8+ T-cell responses, and this may negatively impact disease outcome and progression (Carrington *et al*., 1999; Huang *et al*., 2009; Qing *et al*., 2006).

These data suggest that different HLA types may have differential contribution to immune control of HIV, which makes it sound to construe that HLA diversity and distribution have to be taken into consideration in the design of T-cell-generating vaccines, if a high degree of population coverage is to be attained. Previous studies have shown differential contribution by different HLA molecules to the total anti-HIV cellular immune response (Kiepiela *et al*., 2004). This study suggested a dominant role of HLA-B molecules in shaping the co-evolution of HIV-1 and HLA molecules, through presentation of more epitopes when compared to HLA-A and HLA-C. Interestingly, several studies have shown that certain HLA molecules can present similar or identical viral epitopes despite their polymorphisms, and these have been grouped into genetic families referred to as HLA ‘supertypes’, based on analysis and subsequent clustering of their peptide binding repertoire (Doytchinova *et al*., 2004; Kangueane *et al*., 2005; Sette and

Sydney, 1999). Therefore, knowledge of the distribution of HLA alleles within local populations is essential for the development of peptide-based vaccines with large population coverage.

Table 1.1 . Table of HLA alleles commonly associated with different disease outcomes.

HLA allele	Supporting evidence	Reference
Protective alleles		
B*57	Allele is over-represented in elite controllers, B57 CD8+ T-cells are immunodominant than other CD8+ T-cells	Bailey *et al*., 2006; Jansen *et al*., 2005; Altfeld *et al*., 2003; 2006.
B*5801 and B*63	Have a very similar binding motif to B57 and associated viral control	Carrington *et al*., 2003; Frahm *et al*., 2005.
B*27	Present a conserved Gag epitope and is associated with long-term non-progression and escape in this epitope predicts viral load outcome.	Goulder *et al*., 1997; Carrington *et al*., 2003; Ammaranond *et al*., 2011.
Non protective alleles		
B*35, B*58:02, Cw*04 and B*53	Associated with rapid progression to AIDS. B35 subtypes have different associations: B*35:03 is associated with rapid progression while B*35:01 has a slightly different binding motif to B*35:03 but disease outcome not significantly different from population average.	Carrington *et al*., 2003, Gao *et al*., 2001; Walker and Korber, 2001.

1.4. HIV escapes from immune pressure

It has been established that HIV-1 can escape both the humoral and cell-mediated arms of the immune response. HIV-1 Env is the target for neutralization by the antibody response, but research has demonstrated that HIV-1 can replicate in the face of a strong antibody responses. This is because HIV-1 has evolved mechanisms to escape neutralization. These include glycosylation of regions that encompass epitopes that are targets for neutralization (Johnson *et al*., 2002; Sheehy *et al*., 2002), as well as mutation of regions within the variable loops of the gp120 without a fitness cost to the virus, that result in abrogation of neutralization (Johnson *et al*., 2002). Addressing these issues is extremely important in the light of renewed major interest in developing vaccines capable of generating neutralizing antibodies. Because the HIV vaccine field has largely, in the last ten years, focused its efforts on developing T-cell generating vaccines, much work has focused on escape from CD8+ T-cell-mediated immune pressure.

1.41. HIV escape from CD8+ T-cell pressure and implications for vaccine design

As described in a previous section, CD8+ T-cells recognize epitopes presented to them in the form of an HLA class I-peptide complex. The amino acid sequence of the peptide determines the allele that restricts that peptide. Different HLA alleles demonstrate differential tolerance to amino acid variation within the peptides. However, mutational escape which is a major challenge facing CD8+ T-cell-mediated recognition of HIV can be predicted based on the individual HLA allele background (Brumme *et al*., 2007; Brumme *et al*., 2009; Heckerman *et al*., 2010; Leslie *et al*., 2004; Yang *et al*., 2010).

HIV can escape immune recognition by CD8+ T-cells by various mechanisms. Firstly, since there are various steps that occur during MHC class I peptide processing, mutations in the viral protein may result in abrogation of the peptide processing steps and subsequently loss of recognition by CD8+ T-cells (Koup *et al*., 1994). For example, mutations in the regions that flank the epitope may interfere with peptide processing if the residues are crucial for the processing steps (Milicic *et al*., 2005; Draenert *et al*., 2004). For example, the A69K, A81G and H87R mutations flanking the HLA-B*35-restricted VY8 Nef epitope were shown to alter processing of the epitope by the proteasome (Milicic *et al*., 2005). In addition, the A146P mutation flanking the IW9 HLA-B*57-restricted Gag epitope were shown to prevent the amino terminal trimming of the optimal epitope by the ER amino peptidase I enzyme (Draenert *et al*., 2004).

Secondly, escape can be a result of inhibiting or reducing HLA-peptide binding and presentation to the T cell receptor (TCR). Mutations in the antigenic peptide may result in conformational changes that abrogate binding to the HLA allele, thereby abolishing presentation of the epitope (McAdam *et al*., 1995). An example is a well-characterized epitope in p24 restricted by HLA-B*27 (KRWIILGLNK), in which an R→K or R→G at the second position of the epitope was found to abrogate binding to HLA-B*27 molecule (Goulder *et al*., 1997). Lastly, mutational escape may involve mutations that interfere, reduce or antagonize the TCR-peptide interaction. An example of this is the rhesus macaque *Mamu-A*02*-restricted SIV Nef epitope (199RY), where a variant bound the MHC allele as the wild type sequence, but had reduced affinity for the TCR, resulting in escape of the epitope from recognition (Cale *et al*., 2011).

Mutational escape may therefore result in the virus being able to replicate without being recognized by specific T-cells, or without immune pressure in those viral regions. However, escape can in some instances be disadvantageous to the virus, in that if the targeted regions reside in those regions of the viral protein that have a functional or structural role for the virus, a mutation may result in a fitness cost or a replication-defective virus (Liu *et al*., 2007; Brockman *et al*., 2007; Schneidewind *et al*., 2007; Troyer *et al*., 2009).

For a T cell-based HIV vaccine, targeting such immunologically vulnerable regions may be advantageous for controlling virus replication and hence the course of disease progression, and subsequently reducing transmission rates. Therefore, the design of a vaccine immunogen can exploit knowledge of mutations that result in a fitness cost to the virus, such that vaccine induced immune responses may direct HIV evolution towards a less fit virus. However, this is also not a foolproof strategy; the observation of mutations either upstream or downstream of epitopes that have been shown to have a fitness cost to the virus if mutated has raised concerns about this feasibility (Crawford *et al*., 2007; Gonzalez-Ortega *et al*., 2011; Navis et al., 2007; Verheyen *et al*., 2006). Such 'compensatory' mutations have been shown to partially or fully restore the virus fitness after mutating certain epitopes with a fitness role for the virus (Crawford *et al*., 2007; Gonzalez-Ortega *et al*., 2011). Also, a major concern regarding this may be that escape mutants may accumulate at the population level, leading to a loss of recognized epitopes and vaccine-induced T cell responses with reduced effectiveness over time. Indeed, HIV has been shown to adapt at the population level in the face of immune pressure (Bunnik *et al*., 2010; Kawashima *et al*., 2009; Moore *et al*., 2002; Pond *et al*., 2006). In addition, much of the work on mutational escape has relied on epitopes from the p24 conserved region of HIV that are restricted by particular HLA alleles, and it is not known whether there are many other HIV regions outside this well characterized protein that may be suitable as vaccine immunogens, as well as their relevance in populations lacking these HLA alleles. Also, epitopic regions that undergo reversion to their wild type sequence when the virus is transmitted to HLA mismatched host (as the epitope will no longer be under immune pressure) might have to be taken into consideration in the design of a T cell-based vaccine in populations of diverse HLA backgrounds (Davenport *et al*., 2008; Gesprasert *et al*., 2010; Kearney *et al*., 2009), since these epitopes become lost as the epidemic progresses. However, it could be argued that these reverting epitopes may be less important for viral fitness.

1.5. HIV-specific T-cell immunity

Although the immune correlates of protection in HIV infection are not fully understood, numerous studies suggest that the induction of HIV-specific cellular immune responses is important for an effective vaccine against HIV. Therefore, much of the efforts of the last ten years have focused on developing a vaccine capable of generating T-cell responses. There is a wealth of evidence to show that HIV-specific CD8+ T-cells can control viraemia during acute HIV-1 infection (Borrow *et al*., 1994; Koup *et al*., 1994). For that reason, identification of HIV regions repeatedly targeted by the cellular immune response and associated with viral control will assist in choosing immunogens for vaccine design. In addition, elucidating the functional and phenotypic nature of HIV-specific immune responses will assist in choosing which vectors, adjuvants and immunisation protocols are required to generate effective immune responses against the virus.

For more than 20 years, researchers have focused on different aspects of the immune response to HIV in an attempt to identify the immune correlates of HIV control. Such studies have looked at the specificity of the response to HIV, for example which regions of HIV are frequently targeted, and which regions may be associated with viral control, as well as the magnitude and breadth of HIV-specific responses in infected individuals. Research has also focused on the functional attributes of HIV-specific T cells, including the range of cytokines produced by these cells, their cytotoxic potential, proliferative capacity, memory differentiation status, and CD8 or CD4 phenotype. While earlier studies focused on the specificity and quantity (breadth and magnitude) of HIV-specific T-cell responses and their association with markers of disease progression such as CD4 count and plasma viral load, more recent studies have studied the quality of these responses (their functional nature) in trying to identify the features associated with viral control.

HIV-infected individuals show broad recognition of HIV proteins and high magnitude responses (Addo *et al*., 2003; Masemola *et al*., 2004). These kinds of studies typically employed whole proteome screening of HIV-specific T-cell responses by the IFN-g ELISpot assay, and performed them in cross-sectional cohorts of HIV-infected individuals infected for unknown periods of time. These studies on the total magnitude of HIV-specific T-cells as well as their breadth, indicating the number of epitopes targeted, have generated data to conclude that the overall magnitude of HIV-specific T-cells is not associated with viral control (measured by plasma viral load) or disease progression, using CD4 count or viral load as a surrogate (Addo *et*

al., 2003; Betts *et al*., 2001; Buseyne *et al*., 2002; Edwards *et al*., 2002; Gray *et al*., 2009; Masemola *et al*., 2004; Peretz *et al*., 2005). Several longitudinal studies have also failed to make a connection between the overall magnitude and breadth of responses. For example, no association was observed between the magnitudes of each expressed HIV protein with disease progression or viral set point at 12 months in a subtype C-infected cohort (Gray *et al*., 2009), and no correlation was obtained between viral load and rate of CD4 T-cell decline as when the magnitude or breadth of T-cell responses were compared between typical, fast and slow progressors (Peretz *et al*., 2005). These studies suggested that whilst HIV-infected individuals make robust CD8+ IFN-g responses to the virus, the overall breadth and magnitude of these responses was not a correlate of immune control, and perhaps the specificity or quality of the response may be important for viral control.

Masemola *et al*. (2004) showed that the hierarchical targeting of the HIV Gag protein, rather than the overall magnitude of T-cell responses to HIV proteins, was associated with viral control. This result was confirmed by other independent studies that showed that targeting of the HIV Gag protein was beneficial for viral control (Edwards *et al*., 2002; Kiepiela *et al*., 2007; Mori *et al*., 2011). In one of these studies, specifically it was the magnitude and breadth of the Gag p24 protein that inversely correlated with slower disease progression (Mori *et al*., 2011). In other studies, targeting of other proteins such as Nef and Env did not correlate with viral load or CD4 count (Edwards *et al*., 2002). Interestingly, one of these studies demonstrated that while targeting of Gag was associated with lower plasma viral loads, targeting of Env was associated with higher plasma viral loads (Kiepiela *et al*., 2007). The importance of the specificity of T-cells in viral control has also been further supported by many other studies concluding that targeting of Gag is associated with lower viral loads, higher CD4 counts or slower disease progression (Zuniga *et al*., 2006; Geldmacher *et al*., 2007, Pereyra *et al*., 2008).

Although studies on functional avidity of HIV-specific T-cells have not been conclusive (Miura *et al*., 2009; Migueles *et al*., 2003; Draenert *et al*., 2004), the avidity of HIV-specific T-cells targeting specific HIV epitopes that was associated with control of virus replication (Almeida *et al*., 2007; Miura *et al*., 2009). Taken together, when data on HIV-specific CD8+ T-cells of LTNPs or elite controllers are compared to that progressors no consistent significant differences have been observed in terms of specificity, breadth or magnitude (as reviewed in Migueles *et al*., 2004 and Migueles *et al*., 2006; Goulder *et al*., 1996; Migueles *et al*., 2000; Betts *et al*., 2001).

Having realized that most of these studies looking at the specificity, magnitude and breadth of HIV-specific T-cell responses relied on quantifying response using one immunological readout, namely the cytokine IFN-γ, some researchers sought to exploit advances in flow cytometry and looked at the quality rather than quantity of virus-specific CD8+ T-cell responses. Apart from IFN-γ production, HIV-specific T-cells are capable of producing many other cytokines (*e.g.* IL-2 and TNF-α), chemokines (*e.g.* MIP-1β), as well as performing other cytotoxic functions, where the up-regulation of perforin and granzymes can be measured, or functions such as degranulation (CD107a expression) have been used as a surrogate of cytotoxic potential (De Rosa *et al*., 2004; Roederer *et al*., 2004; Betts *et al*., 2006; Kannanganat *et al*., 2007a and 2007b; Hersperger *et al*., 2010). The proliferative capability of HIV-specific T cells can also be measured (Migueles *et al*., 2002; Lichterfeld *et al*., 2004; Horton *et al*., 2006). Initially, data generated in studies characterizing the quality of HIV-specific T-cells suggested that it is the polyfunctionality, *i.e.* the ability of cells to produce more than three 'functions/cytokines' simultaneously, that may be important in viral control (De Rosa *et al*., 2004; Roederer *et al*., 2004 Betts *et al*., 2006; Kannanganat *et al*., 2007). An additional study observed that HIV-specific CD8+ T-cells from long term nonprogressors (LTNPs) or elite controllers were able to produce perforin immediately after stimulation with HIV peptides, which suggests that this might be one of the mechanisms by which this small group of individuals effectively control HIV during natural infection, as this was not observed in progressors (Hersperger *et al*., 2010). Further dissection and characterization of the different combinations of functions of these HIV-specific T-cells, generated data to suggest that highly functional HIV-specific T-cells are distinguished by perforin and IL-2 up-regulation (Makedonas *et al*., 2010). Furthermore recent study suggested that the ability of HIV-specific T-cells to perform effector functions was associated with their ability to express high levels of the transcription factor T-bet (Hersperger *et al*., 2011).

There is evidence that HIV-specific T-cells from individuals showing better control of the virus have a higher proliferative capacity when compared to those who progress faster (Migueles *et al*., 2002; Litcherfeld *et al*., 2004; Horton *et al*., 2006), and that proliferative capacity of HIV-specific CD8+ T-cells correlate inversely with plasma viral load (Day *et al*., 2007). HLA-B*5701 restricted CD8+ T-cells demonstrated higher proliferation compared to non-HLA-B*5701 restricted T-cells (Horton *et al*., 2006; Altfeld *et al*., 2003). Therefore, study individuals with HLA alleles associated with long term non-progression might be controlling the virus

through targeting those HIV regions which cannot escape CD8+ T cell pressure without fitness cost and generating strong T-cell responses that are highly functional and capable to proliferate. A recent study demonstrated the superiority of the proliferative capacity compared to the magnitude of IFN-γ responses in delaying HIV disease progression, showing that the magnitude of IFN-γ responses did not correlate with delayed disease progression, but rather the proliferative capacity of HIV-specific CD8+ T-cells did (McKinnon *et al*., 2011). In addition, the high magnitude of CD8+ T-cells targeting HIV Gag protein, specifically p17 and p15 with high proliferative capacity, was associated with low viral loads and high CD4 counts (Calarota *et al*., 2008). Therefore, magnitude to specific regions of the virus, function and proliferative capacity of HIV-specific T-cells may be more important than merely the overall breadth and magnitude of HIV-specific CD8+ T-cell responses.

The presentation of antigens to CD8+ T-cells and their activation by antigen presenting cells has been shown to involve CD4+ T-cell help (Xiang *et al*., 2005). CD4+ T-cells play an important role in the clonal expansion and differentiation of CD8+ T-cells (Janssen *et al*., 2003) as well as maintenance of their memory (Jun & Bevan, 2003). In addition, CD4+ T-cell proliferation in the same way as the proliferation of CD8+ T-cells as well IL-2 production has been shown to correlate with HIV-1 clinical status in both humans and macaques (Rosenburg *et al*., 1997; Mckay *et al*., 2003). Therefore, CD4+ T-cell help may be vital for viral control, and a key immune function to generate with a HIV vaccine.

1.6. Cross-reactive immunity to HIV-1

The uneven distribution of HIV-1 subtypes and CRFs has led to the development of some HIV candidate vaccines to match subtypes circulating in specific regions (Ellenberger *et al*., 2003; Jaffray *et al*., 2004). However, whether regional vaccines are necessary or a universal vaccine is possible, is still unknown. Therefore, efforts have also focused on developing broadly cross-reactive vaccines for global use, or for areas in which many clades and recombinants circulate. These different approaches are discussed in detail in the next section. Since vaccine development and clinical trials are lengthy, it is important to characterize the cross-reactive potential of vaccine immunogens that are in different stages of development in HIV-infected populations where different viral subtypes circulate. Although these kinds of studies have their limitations, since those immune responses generated in natural infection might differ from those generated

by a vaccine, they may nevertheless reveal important insights into whether vaccine immunogens may be sufficiently cross-reactive in a given population (with its immunogenetic diversity) with particular HIV-1 subtypes circulating.

The extensive diversity in the HLA gene loci is one of the ways by which the immune system adapts to the recognition of the many variant forms of pathogens (Brennan *et al*., 2002), through promiscuous binding by HLA alleles (Sette and Sydney, 1998), and the involvement of other key players, such as peptide processing and presentation, and the flexibility of the T-cell receptor (Brennan *et al*., 2002; Buseyne and Riviere, 2001; Haanen *et al*., 1999). For that reason, T-cells of an individual are able to recognize an extensive array of epitopes from each viral sequence that can be extended by being heterozygous at HLA loci (Carrington *et al*., 1999).

Immune responses that are cross-reactive between different HIV clades ('cross-clade responses') have been consistently demonstrated in HIV-infected individuals (Aidoo *et al*., 2008; Currier *et al*., 2003; Gillespie *et al*., 2002; Keating *et al*., 2002). Many earlier studies relied on the use of peptide pools or recombinant vaccinia virus constructs expressing selected HIV-1 genes, giving 'bulk' cross-reactivity response readouts of HIV proteins (McAdam *et al*., 1998; Ferrari *et al*., 1997 and Betts *et al*., 1997). Recent studies have assessed cross-clade immune responses at the single peptide level.

In one such important study, cellular immune responses to the three major HIV-1 subtypes (A, B and C) were determined in a cohort of 250 unvaccinated individuals from four countries with HIV-1 subtype A, B and C epidemics (Coplan *et al*., 2005). In this study, Gag, Nef, Pol, Tat and Rev gene products were tested in ELISpot assays, with Gag and Nef contributing most of the responses, which were then further characterized for cross-reactivity. The study defined 'cross-clade reactivity ratios' as the ratio of immune responses against a heterelogous clade (the one not responsible for infection) to that against a homologous clade (the one responsible for infection). The higher the ratio, the greater the immune response an individual could mount against that particular heterologous subtype. A high degree of cross-reactivity was observed, with cross-clade reactivity ratios of 91% and 107% for A versus B (meaning that the reactivity of A was more than that of B, the infecting clade) and C versus B Gag proteins, respectively. Similarly, cross-reactivity rations were high for the Nef protein, with 95% between clades A and B and 96% between A and C for the Nef protein (Coplan *et al*., 2005). Further analysis from the same

cohort extended to 363 HIV-infected individuals (Gupta *et al*., 2006) demonstrated similar extensive cross-reactivity among the three HIV-1 subtypes tested for Gag and Nef. In this larger cohort, cross-reactivity ratios were 99.1% between clade C and clade B Gag proteins and 97.8% between clades A and B Gag proteins.

In a separate study, HIV-specific T-cells from clade B-infected study individuals were assessed for reactivity to clades B and C peptides using the IFN-γ ELISpot assay (Zhao *et al*., 2007). In this study, there were a large number of cross-reactive peptides between clade B and C peptides. This observation demonstrates that HIV-1 clade B infected study individuals can recognize peptides based on clade C sequences. A detailed analysis of reactive peptides in this study showed that cross reactivity was associated with a high degree of homology and lower variability, measured by entropy scores among corresponding peptides from the tested clades (Zhao *et al*., 2007). Further, results from Yu *et al*. (2005) demonstrated that there were amino acid variations in the target epitopes that T-cell receptors could tolerate without loss of epitope recognition, hence explaining the ability of cross-recognition among corresponding but variant viral peptides from different clades. This implies that variation outside HLA anchor residues, unless it interferes with peptide processing or TCR contact, may not have a huge impact on the ability of HIV-specific T-cells to cross-recognize HIV epitope variants (Geels *et al*., 2005).

Many of the studies described assessed the degree of cross-reactivity to the major circulating HIV-clades using peptides based on natural clades or their consensus sequences. A number of studies have suggested the use of sequences generated by use of bioinformatics algorithms or artificial sequences. Such sequences include group M consensus, center of tree (COT), most recent common ancestor (MRCA) and mosaic antigens that are based on potential T-cell epitopes (PTEs).

In a study of 25 clade B HIV-infected individuals, artificial sequences based on consensus, COT and MRCA clade B peptides were tested together with clades A, C and consensus group M peptides. Although there was a slight trend towards higher number of responses for consensus B-based peptides (16 peptides) compared to COT-B (15 peptides), ANC-B (15 peptides) and consensus M (15 peptides), consensus A (14 peptides) and consensus C (13 peptides), no significant differences in the frequency, magnitude and breadth of T-cell responses were obtained between these sets of peptides (Malhotra *et al*., 2007). Detailed characterization of the

responses at the peptide level identified 16 cross-reactive peptides among clades A, B and C, seven of which were identical in sequence. The other nine resulted in 11 variants that had amino acid substitutions outside the HLA anchor residue, except in one. These data further provide evidence of the huge degree of cross-reactivity of HIV-specific T-cells, even to artificial sequences at the peptide level, and further evidence that cross-reactivity may be attributable to conservative or semi-conservative amino acid substitution outside HLA anchor residues across HIV-1 clades (Malhotra *et al*., 2007). This suggests that the use of conserved proteins such as HIV Gag or those regions conserved across group M HIV-1 clades in vaccines (Rolland *et al*., 2007) may result in a high degree of cross-reactivity.

When taken together, data from these studies demonstrate that HIV-1 specific T-cells from individuals infected with a particular HIV-1 clade can cross-recognize HIV proteins or peptides based on other HIV clades. However, some studies stress that even though HIV-specific T-cells can cross-recognize viral variants from other clades, the responses are skewed towards the infecting subtype (Geldmacher *et al*., 2007; McKinnon *et al*., 2005). For vaccine development, this may imply that there is some advantage of matching a vaccine to the infecting subtype, which may be the preferred approach for large mono-clade epidemics, such as that in southern Africa.

A major caveat to the majority of studies that have assessed cross-reactive immune responses is that they have characterised these using the IFN-γ ELISpot assay, where excessive amounts of exogenously loaded peptides are used to detect HIV-specific T cell responses. Two studies have demonstrated that although PBMC from HIV-1 infected people demonstrate substantial cross-recognition of HIV-1 epitope variants, *in vitro* viral suppression of HIV-1 by CD8+ T cells specific for the corresponding epitope sequence was highly impaired (Valentine *et al*., 2008; Bennett *et al*., 2008). In addition, relying on a single immunological readout, namely IFN-γ production, may also mask differences in other functions that variants epitopes may display due to non-optimal TCR contact, for example reduced production of other cytokines, or the ability to stimulate T cell proliferation less well. A recent study suggested that some variant forms of peptides may elicit functions other than IFN-γ from specific T cells, and that amino acid mismatches across peptides variants may result in a different functional profile of HIV-specific T-cell responses to these peptide variants (Richmond *et al*., 2011). In the context of a vaccine, this could impact on the efficacy of the T cell response generated. Furthermore, if the magnitude

of the T cell response is affected by variants binding less well and not stimulating cytokine production from particular T cell clones, this may also have an impact on the effectiveness of that response. Thus, there is need to characterize the cross-reactivity of HIV-specific T-cells to particular variant epitopes by measuring more than one cytokine or function of these cells.

1.7. The significance of HIV-1 genetic variation to vaccine design

Despite the recent failures of a proof-of-concept T-cell based clinical trial in the USA (Buchbinder *et al*., 2008) and South Africa (Gray *et al*., 2011), as well as modest success of the RV144 clinical trial in Thailand in which CD4 T-cell responses targeting mainly Env gp120 and two correlates of risk associated with infection rate (IgG antibodies binding to V1-V2 correlated negatively with infection rate and Env binding to plasma IgA correlated positively with infection rate) in vaccinees were identified (Rerks-Ngarm *et al*., 2009), a T cell-generating vaccine is still considered important for protection, slowing or preventing disease in those who become infected, and reducing secondary HIV transmission. Much of the recent effort has focused on the generation of effective and cross-reactive CD8+ T cell responses specific for HIV (Martinez-Picado *et al*., 2006; Frahm *et al*., 2006).

The current focus on developing CTL-based vaccines is in part due to the failure of generating effective vaccines that can generate neutralizing antibodies. For example, the first empirical approach to neutralizing antibody vaccines, the Vaxgene AIDSVax, was a recombinant form of HIV envelope gp120 protein that failed to protect volunteers from infection because the vaccine did not induce broadly neutralizing antibodies (Johnston *et al*., 2007). This failure to induce broadly neutralizing antibodies is due mainly to the high variability in the envelope protein of HIV, and the fact that broadly neutralizing antibody epitopes are conformationally masked (Burton *et al*., 2004). This failure, and the challenges that it highlighted for developing a vaccine that can induce neutralizing antibodies, shifted the focus to the development of T cell-based vaccines. This shift in focus to T-cell vaccnes is further supported by the phenomenon of HLA-driven viral evolution that was illustrated recently with the first evidence of vaccine-driven T-cell footprints on viral sequences reported in breakthrough infections in the STEP trial (Rolland *et al*., 2011). This section will discuss the efforts underway to cater for the huge genetic diversity the virus exhibits, where various approaches have been taken, including the use of natural

sequences as single or multi-clade vaccine immunogens, centralized sequences, multi-variant or mosaics sequences as well as targeting regions (Table1.2).

1.7.1. *Inserting natural sequence as single or multi-clade vaccines*

The variability of HIV-1 raises concerns about whether HIV-specific T-cell responses would be cross-reactive enough to develop a universal vaccine that would be protective against all the different HIV-1 genetic subtypes and recombinants. Therefore, vaccines based on inserting single or multi-clade natural sequences based on HIV clades circulating in particular regions have been developed. Two vaccines based on this strategy involving sequences from clades B and E is the ALVAC candidate vaccine used as the prime in the RV144 trial in Thailand and the AIDSVAX protein used as the booster in the same trial (Rerks-Ngarm et al., 2009). The ALVAC vaccine is a recombinant canarypox vector vaccine based on Gag and Protease from HIV-1 clade B (LAI strain) and a gp120 from HIV-1 CRF01_AE (92TH023) linked to a trans-membrane gp41 from clade B LAI strain (Rerks-Ngarm et al., 2009). AIDSVAX B/E is a bivalent HIV-1 gp120 glycoprotein vaccine from the same two clades, namely B and E (Pitisuttithum *et al*., 2006; Rerks-Ngarm *et al*., 2006; 2009). When these vaccines were used in a prime-boost regimen in Thailand, modest protection from HIV acquisition (31.2%) was observed (Rerks-Ngarm *et al*., 2009). Another example of a vaccine based on the use of natural sequences is the single clade-based Merck Adenovirus serotype-5 clade B gag/pol/nef (Buchbinder *et al*., 2008). This vaccine was tested in regions with predominantly clade B epidemics, namely North America, the Caribbean, South America and Austria where it was referred to as the STEP Phase IIb trial (Buchbinder *et al*., 2008), as well as in a predominantly clade C epidemic (South Africa) in the Phambili Phase IIb trial (Gray *et al*., 2010; 2011). Although these trials were discontinued due to an increase HIV acquisition in the vaccine group compared to the placebo group, the South African trial showed that 12 % fewer vaccinees mounted a response to clade C peptides compared to clade B, with a 35 % reduction in the overall magnitude of the responses (Buchbinder *et al*., 2008; Gray *et al*., 2011). An important consideration in this type of vaccine design approach is the selection of a vaccine sequence that closely matches the circulating viruses in a given region. It was evident from the STEP study that differences between the vaccine and challenge virus sequences may have limited vaccine efficacy (Rolland *et al*., 2011).

Table 1.2. A summary of the different vaccine immunogen design approaches to cope with viral diversity.

Immunogen design approach	Description	Examples of studies testing the concept	References
Natural sequences (single or multiple-clades)	Based on regional variation of HIV-1. Either single clades or multiple clades can be incorporated into the vaccine immunogen.	The two vaccines used as the prime and boost in the RV144 Phase III trial, ALVAC and AIDSVAX, based on clades B and AE that are the circulating types in Thailand. Merck Ad5 gag/pol/nef vaccine tested in STEP and Phambili Phase IIb trials (clade B-based)	Pitisuttithum *et al*., 2006; Rerks-Ngarm *et al*., 2009; Buchbinder *et al*., 2008 Gray *et al*., 2011.
Centralized sequences	Sequences used are derived using phylogenetic inferences and can be center-of-tree, most recent common ancestor, or consensus sequences. The HIV-1 sequences can be clade-specific when the input sequences are from the same HIV-1 clade, or they can be for a group of clades or for all group M HIV-1 clades.	Clade C Env immunogen consensus and ancestral sequences elicited responses of comparable magnitude and breadth and group M Env immunogens elicited responses comparable to wild type clades A, B and C. Consensus group M responses were comparable to intra-clade responses but higher than inter-clade responses in murine models. Four studies that tested the concept of centralized sequences as immunogens using PBMCs from HIV infected individuals and stimulating them with peptides demonstrated that ancestral, center-of-tree and consensus group M peptides reagents are equally recognized compared to clade-specific centralized peptides reagents in clade infected individuals. However, in a few clade C infected individuals, group M peptide reagents were equally but preferentially recognized compared clade-matched and clade-mismatched peptide reagents, respectively except in Env protein where consensus group M was had significantly higher magnitude of response compared to clade-matched peptide reagents.	Kothe *et al*., 2006; Weaver *et al*., 2006; 2010 Frahm *et al*., 2008; Bansal *et al*., 2006; Malhotra *et al*., 2007; Rutebemberwa *et al*., 2005
Multi-variant or mosaics	Based on computer algorithms to generate and optimize sequences that have a greater depth or coverage of potential T cell epitopes.	Rhesus macaque and human studies have showed that mosaic antigens can be expressed and presented to MHC class I and class II molecules and elicited broader T-cell responses of greater magnitude than consensus M, consensus B, consensus C and natural C immunogens.	Barouch *et al*., 2010; Santra *et al*., 2010; Ndhlovu *et al*., 2011
Conserved regions	Based on conserved regions of HIV across different HIV-1 clades that contain epitopes covering all HLA supertypes.	Humanized mice vaccinated with a vaccine based on 14 conserved regions of HIV proteome elicited HIV-specific T cell responses similar to those observed in natural HIV infection and vaccine trial participants.	Letourneau *et al*., 2007

Other examples of using natural sequences to closely match viruses in a given region are the vaccines being developed for South Africa, where clade C viruses are predominant. These vaccines are based on the Du422 isolate, an R5 virus that was selected based on its similarity (>98%) to a South Africa subtype C consensus sequence (Williamson *et al*., 2003). Indeed, a DNA prime followed by MVA boost regimen based on HIV-1 clade C *gag*, *rt*, *tat*, *nef* and *env* in non-human primates generated potent, broad, multifunctional CD4 and CD8 T-cell responses (Burgers *et al*., 2009).

1.7.2. *Centralized sequences*

Unlike developing vaccines using natural sequences, bioinformatics approaches have been used to design vaccines based on centralized sequences or sequences representative of the different HIV-1 subtypes or strains. Such vaccine immunogens are artificially designed using phylogenetic approaches in three main ways, namely consensus group M which is representative of all the HIV-1 group M viruses sequenced to date (Gaschen *et al*., 2002; Kothe *et al*., 2006), the center-of-tree (COT), which is a sequence equidistant from all sequences used to construct a phylogenetic tree (Nickle *et al*., 2003; Nicle et al., 2007; Rolland *et al*., 2007), and sequence based on the most recent common ancestor (MCRA) of the sequences of interest (Doria-Rose *et al*., 2005; Kothe *et al*., 2006). While the COT and MRCA are based on the root of a tree in phylogenetic analyses, they are different in that if there are outlier sequences in a sample, the algorithm that generates the MRCA can be biased away from the majority of the sample sequences while the algorithm for generating the centre-of-tree will control for this while still maintaining the biological properties of the virus (Rolland *et al*., 2007). Theoretically, in all these methods, the sequence can be generated for a single HIV-1 clade, a selected group of HIV-1 clades, or for all the global HIV-1 clades.

Most of the work performed to test central sequences as T-cell vaccines has been performed in primate studies, using mainly HIV-1 Env immunogens. The magnitude and breadth of IFN-γ responses were comparable between ancestral and consensus clade C Env sequences in mice immunized with ancestral C and consensus C *env* DNA vaccines (Kothe *et al*., 2006). In a separate study, there different strains of mice were immunized with consensus group M and clades A, B and C *env* immunogens in a DNA prime-recombinant vaccinia virus boost regimen (Weaver et al., 2006, 2010). A consensus group M Env immunogen was immunogenic and

induced T-cell responses greater of breadth and magnitude than any other wild subtype A, B or C, but similar to polyvalent immunogen, when the immunogens were administered together (Weaver *et al*., 2006, 2010). Characterization of these responses showed that T-cell responses to the consensus group M vaccine immunogen were comparable to intra-clade responses but significantly greater than inter-clade responses (Weaver *et al*., 2006, 2010).

Although there are no similar studies like these murine-model studies in humans yet, measuring HIV-specific T-cell responses to centralized immunogens in HIV-1 infected persons using assays such as the IFN-γ ELISpot assay are the kinds of theoretical studies to predict whether consensus sequences may be better. Four studies to date have tested the reactivity of at least one of clade-specific central sequences (consensus, center-of-tree and ancestral for clade A, B and C) and compared them to at least one of the group M central sequences (consensus, center-of-tree and ancestral) based on HIV-1 Gag, Nef or Env proteins mainly in clade-B infected study individuals and few a few clade-C infected individuals (Bansal *et al*., 2006; Frahm *et al*., 2008; Malhotra *et al*., 2007; Rutebemberwa *et al*., 2005). In one of these studies, group M peptide reagents were less frequently targeted in clade B infected individuals; contrary to this, clade C infected individuals from the same study found comparable recognition between group M based peptides and clade C based peptides (Frahm *et al*., 2008). In two of these four studies, magnitude and breadth of T-cell responses was comparable among clade-specific peptides based on centralized sequences and group M centralized peptide reagents (Bansal *et al*., 2006; Malhotra *et al*., 2007), except in a few clade C infected individuals in one in which although not significantly different, there was a trend towards preferential recognition of clade C and group M compared to central sequences based on mismatched clades. Overall, the three studies demonstrated that consensus group M based peptide reagents are equally recognized compared to clade-matched peptide reagents and this recognition is more than that to clade-mismatched peptide reagents. On the other hand, the fourth study demonstrated that consensus group M peptides have significantly 2-fold higher magnitude of T-cell response compared to clade-matched consensus peptides (consensus B) or clade matched strain specific peptides (HIV_{MN}, Rutebemberwa *et al*., 2005). However, the fact that this fourth study was based on Env, a more variable protein compared to Gag and Nef that were investigated in at least one of other three studies is possibly an explanation for the difference. However, an important consideration about central sequences generated using phylogenetic means is that these sequences are artificial and depend on the phylogenetic method employed to generate them, and the sequences used in the analyses.

Furthermore, the phylogenetic models employed in generating these central sequences do not take into consideration events that occur during natural infection such as recombination (reviewed in Arenas *et al*., 2010) and dual infection that might require the use of more than one sequence from the same individual for use in generating these sequences. The mosaic vaccine approach discussed below tries to cater for such events as recombination using computerized algorithms to generate mosaic antigens by combination of different sequences.

1.7.3. *Multi-variant or mosaic vaccines*

Mosaic antigens are immunogens that are produced by combining sequences from different HIV variants or clades using artificial recombination methods designed to mimic the recombination process that occur during natural HIV evolution. Vaccines based on the mosaic approach utilize computerized algorithms to generate and optimize sequences similar to naturally circulating HIV sequences such that in their combination, demonstrate a greater depth or coverage of HIV-1 potential T-cell epitopes (PTEs) for different in HIV proteins (Fischer *et al*., 2007; Thurmond *et al*., 2008; Corey and McElrath, 2010). Therefore, theoretically, the optimization process can be performed for a single HIV-1 clade or for the entire group M HIV-1 clades. Preliminary data on this approach suggests that the mosaic approach algorithm provides enhanced coverage of 9-mers when compared to the COT approach when the same data set was used (Fischer *et al*., 2008). Mosaic vaccines encompass the concept of polyvalency, but at the epitope level, since corresponding variants of different epitopes from different clades are included.

Studies in rhesus macaques (Barouch *et al*., 2010; Santra *et al*., 2010) and humans (Ndhlovu *et al*., 2011) have demonstrated that mosaic antigens can be expressed and presented to MHC class I and class II molecules in these models. Mosaic antigens in these studies generated responses of greater magnitude and breadth than consensus reagents (Barouch *et al*., 2010; Santra *et al*., 2010). These macaques were either immunized with plasmid DNA prime and recombinant vaccinia virus boost containing constructs that express Gag and Nef based on either consensus or mosaic antigens (Santra *et al*., 2010), or with Gag, Pol and Env based on natural, consensus or mosaic antigens expressed by replication incompetent adenovirus serotype 26 (Ad26) vectors (Barouch *et al*., 2010). The studies showed that mosaic vaccines, compared to consensus or natural immunogens, elicited CD8+ T-cell responses to more epitopes and more variant sequences (Barouch *et al*., 2010; Santra *et al*., 2010). Mosaic vaccines have not yet been tested

in humans. However, in ICS and tetramer staining assays using PBMC from HIV-1 subtype B and subtype C infected study individuals transduced by a bivalent mosaic Gag vaccine expressed by replication incompetent Ad26 vector, mosaic antigens were expressed and presented to CD8+ T-cells and elicited T-cell responses of great magnitude, breadth, depth function compared to natural clade B and C Gag immunogens (Ndhlovu *et al*., 2011). It is important to note that this study only evaluated this in clade B and C infected individuals, and the utility of this approach in other epidemics, for example where multiple diverse clades co-circulate, may require further testing. Overall, mosaic antigens result in a greater breadth of responses, and an important feature of this approach is the representation of many variant forms of the same epitope.

Studies on HLA-associated amino acid changes in populations sequences have generated data to show that HLA class I-restricted CD8+ T-cell pressure not only results in escape from immune recognition (Carlson *et al*., 2008; Leslie *et al*., 2004), but also in the formation of new CD8+ T-cell epitopes (Bhattacharya *et al*., 2007). This may imply an evolutionary advantage to the virus rather than the host. Of note is the data generated by a recent study that demonstrated that although these newly formed epitopes ('neo-epitopes') could elicit higher avidity and more potent IFN-γ-producing CD8+ T-cells than their parent epitopes, the responses were not more cytotoxic than their parent epitope variants (Keane *et al*., 2011). This ability to make responses to new epitopes may increase the positive impact on CTL cross-reactivity. Since the presence of neo-epitopes and their parent variants are variant forms of corresponding epitopic regions, mimicking a mosaic antigen, it is logical to think that mosaic vaccines might elicit broad responses of high magnitude than other vaccine formulations but the responses might not necessarily have antiviral capacity. Furthermore, partial escape occurs frequently, and new T cell responses are mounted repeatedly in vivo with data being generated providing evidence of continuous T-cell recruitment during chronic HIV infection (Allen *et al*,. 2005). TCR vary widely over time as some clones die, and others are recruited. This likely has a dramatic impact on CTL cross-reactivity over time.

1.7.4. *Targeting conserved regions*

Despite the huge HIV variability within and between populations, there are regions of the virus that are invariable, possibly due to the fact that these regions have structural and functional roles to the virus and are constrained. Therefore, the cost-benefit assessment of mutations in these

regions would suggest that these regions are less likely to be mutated, and escape in these regions might result in a decrease in viral fitness, as been demonstrated in a number of studies (Fernandez *et al*., 2005; Martinez-Picado *et al*., 2006; Leslie *et al*., 2004). There are previous studies that have suggested targeting of conserved regions as possible vaccine immunogens (Altfeld *et al*., 2006; Kent *et al*., 2005).

One group (Rolland *et al*., 2007) that proposes the use of conserved regions suggest a vaccine which they referred to as a conserved elements (CE)–vaccine, that is composed of 45 discrete viral segments of at least eight amino acids (resembling the minimum length of a CD8+ T-cell epitope) that fulfill a certain conservation criteria. This analysis resulted in identification of first tier segments in which there was 98% conservation of a single amino acid at a single position across the group M sequences, and second tier segments in which there was 99% conservation of two variant amino acids at a single position across group M sequences. The segments encompassed motifs in Gag and Pol proteins that covered all nine HLA supertypes (B7, A3, A2, A24, B44, B27, B62 and B58), thus potentially providing 80% coverage regardless of ethnicity (Rolland *et al*., 2007). In another study, a gene coding for the 14 most conserved regions of the HIV proteome (HIV_{CONSV}) was designed (Letourneau *et al*., 2007). The three most studied vaccine vectors, namely plasmid DNA, human Ad5 and MVA were used, and induced HIV-specific T-cells in mice producing IFN-γ, IL-2, TNF-α and capable of killing target cells (Letourneau *et al*., 2007). Together with data generated using PBMC samples from an HIV vaccine clinical trial and from healthy donors, the study demonstrated that the vaccine based on this approach primed subdominant CD4 and CD8 + T-cell responses to conserved epitopes in natural infection (Letourneau *et al*., 2007). In another study, full length HIV-1 proteins were characterized for variation using Shannon entropy for four major HIV-1 clades, namely A1, B, C and D and four highly conserved regions rich in previously described CD8+ T-cell epitopes (Yang *et al*., 2009). Subsequent testing of this immunogen in chronically HIV-infected individuals demonstrated that these regions are commonly immunogenic, that is, targeted in multiple individuals (Yang *et al*., 2009). A detailed characterization of conserved regions in a more recent study using a technique known as random matrix theory showed that there are HIV-1 conserved regions that are immunologically vulnerable which they termed ‘HIV sectors’, which may be important as vaccine immunogens (Dahirel *et al*., 2011). These regions have structural constraints and are highly targeted by HIV-1-specific T-cells, particularly five sectors

located in the HIV-1 Gag protein that co-evolve independent of each other and of the rest of the Gag protein (Dahirel *et al*., 2011).

Of course, all these vaccine design approaches are theoretical and which approaches are superior can only be resolved by demonstration of efficacy in different human populations with different viral clades circulating. This requires Phase III vaccine trials, a lengthy and expensive process. However, using advancements in immunological techniques such as ELISpot and flow cytometry, it is possible to predict the ability of these vaccine immunogens to cross-react to different viral variants and clades, and examine the quality of responses produced, using test reagents based on the different vaccine antigens in HIV-infected people. This is the approach employed in this thesis.

1.8. Aims and objectives of the thesis

It is not questionable that the huge genetic diversity of HIV has made effective vaccine design extremely challenging. The question of whether region-specific vaccines are necessary, or a universal vaccine is feasible, still needs to be answered. Due to increased global travel of people coming from regions with different clade epidemics, and its associated HIV risk, it is likely that the frequencies of different clades and their distribution in different regions are going to change over time. This calls for vaccines with cross-clade immunity for wider protection at the population level. Nevertheless, there are a number of HIV vaccine candidates based on specific HIV-1 subtypes that are currently in development, some in pre-clinical stages and some in clinical trials, and it is important to predict whether these vaccines may be effective in unmatched populations where different HIV-1 subtypes are circulating. Vaccine trials are the only means to definitively answer whether a vaccine is effective in a particular population, but these trials are laborious and lengthy, and before or while they are underway, clues to the potential cross-reactive nature of particular immunogens can be obtained by assessing the identity, degree and nature of cross-reactive HIV-specific T-cell responses in the laboratory, using immunological techniques such as the ELISpot assay. It is also possible to describe the quality of these responses by assessing the functions of the cells beyond IFN-γ production. While these studies are very useful to evaluate cross-reactivity in natural infection, it should be noted that these *in vitro* cross-reactive responses may differ from actual vaccine induced responses. The overall aim of this thesis was to characterize the identity of regions of HIV-1 that

are cross-reactive, assess the degree of cross-reactivity, as well as determine the effect of sequence differences (in cross-reactive but variant epitopes) on the functional nature of cross-reactive HIV-specific T-cells, using peptides based on vaccine immunogens.

Aim 1

To assess the degree of cross-reactive immune responses in subtype C-infected individuals using subtype-specific peptides from five different peptide variants based on two vaccine inserts for two subtype C (South African subtype C and a Chinese subtype C) and three consensus subtypes, A, B and D.

Rationale

South Africa has one of the world's highest HIV prevalence rates and a number of phase I and II trials and one phase IIb efficacy trials have taken place using clade A and B-based candidate vaccines. Further clinical trials are likely to follow using clade C-based immunogens. Although a number of studies have assessed the cross-reactivity of HIV-specific T-cells using various clade-based peptide sequences in HIV-infected individuals, the cross reactivity of these T-cells to sequences based on actual vaccine inserts has not been performed. The ability of candidate vaccines that are based on particular clades to elicit cross-reactive T cell immunity may be predicted by testing reagents based on these vaccine sequences in HIV-infected people using laboratory methods, such as the ELISpot assay.

Specific objective

To determine and characterize reactive peptides in HIV-1 subtype C infected individuals in an IFN-γ ELISpot assay using Gag peptides based on consensus subtypes A, B, D and two subtype C peptide sequences, one from China (C_{CH}), and the other one from South Africa (C_{Du422}), as well as identifying regions within the Gag protein that are exclusively and mutually recognized in a subtype C-infected population.

Aim 2

To characterize cross-clade immune responses in HIV-1 infected individuals using HIV-1 Gag and Nef peptides sets based on central sequences representative of all HIV-1 group M viruses.

Rationale

Africa has regions that demonstrate differential predominance of HIV-1 subtypes and circulating recombinant forms. For example, West/Central Africa has the greatest assemblage of HIV subtypes, circulating recombinant forms and unique recombinant forms, while other regions such as southern and East Africa have particular subtypes dominating. This aim was part of a multi-center study involving five African countries, namely South Africa, Ethiopia, Cameroon, Kenya and Uganda. These countries demonstrate different epidemics with Cameroon and Uganda demonstrating multiclade epidemics with almost all HIV-1 subtypes circulating in Cameroon, and Uganda mainly subtypes A, D and A/D recombinants. South Africa and Ethiopia demonstrate distinct subtype C epidemics, and Kenya mainly HIV-1 subtype A epidemic, with a few recombinant sequences from A, C and D. The five-center study aimed to define genetic variation of HIV in these countries participating in or preparing for vaccine clinical trials, and evaluate T cell responses to group M consensus Gag and Nef peptide reagents in order to identify responses that are cross-reactive across HIV-1 clades in ethnically diverse populations where different HIV-1 clades are circulating. This aim reports on work performed by the South African part of this network, in an effort to address the effect of HIV genetic diversity on cross-clade HIV-specific responses to group M-based peptide reagents.

Specific objective

To characterize the T-cell immune responses to HIV-1 Gag and Nef consensus group M peptides in HIV-1 infected individuals from South Africa, determine the relationship with the infecting viral sequences, and compare responses to group M and subtype C reagents.

Aim 3

To determine whether sequence variation in cross-reactive T-cell epitopes affects the qualitative nature of the HIV-specific T-cell immune response.

Rationale

The majority of studies characterizing cross-reactivity of HIV-specific T-cells have used a single immunological readout, the IFN-γ ELISpot assay, using excessive, non-physiological amounts of peptides. However, there are previous studies that have demonstrated the importance of other functions of T cells in protective responses, such as cytokine production, cytotoxicity and proliferation. The effect of HIV-1 variation on cross-reactivity of these HIV-specific T-cells has

not been assessed in the context of these other functions. Therefore, this study aimed to measure in more detail the true cross-reactive potential of HIV-specific T-cells recognizing corresponding HIV-1 peptides with amino acid mismatches, using peptide dilution assays and multi-parameter flow cytometry-based functional assays.

Specific objective

To characterize the cytokine profile (IFN-γ, IL-2, TNF-α and MIP-1β) and cytotoxic potential (perforin, granzyme B and CD107a expression), as well as proliferative capacity, of HIV-specific T-cells in response to stimulation with corresponding peptide variants with amino acid mismatches.

CHAPTER 2

INVESTIGATING INTRA- AND INTER-CLADE HIV-SPECIFIC T-CELL RESPONSES

Contents

2.1. INTRODUCTION

The development of a safe, globally effective and affordable vaccine offers the best hope for the future control of the HIV pandemic. One of the major challenges in developing such a vaccine is the high degree of genetic diversity the virus exhibits. The extensive genetic variation of HIV is fuelled by high mutation, recombination and replication rates, partly driven by host cellular and humoral immune pressure (Choisy *et al*., 2004; Wolinsky *et al*., 1996). It is anticipated that the next generation phase IIb and phase III vaccines will contain inserts matched to the dominant clade C and E viruses circulating in targeted populations of Southern Africa and Thailand respectively (Glenda Gray; AIDS Vaccine 2011, Thialand; AIDS Vaccine Clinical Trials Update). However, the long and expensive process of product development, together with the high diversity of the global pandemic, makes matching vaccines to circulating viruses difficult. It is thus likely that candidate vaccines designed for one clade will be tested in epidemics where multiple subtypes and recombinant forms co-circulate. It is therefore important to predict cross-clade epitope coverage.

T-cell immunity has been found to play a role in HIV control (Frahm *et al*., 2006; Martinez-Picado *et al*., 2006; Miura *et al*., 2009). The importance of responses to Gag is well documented, with studies showing that the magnitude of anti-Gag CD8+ T-cell responses inversely correlates with plasma viral load, (Edwards *et al*., 2002; Kiepiela *et al*., 2007), and that preferential targeting of this protein during infection correlates positively with lower viral load (Edwards *et al*., 2002; Masemola *et al*., 2004; Zuniga *et al*., 2006). Other studies have shown that the breadth of anti-Gag T-cell responses is associated with lower viral loads Geldmacher *et al*., 2007; Kiepiela *et al*., 2007). Collectively, these data strongly implicate Gag as an important target of HIV-specific T-cells for inclusion in candidate preventative vaccines. Of major importance for preventative vaccine development is the identity of regions within the HIV-1 proteome that can be targeted by T-cells and that are cross-reactive between different viral clades. Previous studies that examined cross-clade HIV-1 Gag T-cell immune responses in an environment of multiple circulating clades (Geels *et al*., 2005; Gudmundsdotter *et al*., 2008; Gupta *et al*., 2006) have found that HIV-infected individuals can mount robust cross-clade HIV-specific T-cell immune responses, but with a preference for the predominant circulating or infecting clade (Geldmacher *et al*., 2007; McKinnon *et al*., 2005).

South Africa has a high incidence of HIV-1, with clade C accounting for over 95% of infections (Iweriebor *et al*., 2011; Musyoki *et al*., 2011; van Harmelen *et al*., 1997). A number of phase I and II trials and one phase IIb efficacy trial have taken place here, testing constructs based on clade A, B and C-based candidate vaccines (Gray *et al*., 2011; Peters *et al*., 2007; Vardas *et al*., 2010). Following on from the first demonstration of vaccine-induced protection from HIV-1 acquisition in the RV144 trial in Thailand (Rerks-Ngarm *et al*., 2009), follow-up trials in high incidence settings such as South Africa are currently being planned. A vaccine will need to protect against the high degree of HIV diversity and it is thus imperative to be able to predict the level of T-cell coverage and cross-reactivity. In this study, examined intra- and inter-clade cross-reactivity of HIV-1-specific T-cell responses to Gag was examined, using peptides matching candidate South African and Chinese clade C vaccine constructs, and compared these with clades A, B and D consensus-based peptides. The South African and Chinese *gag* genes have been included in candidate HIV-1 vaccines that have been tested for safety in phase I clinical trials (Kresge, 2009; Vasan *et al*., 2010a; Vasan *et al*., 2010b. This study was performed in a South African clade C-infected population, where the infecting virus from each individual was sequenced and T-cell responses assessed to the different peptide sets. This allowed two inter-related aims to be explored: a) to identify the location of commonly and exclusively targeted epitope regions in Gag and relate these to the level of virus variability; b) to identify the extent of intra- and inter-clade recognition using peptide sets that match vaccine inserts.

2.2. MATERIALS AND METHODS

2.2.1. Participants

Forty HIV-1 infected individuals were recruited for the study from HIV clinics located in Johannesburg and Cape Town in South Africa. Eligibility criteria were willing and able to provide informed consent, clinically asymptomatic, ART naïve, and with a peripheral blood CD4 count that was above 350 cells/mm^3. The study was approved by the ethical review boards of the University of Cape Town and University of the Witwatersrand and each study participant provided written informed consent. CD4 counts were determined using the FlowCARE PLG CD4 monoclonal antibody reagent kit (Beckman Coulter, CA, USA) according to the manufacturer's instructions. Plasma HIV-1 RNA (viral loads in copies/ml) was determined using the NucliSENS EasyQ® HIV-1 (version 2.0, BioMerieux SA, Lyon, France) according to manufacturer's instructions. This kit has a limit of detection of 50 copies/ml of HIV-1 RNA.

2.2.2. PBMC processing

Sixty ml of blood was drawn by venipuncture in Acid-Citrate-Dextrose (ACD) tubes. Peripheral Blood Mononuclear Cells (PBMCs) were isolated using standard Ficoll-Hypaque density gradient centrifugation. Firstly, the Ficoll-Hypaque was allowed to reach room temperature and 15ml was added into 50ml Leucosep tubes (Sigma-Aldrich, USA). This was centrifuged at 1000g for 10 minutes (Heraeus 1.0R Centrifuge). Study participants' blood in the ACD vacutainers was mixed gently and 30ml poured onto the separating disc. This was centrifuged for 15 minutes at 1000g as before. The centrifugation process resulted in four distinct layers, first the red blood cells, granulocytes and dense solution at the bottom of the tube, secondly the disc separating red blood cells from PBMC, thirdly the ficoll layer containing PBMC and lastly the plasma layer on top. Plasma was removed without disturbing the PBMC layer and stored at -80^{0}C. The PBMC layer was carefully transferred to a 50ml falcon tube, diluted to 50ml using phosphate-buffered saline (PBS) with 1% foetal calf serum (FCS; Invitrogen, USA) and centrifuged at 250g for 10 minutes. The supernatant was discarded and the pellet re-suspended in 10ml PBS and 1% fetal calf serum. PBMC were

counted using a Guava personal counter (Guava Technologies, USA) and stored at 10-20 x 10^6 per cryovial in FCS with 10% dimethyl sulphoxide (DMSO) and stored in liquid nitrogen.

2.2.3. Peptides

A total of 540 peptides were used, 120 each for C_{Du422}, C_{CH} and B and 90 peptides each for the A and D sets used. The clade C and B peptides spanned the full length of the Gag protein while clade A and D peptides covered the p17, p24 and p2 regions. Peptides corresponding to p15 were excluded from data comparisons between peptide sets because this region was omitted from the A and D sets. The peptide sets were derived from HIV-1 Gag clades A, D, consensus B, and vaccine insert matched peptides from C_{Du422} clade C (South Africa) and a Chinese clade C strain (C_{CH}). Table 2.1 details preclinical and clinical trials in which vaccines containing the C_{Du422}, C_{CH} and A sequences were tested. All peptides sets were provided by the International AIDS Vaccine Initiative, apart from the clade B peptides, which were provided by the National Institutes of Health AIDS Research and Reference Reagent Repository.

2.2.4. HLA typing

Professor Clive Gray performed high resolution HLA class I A, B and C typing and the results provided for analysis. This was done using sequence specific PCR. Briefly, DNA was extracted using the QIAGEN DNA isolation kit for blood (QIAGEN, Chatsworth, CA). High-resolution HLA class I genotyping was performed by sequencing of exons 2, 3 and 4 using Atria Allele SEQR kits (Abbott Diagnostics) and Assign SBT 3.5 (Conexio Genomics, Fremantle, Australia).

Table 2.1. Gag peptide sets and sequences

Peptide set name	HIV sequence information and references	Vaccine trials and references
Clade A	Gag A p24/p17 consensus Hanke *et al*., 2000	Peters *et al*., 2007 Jaoko *et al*., 2008
Clade B	Gag B Consensus Catalogue # 8116 NIH AIDS Research Reference Reagent Program https://www.aidsreagent.org	Not Applicable
South Africa clade C (C_{Du422})	C_{Du422} Williamson *et al*., 2003 GenBank: AF544010	Burgers *et al*., 2009
China clade C (C_{CH})	C/B' synthetic Huang *et al*., 2008	Vasan *et al*., 2010a, b, 2011
Clade D	Gag D p24/p17 consensus Courtesy of Bette Korber, LANL, Los Alamos, NM	Not Applicable

2.2.5. IFN-γ ELISpot assay

T-cell responses were assessed by IFN-γ ELISpot assay as previously described (Masemola *et al*., 2004). Briefly, PBMC were plated in 96-well polyvinyledene difluoride-backed plates (Microsep, Millipore Products, France) that had been coated previously with 50µl of anti-IFN-γ mAb 1-D1K (5mg/ml, Mabtech, Sweden) overnight at 4^{o}C. The unbound antibody was washed away three times with 200µl/well of sterile PBS. Peptides arranged in 5 different pools and 24 matrix pools for each of the five peptide sets, were added (50µl/well) directly into the wells in duplicate for the pools and singly for the 24 matrices at a final concentration of 1.5µg/ml. Fifty microlitres of 8-11-mer Cytomegalovirus Epstein bar virus and Flu virus (CEF) peptide pool (National Institute of Health AIDS Research and Reagent Repository) were added at a final concentration of 1.5µg/ml, and 50µl of Phytohaemagglutinin (PHA, BD Biosciences, USA). PBMC were added at 100 000 cells/well and incubated overnight at 5% CO_2, 37^{o}C. On the following day, the plates were washed six times with 200µl/well of PBS containing 0.05% Tween 20 (PBS-Tween, Sigma, USA). Fifty microlitres per well of biotinylated anti-human IFN-γ monoclonal antibody clone 7-B6-1 (mAb7-B6-1, MabTech Sweden) diluted to 2µg/ml in PBS-10% FCS was added and the plates were incubated for 3 hours at room temperature. The plates were washed again six times with 200µl/well PBS 0.05% Tween and Streptavidin-Horse-Radish Peroxidase (HRP) (BD Pharmingen, Canada) at 1:500 with PBS-10% FCS was added. This was incubated at room temperature for an hour.

The plates were washed for the third time, six times with 200µl/well of PBS-Tween. The development step was performed with 100µl/well of Nova Red substrate (Vector Laboratories, CA, USA) for six minutes in the dark. Emptying the wells and rinsing them in cold tap water stopped the reaction. Spots were counted on a CTL Analyzer (CTL Technologies, Cleveland, USA) and expressed as spot forming units per million (SFU/10^6) PBMC. A response was considered positive if the SFU/10^6 exceeded 100 after background subtraction. An assay passed if there were <5 spots in each of the media control wells, not >100 spots in each of the media and cell (background) wells and >400 spots in the PHA control wells. Single peptide reactivity was confirmed after deconvoluting the pool/matrix reactive peptides in the initial screen. The number of epitopes was determined after taking into consideration overlapping peptides. Two consecutive reactive peptides were considered as one response and three consecutive reactive peptides as two responses.

2.2.6. Epitope prediction, peptide mapping, binding affinity and processing prediction

Epitope prediction within confirmed reactive peptides was determined using the Epitope Location Finder (ELF) tool on the Los Alamos HIV database (www.hiv.lanl.gov) and the NetMHC version 3.2 server (www.cbs.dtu.dk/services/NetMHC). The tool predicts binding of peptides to a number of different HLA alleles using a novel approximation algorithm that employs artificial neural networks and weight matrices (Nielsen *et al*., 2003). The peptide map was generated using the PepMap tool on the Los Alamos HIV database (www.hiv.lanl.gov). The tool maps an input of peptides on the HIV reference sequence HXB2. All the reactive peptides in the study were loaded onto the tool and mapped onto the HXB2 reference sequence.

Binding affinities were predicted using NetMHC3.2. To further characterize why some peptide variants may be more cross-reactive than others, the effect of amino acid mutations on the different steps involved in epitope processing and presentation by their respective HLA alleles was investigated by predicting proteasomal cleavage scores, transport by transport associated with antigen processing (TAP) and MHC class I binding scores using a computational method (www.immuneepitope.org). The analysis was performed on all peptides

reactive in the study and the proteasome, TAP, MHC processing and total scores were predicted as well as the MHC IC_{50}.

2.2.7. Statistical analyses

Statistical analyses were performed using GraphPad Prism version 5.00 for Windows (GraphPad Prism Software, San Diego, CA, USA). All data were analyzed by use of non-parametric statistics. The Friedman one-way ANOVA test for matched pairs followed by the Dunn's post-test was performed to test for any significant differences in genetic distances, magnitude and breadth of responses among the different peptide sets. The non-parametric Kruskal-Wallis test for unmatched pairs was performed to test for differences in the median entropies of peptides in different recognition categories and for differences in scores from epitope prediction algorithm. All tests were two-tailed and $p < 0.05$ was considered significant. Correlation analysis was performed using the non-parametric Spearman method.

2.3. RESULTS

2.3.1. Characteristics of study individuals

Immunological data were available for 39 participants. The median age was 28 years (range 22-47 years). The median CD4 count was 492 cells/mm^3; (range 295-1437 cells/mm^3; Table 2.2). The study cohort had a median viral load of 11 000 copies/ml (range 200-260 000). Thirty-six participants had a response to at least one Gag peptide from one of the five peptide sets used. Seventeen study participants who had Human Leukocyte Antigen (HLA) A and B typing data available (Table 2.3) had their reactive peptides further characterized.

Table 2.2. Summary of clinical data of study participants

Characteristic	**Median (Range)**				
Age (Years)	28 (22-47)				
Plasma HIV RNA (Copies/ml)	11000 (200-260000)				
CD4 Count (cells/µl)	492 (295-1437)				
Frequency of HIV-specific T-cell responses	**C_{Du422}**	**C_{CH}**	**B**	**A**	**D**
N (%)	36 (92)	35	27	30	34

N (number of participants) = 40. N=39 participants were screened for immunological responses and 36 had responses to at least one Gag peptide. C_{Du422}: South African clade C, C_{CH}: Chinese clade C; A, B and D are clades A, B, and D sequences.

2.3.2. Genetic relatedness and epitope coverage of infecting virus by ELISpot peptide reagents

The peptide reagents used in the ELISpot assays are detailed in Table 2.1. The genetic distance between these reagent sequences from infecting viral sequences from each of the study participants was determined. These results have been reported previously (Lycias Zembe MSc thesis, UCT, 2007) and are provided in Appendix B for reference. Briefly, all study participants were confirmed as being infected with HIV-1 clade C (Figure B3.1, Appendix B3 and Figure B4.1 Appendix B4). Genetic distances for the peptide sets South African clade C_{Du422} and Chinese clade C (a synthetic B/C recombinant, C_{CH}) and clade B peptide sets were based on the full-length Gag amino acid sequences, whilst those for the other clades were based on only p17p24p2 regions. The genetic distance of the C_{Du422} and

Table 2.3. Clinical characteristics of selected study participants

P.I.D	Age (years)	CD4 count (cells/µl)	Viral load (copies/ml)	HLA-A		HLA-B		HLA-C	
CC05	32	610	3900	A29:02	A30:02	B42:01	B42:01	C17:01	C17:01
CC12	47	536	11000	A 30:01	A30:02	B07:02:01	B58:01	C03:02	C07:02
CC15	27	452	180000	A 02:01	A29:02	B08:01:01	B42:01	C17:01	C17:01
CC17	40	884	43000	A 02:05	A26:01	B51:01	B14:01	n/a	C18:04
CC19	23	569	7800	A 30:02	A74:01; A74:02	B07:02:01	B15:03; B95:03	C02:02; C02:10	C07:02
CC22	31	420	15000	A 23:01	A33:03	B08:01; B08:03	B53:01; B35:29	n/a	n/a
CC23	31	565	72000	A02:01	A30:02	B08:01:01	B45:01	n/a	C16:01
ST03	29	578	21000	A30:01	A74:01; A74:02	B42:01	B58:01	C03:02	C17:01
NM06	27	1321	130000	A30:02	A43:01	B15:10	B57:03:01	C04:01; C04:09N	C18:01; C18:02
NM07	35	1010	18000	A01:01	A30:01	B15:03; B95:03	B45:01	C04:01; C04:09N	C06:02
BN08	28	1437	740	A01:01	A30:01	B15:03; B95:03	B45:01	C04:01; C04:09N	06:02
MT09	33	605	2700	A23:01;A23:04	A2902; A29:03	B15:03; B95:03	B42:01	C02:10	C17:01
RL12	19	628	3500	A23:01; A23:01	A29:02:01;A29:03	B51:01:01	B14:01	C07:02	C07:27
TM16	29	702	15000	A23:01	A23:01	B08:01	B45:01	C06:02; C06:11	C03:04; C03:07
ER17	35	999	2000	A29:01	A34:04	B18:01; B18:17N	B8101; B81:02	C07:04; C07:11	C08:04
HN18	29	471	11000	A29:02	A74:01	B1503;B9503	B58:02	C02:10; C02:02	C06:02
NS19	34	1167	3200	A30:01	A74:01; A74:02	B42:01	B58:02	C17:01	C03:02

N = 17 had complete HLA typing for loci A and B. N = 2 had no complete HLA typing for C locus. Suffix N denotes HLA alleles whose protein is known not to be expressed
n/a denotes HLA alleles' data not available.

and C_{CH} peptide sequences to the infecting viral sequences of the participants were similar, with a median amino acid distance for C_{Du422} of 5.5%, and for C_{CH} of 6% (range 4-15%, Figure B3.2, Appendix B3). However, the genetic distance between the peptide sets matching consensus clades A, B and D with the infecting clade C sequence was significantly greater ($p<0.0001$), with median distances of 13%, 12% and 11%, respectively (Figure B3.2, Appendix B3).

Differences in antigenic potential of the sequences were investigated next, by examining the proportion of matching T-cell epitope-length peptides within the infecting viral sequences, compared to the peptide sets being tested. The coverage of putative 9mer epitopes within infecting sequences was similar for C_{Du422} and C_{CH} (55.28% and 53.30% respectively, Figure 2.1A, red bars). When 9mers were aligned that differed by one amino acid, there was a similar trend and an increased frequency of epitope coverage, at 85%, 81%, 63%, 64% and 71% for peptide sets C_{Du422}, C_{CH}, A, B and D, respectively (Figure 2.1A, orange bars). Inclusion of 9mer that differed by two amino acids further increased the coverage in all peptide sets, 94.2%, 93%, 79%, 81%, 86% and 96% for C_{Du422}, C_{CH}, A, B, D and all peptide reagents combined, respectively (Figure 2.1A, yellow bars). Consistent with the increased amino acid divergence in peptide sets A, B and D from the infecting virus (Figure B3.2, Appendix B3), matched epitope coverage was lower for these peptide sets, at 33.21%, 30.35% and 37.39% respectively (Figure 2.1A). Conversely, the proportion of mismatched epitopes (by one or 2 amino acids) was higher for the subtype B peptide reagent and least for the subtype C peptide sets (Figure 2.1C). In addition, the proportion of mismatched epitopes was reduced by use of all the five peptide reagent sequences together (Figure 2.1C, also evident in Figure 2.1A). These data suggest that intra-clade T-cell reactivity may be similar due to lower genetic divergence compared to between clades.

Epitope coverage analysis was dissected to the single amino acid level across the alignment to see the positional epitope coverage of subtype C-viruses by the five peptide sets. Consistent with genetic distance data and epitope coverage above, these subtype C-sequences were equally covered by the two clade C vaccine inserts, C_{Du422} and C_{CH} (Figure 2.1B and C). The use of only C_{Du422} vaccine candidate had lower coverage; at 55.28% (Figure 2.1D and E) when compared to using a combination of all the five reagent sequences as shown by the higher overall coverage at 68.31% (Figure 2.1F and G). Overall, 9-mer matched coverage was high within the p24 region compared to p17 and p2 regions of Gag protein (Figure 2.1D and F), as expected (Appendix C, Figure C1-C5).

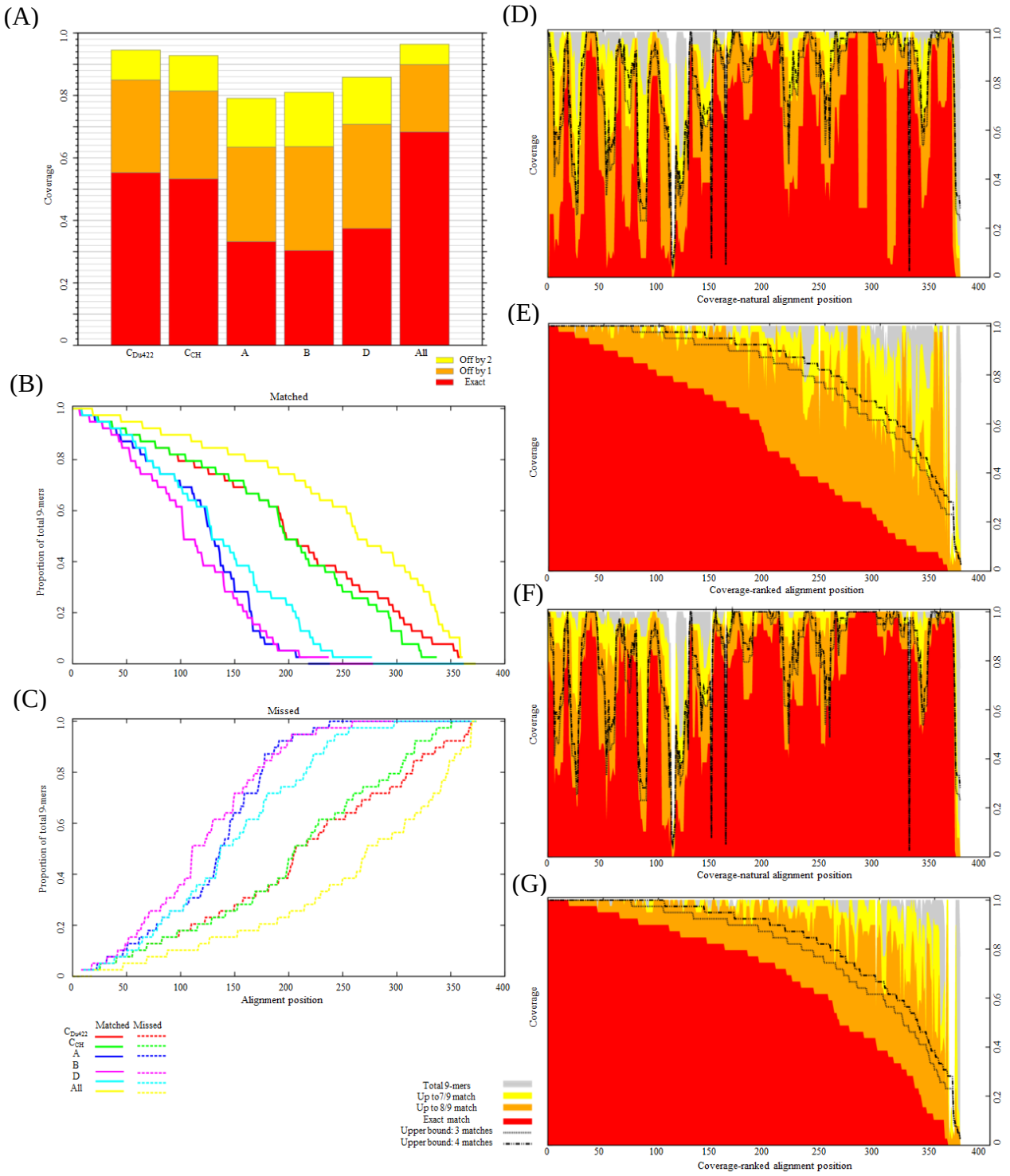


Figure 2.1. The coverage of putative epitopes of current vaccine inserts and consensus sequences in 39 HIV-1 subtype-C infected study participants. (A) Epicover: Epitope coverage of 39 HIV-1 subtype C viruses by two vaccine candidates (C_{Du244} and C_{CH}), subtype B and two consensus sequences (A and D). (B-G) Posicover: **(B)** Nine-mers matched to of five peptide sets in the 39 study participants ranked by coverage. **(C)** Nine-mers missed by use of the five peptides sets in the 39 study participants ranked by coverage. **(D)** Nine-mers matched in their natural position and **(E)** ranked by coverage, by using C_{Du422} vaccine candidate. **(F)** Nine-mers matched for five peptide sets in their natural position and **(G)** ranked by coverage. Analysis was based on the p17p24p2 region of the Gag protein.

2.3.3. Preferential recognition of the infecting clade by HIV-specific T-cells

The first level of cross-clade analysis involved identification of whether the magnitude of ELISPOT responses was equally distributed between the different peptide sets tested. All comparisons between peptide sets were based on the p17p24p2 region of the Gag protein. For the C_{Du422} and C_{CH} sets, there was no difference between the median magnitude of response expressed as SFU/10^6 PBMC, at 2690 (range 0 to 24550) for C_{Du422} and 2828 (range 0 to 19407) for C_{CH} ($p > 0.05$; Figure 2.2A). Conversely, the magnitude of responses to other clades were significantly lower, at 750, 810 and 1390 for clades A, B and D respectively ($p<0.05$; Figure 2.2A). A similar trend was observed for the breadth of responses, with a larger median number of reactive peptides being recognized in the two clade C based peptide reagents, C_{Du422}; 4 (range 1-11) and C_{CH}; 4 (range 0-10) when compared to clades A; 1 (range 0-6), B; 2 (range 0-6) and D; 3 (range 0-8) peptide sets (Figure 2.2B). Collectively, these data show that peptides more closely matched to the infecting autologous sequence result in high magnitude responses and wider breadth of coverage, consistent with the predicted epitope coverage shown in Figure 2.1.

The second level of cross-clade analysis involved assessing the reactivity of each peptide at the single study individual level and the contribution of each peptide set to the total response. An example is shown for study individual CC23 (Figure 2.3). This study individual demonstrated substantial cross-reactivity to different peptides, recognizing a total of 11 peptides. When the sequences of these peptides were taken into consideration, the 11 peptides made up 39 different variants, of which 22 variants were reactive. Six out of eleven peptides were cross reactive (recognized in at least two peptide sets) with differing degrees of cross reactivity. Each peptide set contributed the following responses; 8, 7, 4, 5 and 8 for C_{Du422}, C_{CH}, A, B and D, respectively. Peptides 19, 63 and 80 were the most cross-reactive peptides in this individual, being recognized in all the four clades, with two variants having the same sequence for peptides 63 and 80. In addition, these two peptides had the highest magnitude of response in this study individual (6670 SFU/10^6 PBMC and 7470 SFU/10^6 PBMC, respectively, Figure 2.3). The reasons for different variants being recognized or not are explored further in section 2.3.5.

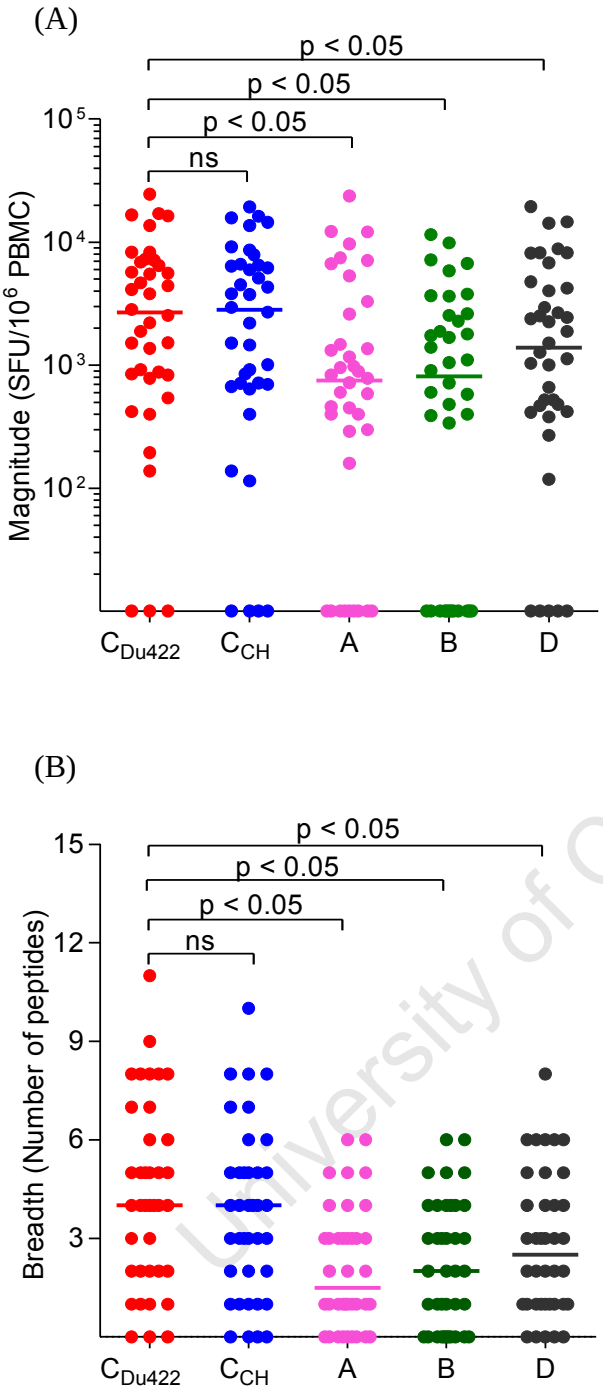


Figure 2.2. Recognition of HIV-1 Gag peptides from clades A, B, CDu422, CCH and D sequences in HIV-1 clade C infected individuals from South Africa. (A) Total magnitude of Gag-specific T-cells (SFU/10^6 PBMC) against individual peptides for each study individual. Values are shown in log scale. **(B)** The minimum number of responses per peptide set for each study individual. Data is based on the p17p24p2 region of the Gag protein. Significant differences were tested for using one-way ANOVA followed by a Dunn's post-test.

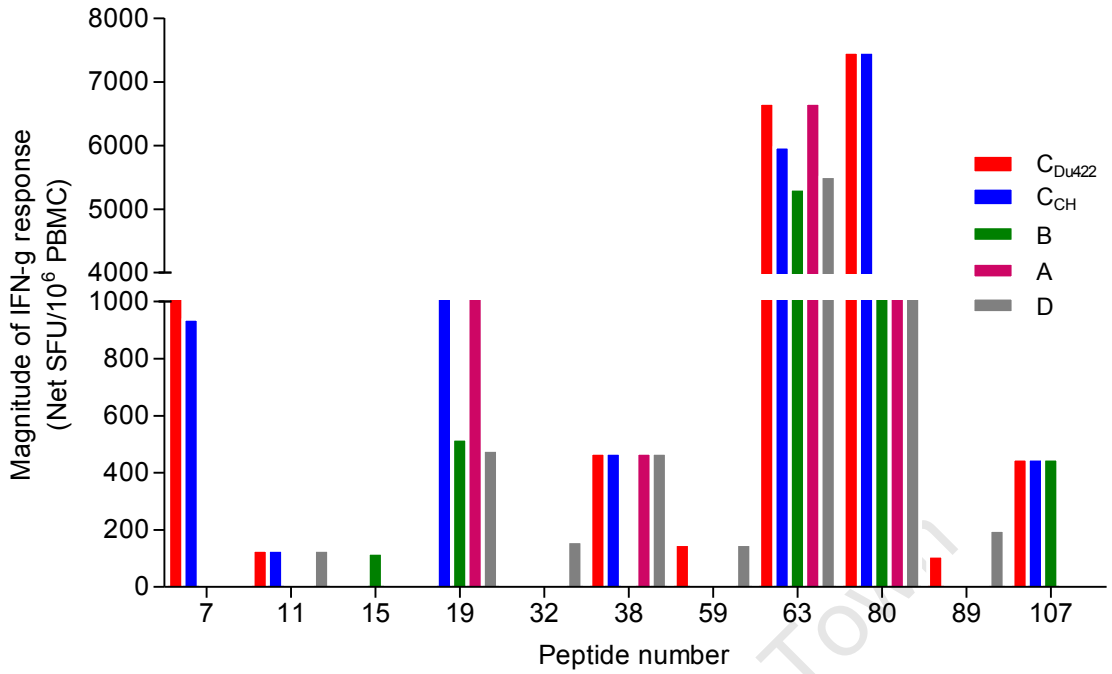


Figure 2.3. Comprehensive assessment of HIV-1-specific T-cells cross-reactive among the clades A, B, C_{Du422}, C_{CH} and D sequences at the single peptide level. The recognition of peptides from the clade A, B, C_{CH}, C_{Du422} and D sequences in study individual CC 23. Peptide numbers are based on the subtype C_{Du422} numbers.

Similar analysis was performed for all study individuals and for all peptides that were reactive in the study. In the test peptides used for the ELISpot assay (a total of 540 peptides), there were 26 peptides that had similar sequences between two variants (6 between A and D, 6 between C_{Du422} and C_{CH}, 1 between C_{CH} and B. 2 between C_{Du422} and D, 5 between C_{Du422} and A, 6 between C_{Du422} and C_{CH} and between A and D peptide variants). A total of 22 peptides had similar sequences among three variants (8 between C_{Du422}, C_{CH} and D variants, 3 between C_{Du422}, C_{CH} and A variants, 6 between Du422 C_{CH} and B variants and 5 between C_{Du422}, A and D variants).

Peptide reactivity was normalized for the size of the protein into p17, p24 and p15 (for peptides sets C_{Du422}, C_{CH} and B which had p15). Overall, reactivity normalized per amino acid was higher for p24 (0.20) when compared to p17 (0.18) and p15 (0.10, Table 2.4).

Table 2.4. Immunogenicity and cross-reactivity of the different HIV-1 Gag regions

Gag region	[a]Number of amino acids	Number of peptides	Number of reactive peptides	[b]Reactivity per peptide	[c]Reactivity per amino acid	[d]Cross-reactivity per peptide
p17 (MA)	131	30	24	0.80	0.18	2.5
p24 (CA)	231	59	47	0.79	0.20	3.2
p15 (NC)	137	22	14	0.67	0.10	2.0

[a]The number of amino acids making up the Gag region.
[b]The proportion of peptides that are reactive given all the peptides tested for that particular region.
[c]The reactivity of the Gag region normalized to per amino acid making up that particular region.
[d]The average number of variants recognized per peptide for that particular region.

Reactive peptides that were recognized by at least five study participants were classified as immunodominant peptides and are shown in Table 2.5. The location of these peptides in the three regions of the Gag protein and the previously determined restricting HLA alleles in the study individuals are shown. A total of 15 immunodominant peptides were identified in the study using this analysis. The magnitude of response of these immunodominant peptides ranged from 480-10150 SFU/10^6 PBMC (Table 2.5). Four out of 15 (27%) of these peptides were from the p17 region of the Gag protein and 11/15 (73%) were from the p24 region of Gag protein. There were no peptides from the p15 region of Gag that were classified as immunodominant peptides. Peptides 10 and 33 contain epitopes restricted by HLA alleles previously shown to be protective, while peptide 77 has an epitope restricted by an allele previously shown to be associated with rapid progression. Of these peptides, 13/15 (87%) were low entropy peptides and 2/15 (13%) were high entropy peptides (data not shown).

2.3.4. Peptide variability and cross-reactivity

The second level of analysis consisted of identifying the numbers of mutually (cross-reactive) and exclusively recognized (recognition by one clade only) peptides within the five peptide sets tested for reactivity. The study participants recognized a total of 84 peptides, with 29 peptides being exclusively recognized in one clade only. Nineteen peptides were exclusive to clade C (either C_{Du422} or C_{CH}), 8 to clade B, and one peptide each to clades A and D peptide sets (Figure 2.4A). Between the two clade C peptide sets, 6/19 peptides were recognized exclusively in the C_{Du422} peptide set and 13/19 peptides were common between the two peptide sets. There were 17 peptides that were mutually recognized across all four clades. Of these, 7 were positioned in p17 and 10 within the p24 region of the Gag protein (Figure 2.4B). The remainder of the peptides was recognized in two or three of the clades.

Table 2.5. Immunodominant peptides recognized in the study

Peptide no	Peptide sequence	Peptide location	[a]HLA allele	[b]Cross-reactivity	[c]SFU/10^6 PBMC	Reference
5	KIRLRPGGKKHYMLK	p17	Cw*4	5	660	Geels *et al.*, 2005
6	RPGGKKHYMLKHIV	p17	A*2301	3	1360	Kiepiela *et al.*, 2007
10	SRELERFALNPGLLE	p17	B*27	4	1750	Shellens *et al.*, 2008
19	ELKSLYNTVATLYCV	p17	A*2902	4	5375	Llano *et al.*, 2009
33	PIVQNLQGQMVHQAI	p24	B*13	4	2130	Llano et al., 2009
35	MVHQAISPRTLNAWV	p24	Cw*0602	5	1120	Allen *et al.*, 2005a
37	RTLNAWVKVIEEKAF	p24	A*0201	4	1090	Schaubert *et al.*, 2007
44	GATPQDLNTMLNTVG	p24	B*4201	4	6430	Day *et al.*, 2007
61	EQIAWMTSNPPIPV	p24	A*0201	3	480	McKinney *et al.*, 2004
53	PPIPVGDIYKRWILL	p24	B*0801	5	6670	Gillespie *et al*., 2007
73	RDYVDRFFKTLRAEQ	p24	B*1510	5	10150	Gray *et al.*, 2009
74	DRFFKTLRAEQATQE	p24	B*1401	5	6965	Mathews *et al.*, 2008
75	KTLRAEQATQEVKNW	p24	Cw*0304	4	3990	Masemola *et al.*,
77	TQEVKNWMTDTLLVQ	p24	B*53	2	520	Kaul *et al.*, 2001a
83	LRALGPGATLEEMMT	p24	B*7	3	3830	Perez *et al.*, 2008

[a]The restricting HLA allele in the individual that gave the highest magnitude of response to the peptide out of the individuals who responded to the peptide.
[b]The cross-reactivity is given by the number of variants recognized for that particular peptide
[c]The magnitude of IFN-γ response shown in the table is the magnitude of response to the variant that gave the highest response for that particular peptide out of all variants that were reactive in the study individuals.

In an attempt to understand the basis of mutual or exclusive recognition of peptides, we compared the Shannon entropy score for the 17 peptides cross-recognized and the 19 peptides exclusively recognized from the clade C peptide sets (Figure 2.5A). It was evident that peptides exclusively recognized had significantly higher entropy than the mutually recognized peptides ($p = 0.0128$). The bulk of these high entropy peptides were also the least recognized within the cohort (Figure 2.5B). These data suggest that exclusive recognition of peptides is related to clade C-specific variability within the epitopes. Not all peptides exhibited this pattern of low entropy and mutual recognition or high entropy and exclusive recognition, and Table 2.6 shows a representative example of an individual in the cohort (CC23) recognizing 11 peptides, in some cases despite extensive amino acid variability. Peptides 63 and 80 are examples of mutually recognized peptides regardless of amino acid change D→E and T→S within the each of the peptides, respectively (Table 2.6). These variations are most likely tolerated as they fall outside HLA anchor motifs or T-cell receptor (TCR) contact residues. In contrast, peptides 7, 15 and 32 were

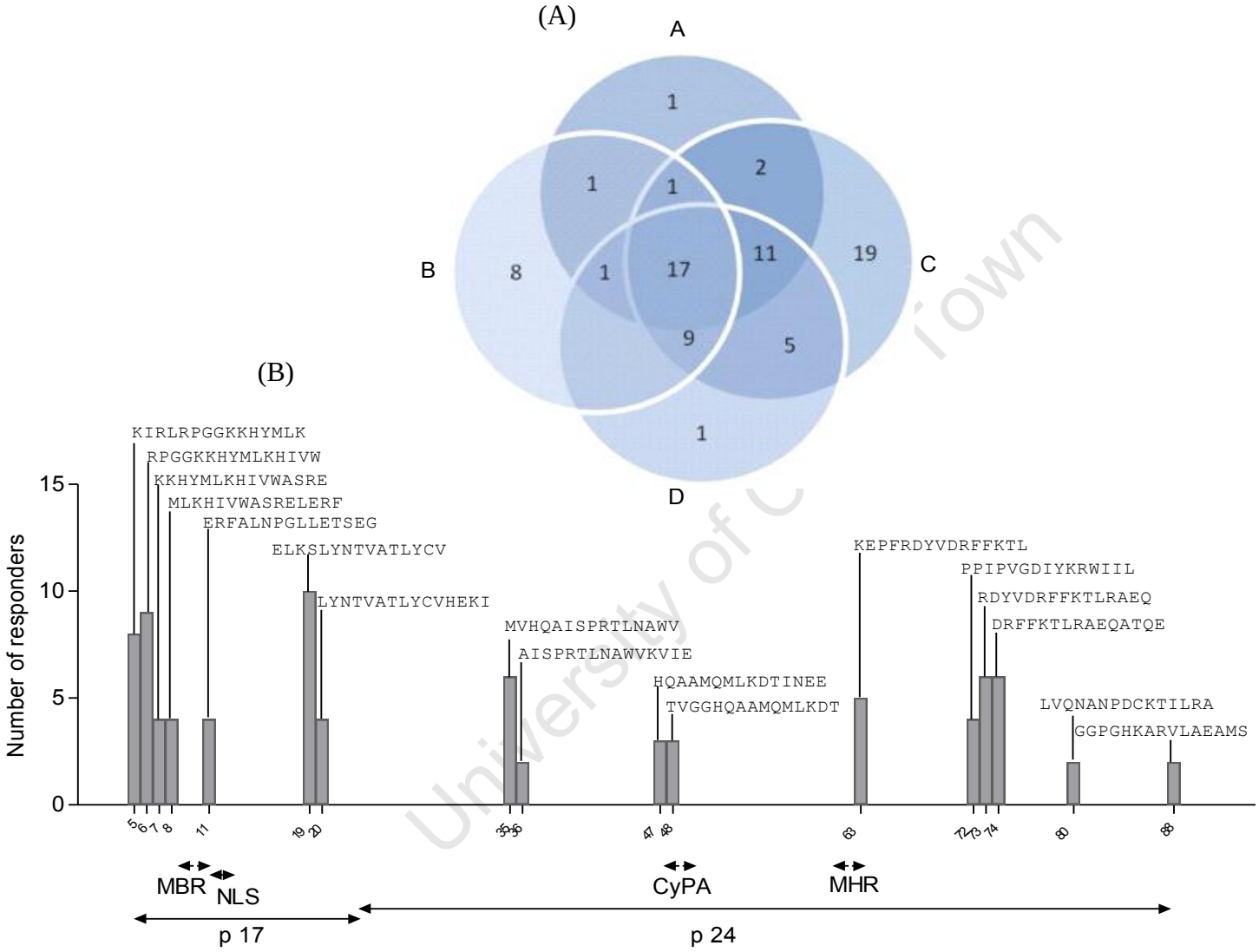


Figure 2.4. Characterization of reactive peptides. (A) The contribution of each clade-specific peptide set to the overall HIV-specific T-cell responses. Shown in each portion of the Venn diagram are the numbers of peptides recognized in each clade and those common among the different combination of clades**. (B)** HIV-1 Gag cross-reactive epitope hotspots. The Gag region in which the peptides are located is shown and drawn to scale, MBR: membrane binding region; NLS: nuclear localization signal; CyPA: Cyclophilin A binding region; and MHR: major homology region. The two clade C sequences were combined as C for this analysis.

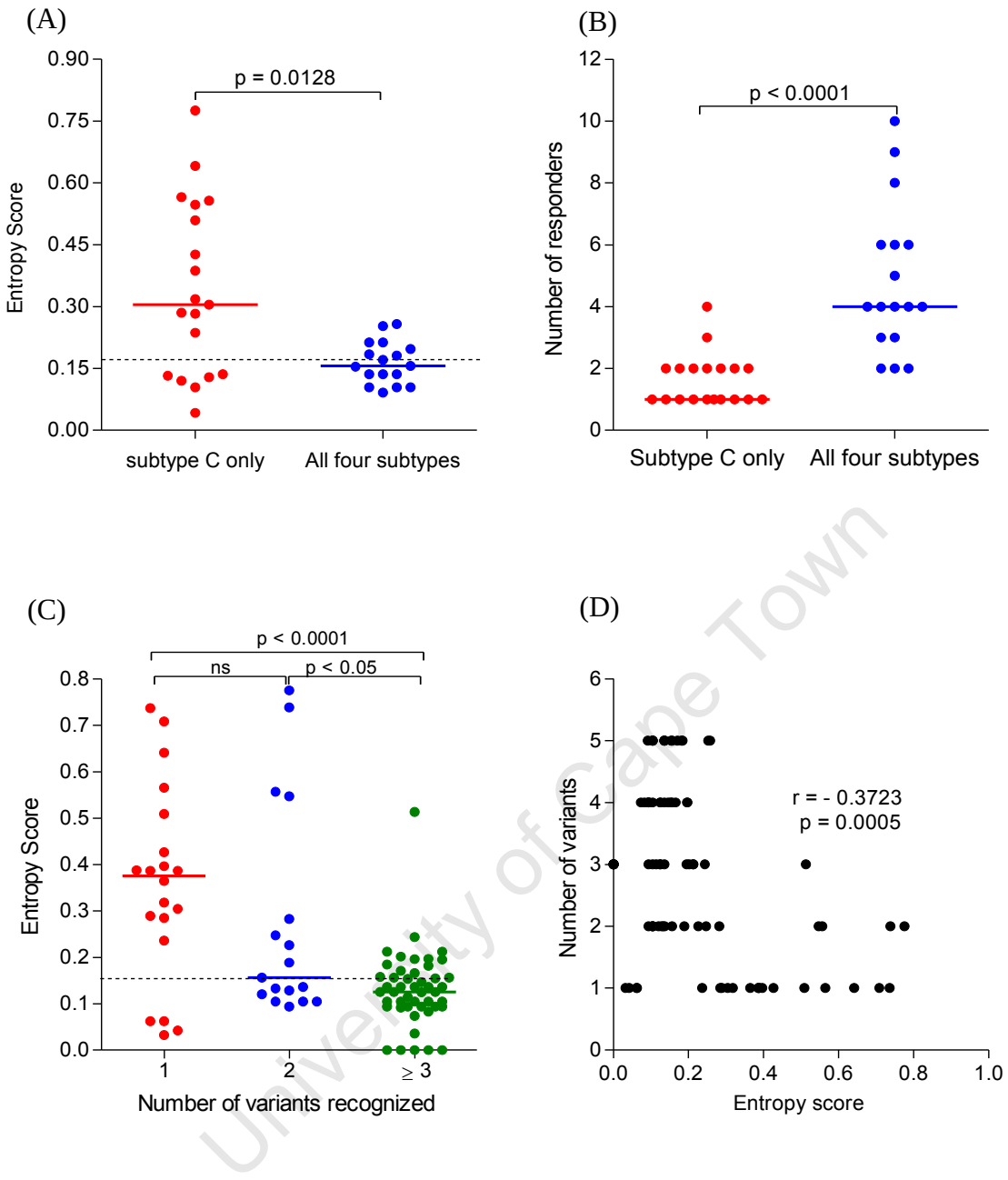


Figure 2.5. Characterization of peptide cross-reactivity and the basis for mutual and exclusive recognition. **(A)** Comparison of entropy scores between peptides recognized exclusively by clade C peptide reagents (red) and those mutually recognized in all the four clades tested (blue). **(B)** Frequency of responders to epitopes recognized by clade C reagents (red) compared to those mutually recognized in all four clades (blue). **(C)** Comparison of entropy score of peptides with different levels of cross-reactivities. Peptides were categorized into those recognized in one (red), two (blue) or three or more (green) variants of the peptide sets. **(D)** Association between number of variants recognized and entropy score, determined by Spearman correlation. The non-parametric one way ANOVA and Kruskal-Wallis test was used to test for differences in the median entropy scores among the different recognition categories.

exclusively recognized due to variation in residues important for recognition, and hence not cross-reactive (Table 2.6). Peptides that were reactive despite differences from the infecting viral sequence may have had variations in 'tolerated' residues for peptide binding and conformation, or the epitopes may indeed have been presented in infected persons as a result of minor viral variants that we did not detect by population sequencing of the dominant virus.

To have an overall picture of the pattern of recognition of reactive peptides in the study, all the reactive peptides from the study were categorized into those that are recognized in 1, 2 or ≥3 variants and their entropy scores compared. Reactive peptides recognized in 2 or ≥3 corresponding variants had significantly lower entropy scores when compared to peptides recognized in only one peptide variant ($p < 0.0001$ and $p < 0.05$ respectively, Figure 2.5C). Furthermore, there was a significant negative correlation between the entropy score and number of variants recognized for each peptide when all reactive peptides in the study were assessed ($r = -0.3723$, $p = 0.0005$, Figure 2.5D).

A graphical representation of these analyses for all the peptides recognized in the study, showing the number of responders to each reactive peptide, the number of variants recognized for each reactive peptide and the entropy score of the particular reactive peptide is shown in Figure 2.6A, B and C respectively. It is evident from this representation that there is no general trend for the number of responders to each reactive peptide in relation to entropy or number of corresponding variants recognized (Figure 2.6A). In other words, there was no correlation between immunodominance of peptides and their entropy ($r = 0.0238$, $p = 0.8297$; data not shown). However, peptides that were recognized in more variants (generally three or more) matched those peptides with lower entropy score, as mentioned above (Figure 2.6B and C).

Table 2.6. Reactive and corresponding non-reactive peptides in study individual CC 23

Peptide number	Name	Sequence	Restricting HLA	SFU/10^6 PBMC
[a]7	Autologous	GKKRYMLKHIVWASRE	A*0201	
	C_{Du422}	KKHYMLKHIVWASRE		1750
	C_{CH}	KKHYMLKHLVWASRE		940
	B	GKKKYKLKHIVWASR		0
	A	KKKYRLKHLVWASRE		0
	D	KKKYRLKHLVWASRE		0
[a]11	Autologous	EKFALNPGLLETSDG	A*0201	
	C_{Du422}	ERFALNPGLLETSEG		130
	C_{CH}	ERFALNPGLLETSEG		130
	B	LERFAVNPGLLETSE		0
	A	ERFALNPSLLETAEG		0
	D	ERFALNPGLLETSEG		130
15	Autologous	KQIIKQLQPALQTGT	A*02	
	C_{Du422}	KQIMKQLQPALQTGT		0
	C_{CH}	KQIIKQLQPALQTGT		0
	B	CRQILGQLQPSLQTG		120
	A	CQQIMEQLQSALKTSE		0
	D	KQIIGQLQPAIQTGS		0
19	Autologous	EELKSLFNTVATLYCV	B*0801	
	C_{Du422}	ELKSLYNTVATLYCV		0
	C_{CH}	ELRSLFNTVATLYCV		1280
	B	EELRSLYNTVATLYC		520
	A	LKSLFNTVATLYCVH		2353
	D	ELRSLYNTVATLYCV		480
[b]32	Autologous	SQVSQNYPIVQNLQGQMV	Unknown	
	C_{Du422}	SQNYPIVQNLQGQMV		0
	C_{CH}	SQNYPIVQNLQGQMV		0
	B	SQNYPIVQNLQGQMV		0
	A	KVSQNYALKHRAYEL		0
	D	SQVSQNYPIVQNLQG		160
38	Autologous	NAWVKVIEEKAFSPEI	B*4501	
	C_{Du422}	AWVKVIEEKAFSPEV		470
	C_{CH}	AWVKVIEEKAFSPEV		470
	B	NAWVKVVEEKAFSPE		0
	A	AWVKVIEEKAFSPEV		470
	D	AWVKVIEEKAFSPEV		470
59	Autologous	AGTTSTLQEQIAWMTS	A*0201	
	C_{Du422}	GTTSTLQEQIAWMTS		150
	C_{CH}	GTTSTLQGQIAWMTS		0
	B	AGTTSTLQEQIGWMT		0
	A	GTTSTLQEQIGWMTS		0
	D	GTTSTLQEQIAWMTS		150
63	Autologous	NPPIPVGEIYKRWIIL	B*0801	
	C_{Du422}	PPIPVGDIYKRWIIL		6670
	C_{CH}	PPVPVGEIYKRWIIL		5980
	B	NPPIPVGEIYKRWII		5320
	A	PPIPVGDIYKRWIIL		6670
	D	PPIPVGEIYKRWIIL		5520
80	Autologous	LLTQNANPDCKTILRA	B*0801	
	C_{Du422}	LVQNANPDCKTILRA		7470
	C_{CH}	LVQNANPDCKTILRA		7470
	B	LLVQNANPDCKTILK		3920
	A	LVQNANPDCKSILRA		2653
	D	LVQNANPDCKTILKA		1240
89	Autologous	GHKARVLAEAMSQVGH	A*0201	
	C_{Du422}	HKARVLAEAMSQTNS		110
	C_{CH}	HKARVLAEAMSQANG		0
	B	GHKARVLAEAMSQVT		0
	A	HKARVLGTGARASVL		0
	D	HKARVLAEAMSQATN		200
107	Autologous	FLGKIWPSHKGRPGN	A*0201	
	C_{Du422}	FLGKIWPSHKGRPGN		450
	C_{CH}	FLGKIWPSHKGRPGN		450
	B	FLGKIWPSHKGRPGN		450
	A	Not available[c]		N/A
	D	Not available		N/A

[a]The previously described HLA allele to restrict the epitope was absent in the individual, however HLA A*0201 was found to be a strong binder to the epitope within this peptide with a binding affinity of 5nM for peptide 7 and a weak binder with an affinity of 135nM for peptide 11, using NetMHC. [b]No epitope described nor predicted to bind to HLA alleles of this participant in this peptide. [c]Clades A and D had no p15 region of the Gag protein. Some peptide variants with substitutions in regions flanking the epitope showed discordant recognition patterns regardless of matching epitope sequence, possibly due to additional non hydrophobic amino acids at the C-terminal or N-terminal that are not well tolerated by class I alleles [63]. ELISPOT reactivity is shown for variants that were reactive, while 0 denotes those that were not reactive. Red letters in the peptide sequence indicate amino acid mismatches between that peptide variant and the C_{Du422} sequence. The green letters show amino acid mismatches between the peptide variant sequence and the infecting virus. The predicted epitope in each variant is underlined. SFU: Spot Forming Units (per 10^6 PBMC).

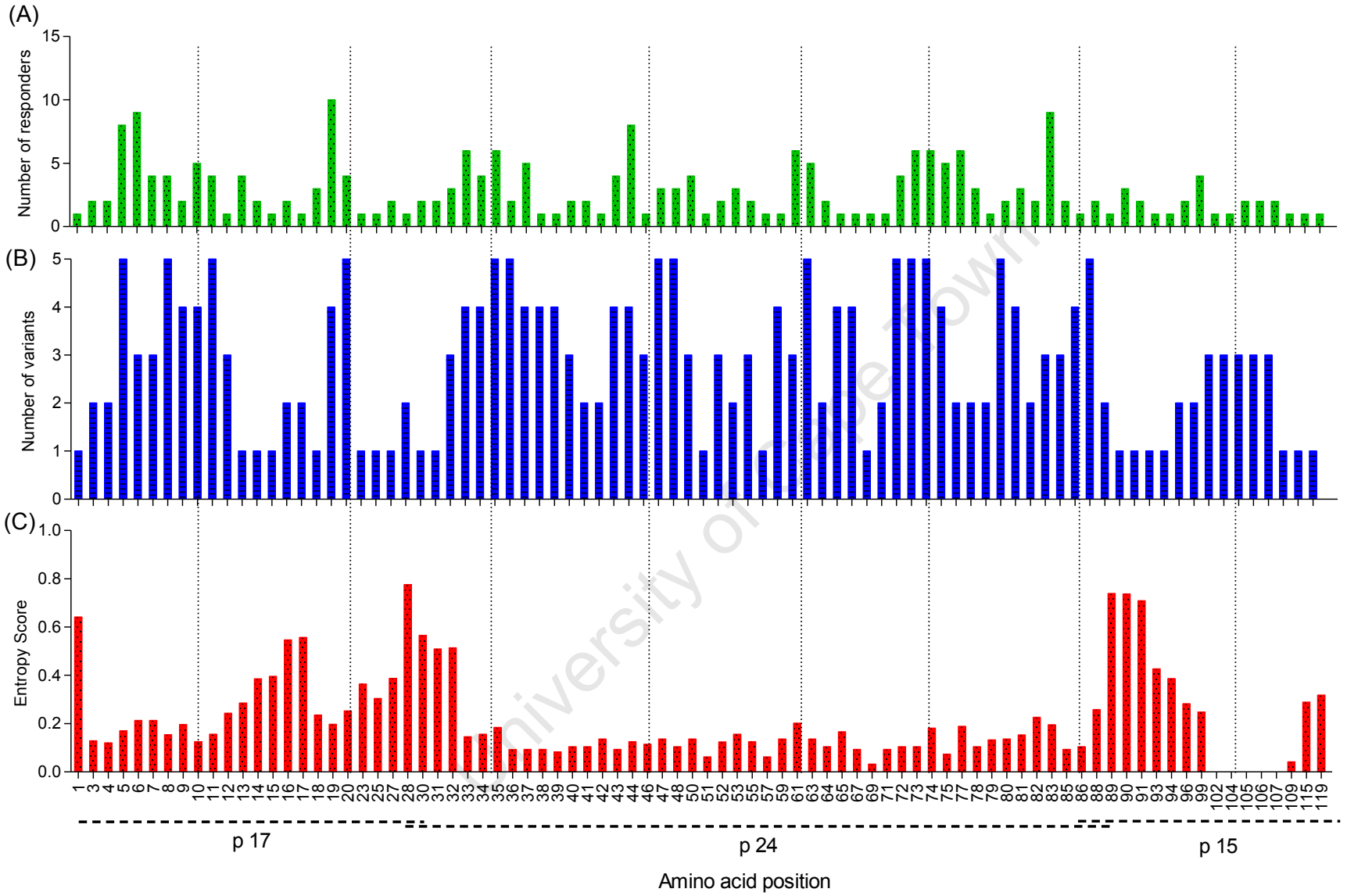


Figure 2.6. Reactive peptides from all the study individuals. (A) Number of individuals responding to each peptide. **(B)** Number of peptide variants recognized for each reactive peptide. **(C)** Average entropy score of the five peptide variants for each peptide recognized.

2.3.5. The Impact of Host HLA on Epitope Recognition by HIV-1-specific T-cells

To test the hypothesis that the level of peptide recognition was most likely governed by mutations in key residues that are associated with peptide processing and binding to restricting HLA molecules, an algorithm prediction tool was applied (www.immuneepitope.org). The tool assesses predicted scores for Transport Associated with Antigen Processing (TAP) binding (an estimate of the affinity of the peptide with the TAP molecule), MHC binding scores (an estimate of the efficiency of binding to an MHC molecule), proteasome scores (an estimate of cleavage site usage) and processing scores (an estimate of the quantity of peptide present in the endoplasmic reticulum that is available for MHC binding, from a combination of cleavage and transport predictions), and was applied to those participants whose HLA class I A and B alleles were typed in this study and were also available in the database (Table 2.3).

The total epitope score, which is a summary of the proteasomal, TAP and MHC scores, as well as the MHC IC_{50} of the epitopes restricted by the predicted HLA, are shown in Table 2.7 (shown for study individual CC23). There were higher scores for proteasome activity, TAP and MHC binding, and lower MHC IC_{50} in peptide variants that were recognized in the IFN-γ ELISPOT assay. Overall, when peptide variants were classified into those that were recognized in the ELISPOT assay and those that were not, MHC binding score of reactive peptides was significantly higher than that of non-reactive variants ($p<0.0001$, Figure 2.7A). In addition, proteasomal cleavage scores (Figure 2.7B), TAP scores (Figure 2.7C) and processing scores (Figure 2.7D) were significantly higher for reactive peptide variants compared to their non-reactive counterparts ($p=0.0102$, $p=0.0427$ and $p=0.0161$, respectively). When all these scores were summated, the total score of the reactive peptides was significantly higher than non-reactive peptide variants ($p=0.0001$, Figure 2.7E). Furthermore, MHC IC_{50} scores of the reactive variants were significantly lower for reactive variants when compared to their corresponding non-reactive variants ($p<0.0001$, Figure 2.7F). Thus, non-recognized peptides, with higher entropy, may have at some stage mutated under immune-mediated selection pressure, showing that variability in key residues is important for MHC presentation and recognition by the TCR. Overall, these data provide support for the notion that different clades of HIV-1 may have been shaped by class I HLA restricted epitope diversity through probable selective immune pressures in different populations. However, it cannot be discounted that the non-reactivity of these peptides might be due to random sequence variability, unrelated to immune selection, and that these regions may never be recognized as epitopes.

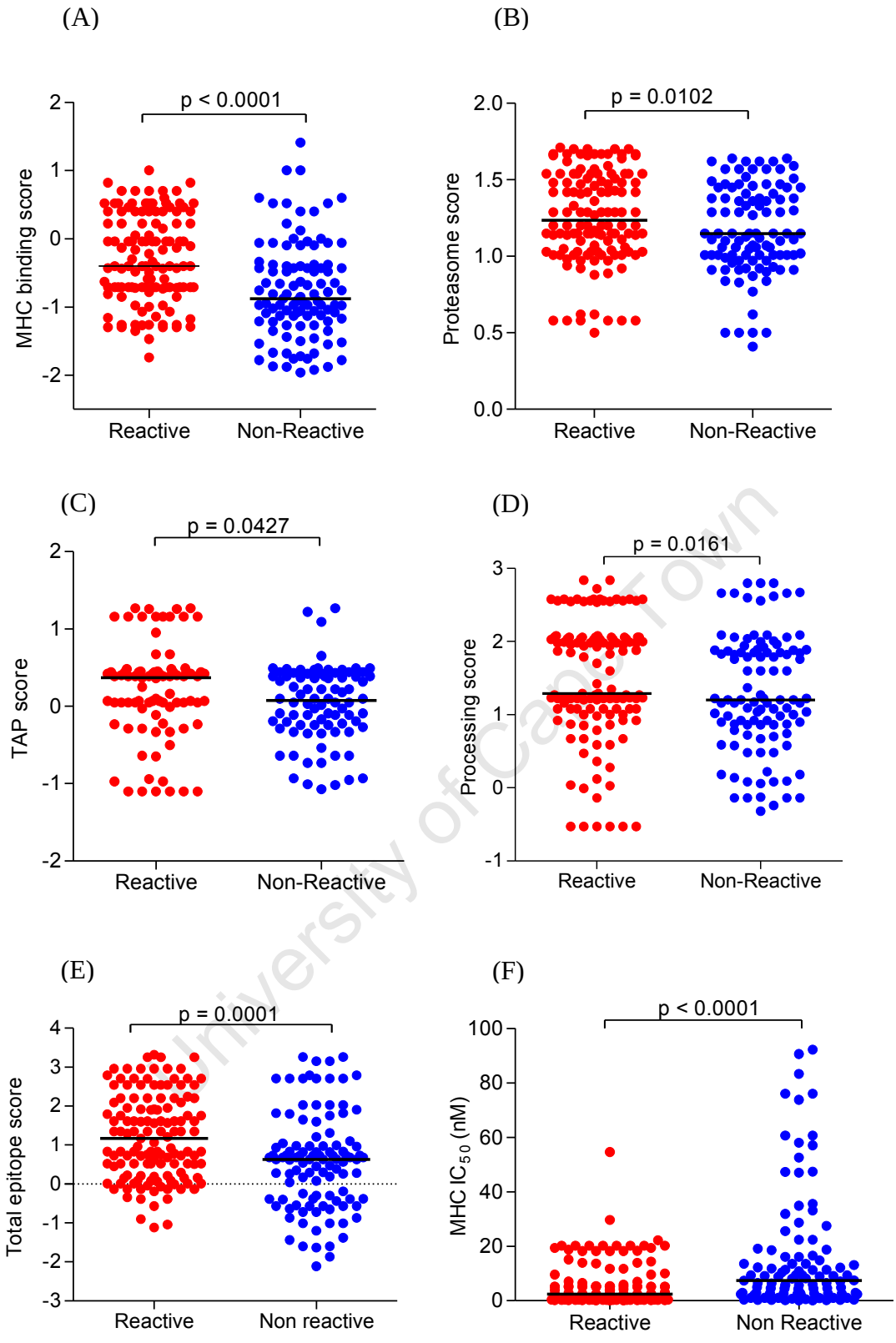


Figure 2.7. Characterization of peptide processing and MHC class I presentation of epitopes and their variants. Comparison of **(A)** Major Histocompatibility Complex (MHC) binding score, **(B)** proteasome cleavage score, **(C)** transporter associated with antigen processing (TAP) score, **(D)** processing score, which combines proteasomal cleavage and TAP transport scores, **(E)** total epitope score, and **(F)** MHC IC_{50}, in nM. The total epitope score is a summary of the proteasomal, TAP and MHC binding scores and between reactive peptide variants (red) and their corresponding non-reactive variants (purple). Peptides binding to HLA class I A and B alleles had their total processing scores and MHC IC_{50} predicted using the Immune Epitope Database (www.immuneepitope.org), and scores were compared using the non-parametric Mann-Whitney test.

Table 2.7. The effect of amino acid mutation on predictions of epitope processing

Peptide	Peptide	Peptide	SFU/10^6 PBMC	Proteasome score	TAP score	MHC binding	Processing score	Total score	MHC
7	C_{Dv422}	-KKHYMLKHIVWASRE	1750	1.1	-0.23	0.7	0.87	1.57	0.20
	C_{CH}	-.....L.....	940	1.1	-0.23	0.7	0.86	1.56	0.20
	B	G..K.K.....	0	1.1	-0.24	-1.12	0.86	-0.25	13.1
	A	-..K.R...L.....	0	1.1	-0.19	-1.21	0.90	-0.30	16.1
	D	-..K.R...L.....	0	1.1	-0.19	-1.21	0.90	-0.30	16.1
11	C_{Dv422}	-ERFALNPGLLETSEG	130	0.97	-0.29	-0.59	0.67	0.08	3.9
	C_{CH}	-..............	130	0.97	-0.29	-0.59	0.67	0.08	3.9
	B	L....V.....	0	0.97	-0.27	-1.09	0.70	-0.39	12.3
	A	-.......S....A..	0	0.96	-0.29	-0.63	0.67	0.04	4.3
	D	-..............	130	0.97	-0.29	-0.59	0.67	0.08	3.9
15	C_{Dv422}	-KQIMKQLQPALQTGT	0	0.84	-0.35	-1.35	0.48	-0.87	22.4
	C_{CH}	-...I..........	0	0.84	-0.35	-1.35	0.48	-0.87	22.4
	B	CR..LG....S....	120	1.01	-0.65	-1.47	0.36	-1.12	29.6
	A	CQ...E...S..K.SE	0	0.87	-0.73	-1.52	0.14	-1.38	33.1
	D	-...IG.....I...S	0	0.90	-1.02	-1.50	-0.13	-1.63	31.9
19	C_{Dv422}	-ELKSLYNTVATLYCV-	0	1.15	0.05	-0.38	1.20	0.82	2.40
	C_{CH}	-..R..F.....	1280	1.15	0.07	0.52	1.22	1.75	0.30
	B	E..R.....	520	1.15	0.07	-0.40	1.22	0.83	2.50
	A	--....F.....H	2353	1.15	0.05	0.40	1.20	1.60	0.40
	D	-..R.....	480	1.15	0.07	-0.40	1.22	0.83	2.50
38	C_{Dv422}	-AWVKVIEEKAFSPEV	470	1.03	0.05	0.52	1.08	1.61	0.30
	C_{CH}	-..............	470	1.03	0.05	0.52	1.08	1.61	0.30
	B	N.....V.....	0	0.41	-0.73	-0.88	-0.32	-1.20	7.50
	A	-..............	470	1.03	0.05	0.52	1.08	1.61	0.30
	D	-..............	470	1.03	0.05	0.52	1.08	1.61	0.30
63	C_{Dv422}	-PPIPVGDIYKRWIIL	6670	1.54	0.45	-0.44	1.99	1.95	1.10
	C_{CH}	-..V...E.....	5980	1.54	0.44	0.22	1.98	2.20	0.60
	B	N......E.....	5320	1.54	0.44	0.22	1.98	2.20	0.60
	A	-..............	6670	1.54	0.45	-0.44	1.99	1.95	1.10
	D	-......E.....	5520	1.54	0.44	0.22	1.98	2.20	0.60
80	C_{Dv422}	-LVQNANPDCKTILRA	7470	1.67	0.39	-0.71	2.06	1.35	5.20
	C_{CH}	-..............	7470	1.67	0.39	-0.71	2.06	1.35	5.20
	B	L.............K	3920	1.67	0.39	-0.71	2.06	1.35	5.20
	A	-.........S....	2653	1.17	0.25	-1.26	1.42	0.16	18.4
	D	-.............K.	1240	1.67	0.39	-0.71	2.06	1.35	5.20
89	C_{Dv422}	-HKARVLAEAMSQTNS	110	0.94	-0.94	-0.22	-0.01	0.22	0.60
	C_{CH}	-.............A.G	0	0.77	-0.54	0.12	0.22	0.35	0.80
	B	G.............VT	0	1.05	0.32	1.41	0.72	0.68	25.6
	A	-......GTGARASVL	0	1.52	0.33	-1.76	1.85	0.08	58.0
	D	-.............ATN	200	0.98	-0.50	0.15	0.47	0.63	0.70
107	C_{Dv422}	FLGKIWPSHKGRPGN	450	0.58	-1.10	0.46	-0.53	-0.13	0.40
	C_{CH}	....	450	0.58	-1.10	0.46	-0.53	-0.13	0.40
	B	....	450	0.58	-1.10	0.46	-0.53	-0.13	0.40

The different stages of epitope processing were predicted using algorithms for MHC class I T-cell epitope processing for reactive peptides in the study. Results shown in the table are for individual CC23. For peptide 107, the A and D variants are missing due to the absence of the p15 region in these peptide reagent sets as noted in the text.

2.4. DISCUSSION

Immunogens designed to elicit T-cell responses remain a major focus of HIV vaccine development. Because of the significant sequence variation that exists between HIV-1 clades (7-15%, (Korber *et al*., 2001), there is an acknowledgement that the HIV clades on which vaccines are based will have an impact on the immune response elicited, and very likely the subsequent efficacy of vaccines.

The predicted correlates of protection (neutralizing antibodies and CTL) might be changing with correlates analysis from the RV144 trial, which has shown that V2 binding antibodies and CD4 responses were present in those vaccinated and protected from infection (Haynes *et al*., 2012). These data highlight the need for empirical studies. A vaccine to prevent infection has to address the issue of the high mutation rate of HIV and hence the huge genetic diversity of the virus. Therefore approaches that can predict possible mutation states and highly mutable regions of HIV will be useful for designing these kinds of vaccines. Using such methods, perhaps it will be possible to predict which regions are likely to mutate to which amino acids in people expressing particular HLA alleles, and then design immunogens that address that. The concept of cross-reactivity is important for both vaccine designs as in either case a vaccine has to encompass the ever increasing genetic diversity of the virus. Therefore, testing HIV-infected persons serves as a proxy for assessing T-cell cross-reactivity of highly immunogenic vaccines. It is useful for determining whether HIV-specific responses that particular populations with specific HLA alleles mount, and their degree of cross-reactivity to vaccine sequences that may be tested in that region; infecting viruses represent future 'challenge' viruses that vaccine trial volunteers may encounter. Previous studies have shown that HIV-infected individuals can mount cross-reactive T-cell responses against different HIV-1 clades (Buseyne *et al*., 1998; Cao *et al*., 2000; Currier *et al*., 2003; Ferrari *et al*., 1997). While early studies focused on a limited number of selected epitopes, and relied on the use of pools of peptides or cells infected with recombinant vaccinia virus expressing HIV proteins, more recent studies have assessed the degree of cross-clade recognition at the peptide or epitope level Aidoo *et al*., 2008; Gupta *et al*., 2006; Yu *et al*., 2005). This chapter describes a study which is the first to comprehensively look at the ability of clade C infected individuals to recognize peptides included in vaccines currently being tested, and to investigate both intra-clade and inter-clade cross recognition.

This study investigated T-cell reactivity in 39 individuals where the sequence of the infecting virus was determined by population sequencing of the dominant virus, and high resolution HLA typing was performed in a sub-group of these individuals. Although it was not directly shown that responses were due to CD8+ T-cells, reactivity was assumed to be CD8-mediated as previous studies have shown that ELISpot responses are predominantly CD8+ T-cell mediated (Addo *et al*., 2003; Bansal *et al*., 2003). However, it cannot be ruled out that some of the responses may have been CD4-mediated. In this study, responses against Gag peptides from five sequence variants, clade C_{Du422} from South Africa based on the Du422 sequence (Williamson *et al*., 2003), C_{CH} (Chinese clade C, for intra-clade responses), and clades B, A and D, at the single peptide level were assessed and were in the same order as predicted by coverage algorithms with slight differences possibly due to the fact that recognition of peptides is affected by many other variables in particular the HLA background of the individual. Although South Africa and China have genetically distinct clade C epidemics, we found that neither the magnitude nor the breadth of HIV-specific T-cell responses to Chinese and South African clade C peptide variants differed significantly. However, the magnitude and breadth of the responses to these two clade C peptide variants was significantly higher than that of Gag peptide reagents based on clades B, A and D sequences. Although D is closer to B than C, D was recognized with higher magnitude and breadth than other clade-mismatched peptides. Perhaps the greater variability in D does not reside in immunogenic regions and therefore preferential recognition of D compared to clades A and B. Overall, the data is further corroboration of results from previous studies which have shown that HIV-specific T-cells are cross-reactive among different HIV clades but with a preference for the infecting clade (Geldmacher *et al*., 2007; McKinnon *et al*., 2005). A similar reduction in epitope breadth for non-infecting clade peptide sets of approximately 50-70% was observed in a clade B-infected population when comparing recognition of clade B peptides to C and A peptides sets spanning the whole HIV genome (Yu *et al*., 2005). These data suggest that vaccines based on other clades may be cross-reactive and therefore warrant their testing in HIV-1 clade C-epidemic regions. However, whilst vaccines based on non-matching HIV clades may still induce cross-reactive responses, this reactivity may be less than that for clade-matched vaccines. Indeed, Gray *et al* (2011) demonstrated recently in the Phambili phase IIb trial of a clade B-based Ad5 immunogen in a South African population that 12% fewer vaccinees mounted a response to clade C peptides compared to clade B, with a 35% reduction in the overall magnitude of the responses. Whilst the vaccine was not protective in this population, it was also not effective in

a clade B-infected population (Buchbinder *et al*., 2008), so no conclusions can be drawn regarding the lower clade C-specific responses and vaccine efficacy. Clade-matching vaccines may represent a viable approach for regions where single clades circulate, such as Southern Africa, but this approach is highly limited for regions where multiple clades circulate, and increasing global HIV-1 diversity is a major challenge (Hemelaar *et al*., 2011).

The results from this study demonstrate that within a single individual, some HIV peptides were exclusively recognized in the clade C sequence variants (C_{Du422} and C_{CH}), whilst others were uniquely recognized in the clades B, A and D peptide variant. The recognition of clades B, A and D peptide variants and not the corresponding clade C peptide variants is of importance, as it demonstrates that using a single peptide reagent set leads to a considerable number of responses being missed when investigating T-cell immune responses (Frahm *et al*., 2007; Rolland *et al*., 2011). Of course, this approach is very likely still an underestimation of actual responses, since using peptide reagents matching the autologous virus demonstrates an increase in detectable T-cell responses of 29%, even in more conserved parts of the genome such as p24 (Altfeld *et al*., 2003). Importantly, an even greater increase of 37% in detectable epitopic regions was demonstrated when using a clade B Nef PTE (potential T-cell epitope) peptide set compared to clade B consensus peptides (Malhotra *et al*., 2007); indeed, preclinical vaccine studies show that these synthetic mosaic immunogens based on the PTE approach expand both the breadth and depth of T-cell responses (Barouch *et al*., 2010; Kong *et al*., 2009; Santra *et al*., 2010), as discussed in chapter one. Whether these increases in cross-reactive breadth are sufficient to be cross-protective remains to be elucidated in clinical trials.

It was surprising to note that only 14% of the reduction in cross-reactivity could be explained by the variability in peptide sequences. Although this was a low predictive value, it was significant and but suggested that there could be many other variables involved including HLA alleles that could be affecting cross-reactivity of these peptides. Therefore further characterization of reactive peptides in the study identified highly cross-reactive peptides with low intra- and inter-clade diversity, as shown by their lower entropy scores. Peptides that were recognized in two or more variant forms had significantly lower entropy scores when compared to peptides recognized once across the peptide sets. The pattern of recognition observed in the mutually recognized peptides may imply that HLA alleles restricting these peptides are driving mutations in the epitopes, as shown by loss of recognition of certain variants. Further characterization of the mutually recognized peptides identified other key

factors including TAP and MHC binding and proteasomal cleavage as playing a role in the recognition of specific variants and not others. This may illustrate the evolution of HIV due to T-cell pressure in HLA class I-restricted epitopes (Allen *et al*., 2004; Carlson *et al*., 2008; Dong *et al*., 2011; Kawashima *et al*., 2009; Kiepiela *et al*., 2004; Rolland *et al*., 2011), which is evidenced by HLA footprints observed in specific regions of the viral proteome containing HIV-specific HLA class I restricted T-cell epitopes (John *et al*., 2010; Moore *et al*., 2002). This phenomenon of HLA-driven viral evolution was illustrated recently with the first evidence of vaccine-driven T-cell footprints on viral sequences reported in breakthrough infections in the STEP trial (Rolland *et al*., 2011).

Overall, these data further corroborate previous findings which suggest that within the clades A, B, C and D sequences, some corresponding viral regions share a similar degree of conservation, possibly due to structural constraints that prevent sequence mutations in specific parts of the viral genome (Woo *et al*., 2010). This is further supported by the finding that most of the highly cross-recognized peptides were from the p24 region of the Gag protein which is known to be highly conserved and play a structural role in the HIV proteome. Yet, cross-clade recognition of peptides with considerable differences in their amino acid composition was also observed; most of the amino acid changes were semi-conserved, that is between amino acids with closely related side chains, and therefore did not have a significant impact on the processing of the epitopes for presentation by HLA alleles. This suggests that T-cell receptors of HIV-specific T-cells as well as HLA molecules can tolerate some degree of amino acid substitution in their epitopes without total loss of epitope recognition or binding as previously found in other studies (Addo *et al*., 2003; Geels *et al*., 2005; McKinney *et al*., 2004). Interestingly, even the same peptide was recognized to different degrees of cross-reactivity in different individuals, showing that different HLA molecules tolerate amino acid changes to different extents. Furthermore, use of different T-cell clonotypes through different TCR recruitment might have impacted recognition (Litcherfeld *et al*., 2007). Together, these effects of HLA and TCR have to be noted when designing vaccine immunogens that elicit cross-reactive responses in different populations. In addition, measuring cross-reactivity by performing viral inhibition assays could be an alternative way of performing these studies, than to reply on the use of immunological readouts as surrogates of predicting cross-reactivity.

In conclusion, the results in this chapter have shown that clade C HIV-infected individuals recognize peptides based on Chinese and South African sequences equally well, suggesting

that intra-clade variability from diverse geographic regions may not necessarily be an impediment to vaccine designed to elicit T-cell responses. However, while extensive cross-clade recognition was detected, the total magnitude was lower and the breadth of T-cell recognition narrower when compared with intra-clade C T-cell responses suggesting that vaccine-induced T-cell immunity of clade-mismatched vaccines would result in lower immunogenicity at the epitope level.

A range of approaches are currently being pursued to develop cross-reactive HIV vaccines, including those containing only conserved regions of the HIV proteome among clades (Letourneau *et al*., 2007; Rolland *et al*., 2007; Rosario *et al*., 2010), as well as mosaic approaches that seek to represent the majority of the diversity within clades (Fischer *et al*., 2007). Immunogens containing conserved regions would serve to focus the T-cell vaccine-induced response towards regions that are less likely to mutate due to structural constraints, and specifically exclude responses to variable regions that may be easily escapable, have little consequence on viral control, and may even act as decoys masking responses to conserved regions (Li *et al*., 2011).

The identification in this study of mutually reactive epitopes within conserved regions of the Gag protein support vaccine design strategies that incorporate conserved regions of the viral genome. Alternatively, T-cell mosaic antigens seek to increase cross-clade reactivity by maximizing the T-cell epitope coverage for most variants (Fischer *et al*., 2007). Ultimately, only testing these different vaccine approaches in clinical efficacy trials will inform us of what the best approach is for long-term protection from HIV acquisition or disease and the data from this part of the study support the use of conserved regions in developing cross-reactive vaccines. However, whether there are enough conserved regions for inclusion in vaccine or whether true cross-reactivity will be an issue is unknown, since it remains to be determined how many epitopes are needed to confer protection from infection or disease progression. The concept of HLA supertypes, where groups of HLA alleles can restrict the same epitopes, and the demonstration of HLA promiscuity, where many HLA alleles can restrict more than one epitope or tolerate epitope variants, might provide possibilities for good MHC coverage in different populations by conserved epitopes. Since similar analyses have not been extensively performed on HLA class II alleles, it is not known whether this could be the case with CD4 epitopes. From the observation that some long term non-progressors may target a single immunodominant epitope and this correlates with control of viral replication for extended

periods, one could speculate that a handful of conserved epitopes that result in a fitness cost to the virus and that were promiscuous in their binding to multiple common HLA could be 'enough' for inclusion in a vaccine.

CHAPTER 3

CHARACTERIZATION OF HIV-1-SPECIFIC T-CELL RESPONSES TO GROUP M CONSENSUS PEPTIDE REAGENTS

Table of Contents

3.1. INTRODUCTION

The high level of genetic diversity of HIV-1 poses a major challenge for global vaccine development (Garber *et al*., 2004). The use of centralized sequences in HIV vaccine immunogen design has been proposed in order to minimise the genetic distance to multiple clades, and potentially maximise cross-clade reactivity. Central sequences can be derived using various approaches, including (i) consensus (the most frequent base at a given position across corresponding viral sequences); (ii) centre-of-tree (COT; the point in a phylogenetic tree that minimizes the evolutionary distances to all sampled viruses); or (iii) most recent common ancestor (MRCA; the derived progenitor to a set of sequences; Ellenberger *et al*., 2002; Gaschen *et al*., 2002; Korber *et al*., 2001; Mullins *et al*., 2004; Nickle *et al*., 2003; Novitsky *et al*., 2002; Rolland *et al*., 2007). A detailed characterization of these centralized sequences, particularly consensus sequences based on group M viruses, has not been performed in large data sets of clade C infected individuals.

Several studies have investigated the feasibility of vaccines based on centralized sequences in murine or non-human primate models (Doria-Rose *et al*., 2005; Kothe *et al*., 2006; Rolland *et al*., 2007). In a murine model, broader HIV-specific T-cell responses were detected against group M consensus Env immunogens than against any single wild type subtype A, B and C Env immunogens tested (Weaver *et al*., 2006). In a non-human primate study, a candidate vaccine expressing a group M consensus Env immunogen that was optimized using phylogenetic approaches for a shorter variable region compared to wild type Env (Weaver *et al*., 2010) elicited T-cell responses that were significantly broader compared to single clade vaccine sequences based on clades A, C and G (Santra *et al*., 2008). To date, candidate vaccine immunogens based on group M consensus sequences have not been tested in human vaccine trials, so their potential for eliciting broad immune responses and for providing protection against multiple clades from group M is unknown.

One way to evaluate cross-clade reactivity is to assess HIV-specific T-cell responses to these centralized reagents in HIV-infected individuals using the IFN-γ ELISPOT assay. There have been several studies that have characterised the immunogenicity of group M consensus sequences in comparison to subtype-specific sequences. One such study of 43 subtype B-infected individuals from the US evaluated HIV-1 specific T-cell responses to seven different

Gag peptide reagents, including peptides based on clade B consensus, subtype B isolate HXB2, ancestral group M, ancestral subtype B, and consensus A, C and group M (Bansal *et al*., 2006). Out of the 42 reactive peptides identified, 29 (69%) peptides demonstrated cross-reactivity among all the seven Gag peptide reagents tested in at least one of the study individuals investigated. Cross-reactivity was highest for p24 Gag (71%), followed by p15 (67%) and lastly p17 (63%). This study also evaluated responses to these peptide reagents in a limited number of subtype C-infected participants from Zambia (n=13) of which reactive peptides were only confirmed in six individuals. Overall, when subtype B and subtype C infected populations were assessed, the subtype consensus peptides that matched the infecting virus clade detected responses of similar magnitude and breadth to consensus M and ancestral M reagents, and these were higher than responses to clade consensus peptides that were not matched to the infecting virus (Bansal *et al*., 2006). However, this cross reactivity was far lower when evaluating reagents based on the more variable HIV proteins such as Env. In a study of 17 subtype B-infected participants, Rutebemberwa *et al.* (2005), found limited cross-reactivity (only 6% compared to 69% from the Bansal *et al.* study) between consensus B, consensus group M and HIV-1_{MN} (clade B strain-specific) sequences using the Env protein. In this study, 16, 14 and 11 individuals responded to HIV-1_{MN}, consensus clade B and consensus group M peptide sets, respectively (Rutebemberwa *et al*., 2005). In addition, out of the 32 unique responses detected by the three peptide sets, 7/32 were common between consensus clade B and consensus group M, 10/32 were common between HIV-1_{MN} and consensus clade B, no peptides were unique to HIV-1MN and consensus group M peptides, and only two peptides were cross reactive across the three peptide sets. Seventeen peptides were exclusively recognized in the HIV-1_{MN} peptide set and one peptide each exclusively in the consensus group M and consensus clade B peptide set (Rutebemberwa *et al*., 2005).

A more recent study evaluated responses in subtype B infected individuals from Boston, Peru and Barbados (n=54), and clade C infected South Africans (n = 10) to consensus and ancestral clades B and C Gag and Nef peptide sets, as well as centre-of-tree for clade B, Consensus M and centre-of-tree for group M (Frahm *et al*., 2008). HIV-1 specific T cell responses were equally well detected in terms of both magnitude and breadth of responses by each clade-specific central sequence. Although not significantly different, the number of targeted peptides to these centralised peptide sets was lower for Peru and Barbados compared to the US (Frahm *et al*., 2008). It was interesting to note that Peru and Barbados had fewer sequences that contributed to

the phylogenetic design of these centralised reagents, suggesting that lack of adequate representation of population sequences in centralised sequences may limit detectable responses. In addition, in clade C infected individuals, the breadth of clade B peptides was reduced significantly compared to consensus clade C peptides ($p = 0.0039$), while the number of targeted peptides to group M and clade C sequences did not differ significantly ($p > 0.05$), Frahm *et al*., 2008). Even among the clade B based peptides, ancenstral B responses were broader compared to consensus B, possibly due to the fact that ancestral B is more central and thus closer to clade C sequences than the consensus clade B sequence. Also, detailed assessment of the data from this study demonstrated that the number of targeted peptides to group M were higher in clade C- than in clade B-infected individuals (Frahm *et al*., 2008). Overall, there was a significant increase in the breadth of responses when sequences were combined ($p < 0.0001$), with individuals having double the number of responses detected.

A fourth study, of 25 early HIV-1 clade B-infected individuals, similarly evaluated HIV-1 Gag-specific T-cell responses to CON (consensus)-, COT (centre-of-tree)- and ANC (ancestral)-clade B. As in previous studies, they found responses were similar in terms of magnitude, breadth, frequency and epitope specificites of reactive peptides (Malhotra *et al*., 2007). However, clade-mismatched consensus peptides based on clades A and C had slightly lower magnitude, specificity and breadth of responses compared to clade-matched consensus reagents (Malholtra *et al*., 2007). Detailed sequence charaterization of the detected epitopes identified 16 distinct epitopes; with almost half (seven) being invariant across the three clades tested, and 15/16 epitopes were cross-reactive with at least two clades (A, C or both). Amino acid substitutions in 9/16 epitopes were located outside HLA anchor residues (Malhotra *et al*., 2007).

These four studies illustrate the value of investigating different types of centralised or clade-specific peptide reagents to assess the effect of genetic diversity on HIV-specific T-cell responses in infected individuals. To summarise these findings, in some cases there was preferential recognition of clade-matched peptide reagents compared to group M or ancestral-based peptides in terms of both magnitude and breadth of the responses, and in other cases these were equally recognised or group M preferentially recognized compard to clade-mismatched peptides. The results are consistent with the fact that the genetic distance between group M consensus sequences and circulating viruses ranges from 5-15%, which is approximately half the inter-clade distance (Gaschen *et al*., 2002; Nickle *et al*., 2003). Also, within a single clade,

ancestral sequences detect broader responses than consensus sequences of that clade, highlighting the importance of evaluating the different approaches proposed for generating these centralized sequences. However, most of these studies were performed in predominantly clade B-infected populations, with a limited number of clade C-infected individuals. Furthermore, it is important to consider in all these studies that the HLA background of the population under study plays an important role in epitope recognition.

This chapter reports on a study, which forms part of a five-country investigation into HIV-1 diversity and immunity in Africa through the African AIDS Vaccine Programme, involving South Africa, Cameroon, Uganda, Kenya and Ethiopia. The overall aim of the study was to characterize HIV-specific T-cell responses to group M consensus Gag and Nef proteins in HIV-infected individuals living in diverse geographical regions in Africa, where different HIV-1 clades are circulating, and ethnically-diverse population live. Countries were selected based on the different subtypes responsible for their HIV epidemics: Cameroon has one of the greatest assemblage of HIV-1 subtypes, circulating recombinant forms and unique recombinant forms in the world; South Africa and Ethiopia have mainly a subtype C epidemic; and subtypes A and D circulate in Kenya and Uganda, with a high proportion of recombinants from clades A, C and D (Figure 3.1). The results from Uganda have recently been published (Serwanga *et al*., 2011), and this chapter reports on the results from South Africa. In this study, two HIV proteins, Gag and Nef, were chosen for investigating sequence diversity and cellular immune responses, as HIV-infected individuals commonly respond to these proteins, and the conserved and variable regions intrinsic to these proteins enable the effect of viral diversity to be evaluated. In this chapter, two interrelated aims were investigated, namely: a) characterization of HIV-specific T-cell immune responses to consensus group M Gag and Nef peptide reagents; and b), comparison of these responses to those generated against the near to consensus clade C_{Du422} peptide reagent in an HIV-1 clade C-infected cohort. These results were related to the sequences of the peptide reagents and the infecting viruses.

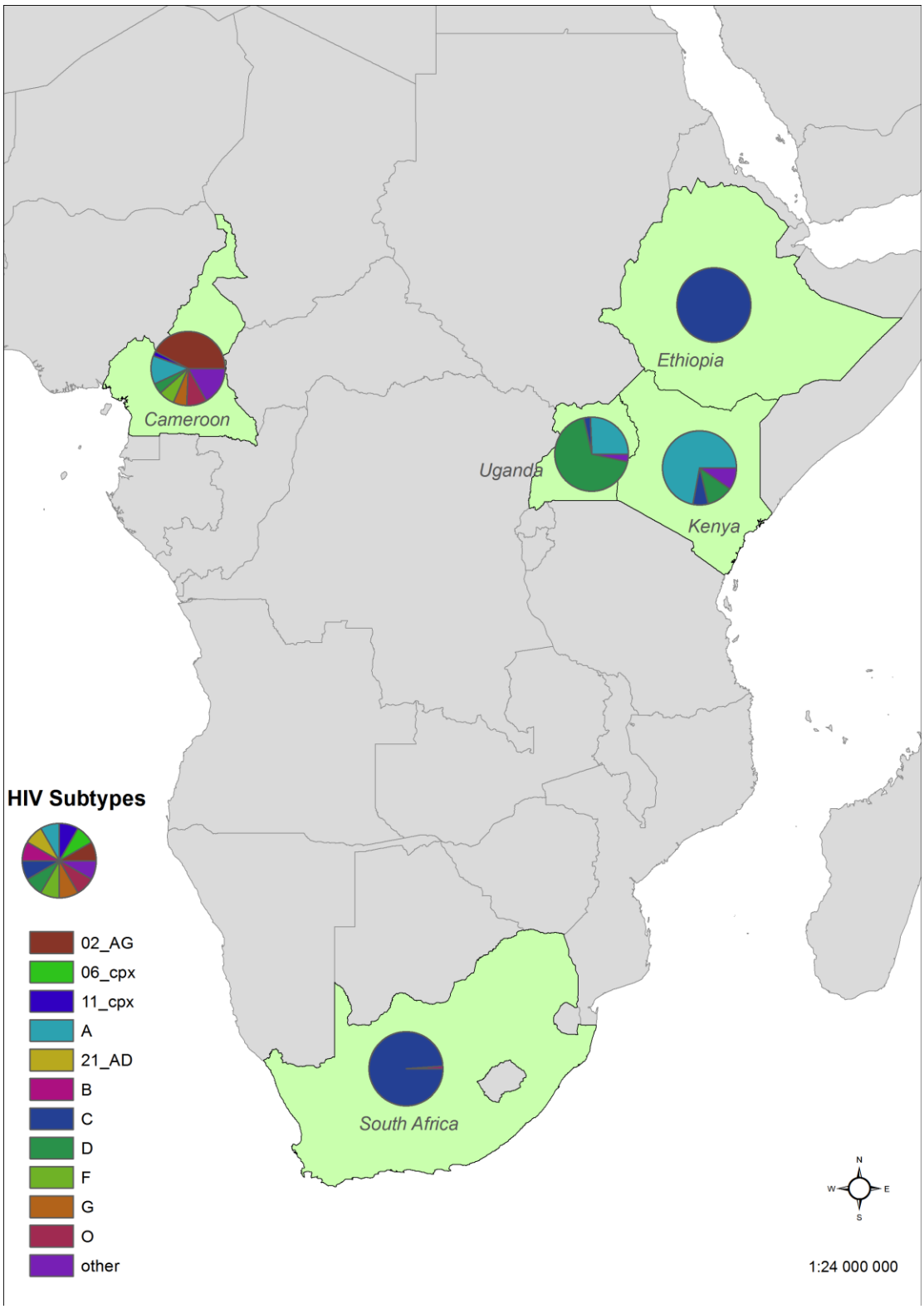


Figure 3.1. Map showing the HIV epidemics of five selected African countries. South Africa and Ethiopia demonstrate predominantly mono-clade epidemics, while multiple clades circulate in Uganda, Kenya and Cameroon, with the latter demonstrating extensive diversity of the virus. Table of names and geographical locations of countries were obtained from Natural Earth Data website (www.naturalearthdata.com). Sequence data for the five countries was obtained from the HIV geography site of the Los Alamos Sequence database (www.hiv.lanl.gov) as of 17 January 2011. The tables were joined and modified and map drawn using the Arc Geographical Information System (GIS) software version 9.3.1 (Arc GIS, ESRI ArcMap™, CA).

3.2. MATERIALS AND METHODS

3.2.1. Study participants

Eligibility criteria for the study were willingness and ability to provide informed consent and a CD4 count >200 cells/mm^3. For those who were on the ART start criteria was CD4 count <200cells/µl according to South Africa national guidelines.The ethical review board of the University of Cape Town approved the study and each study participant provided written informed consent. CD4 counts and viral loads were performed as for the study participants described in Chapter 2. CD4 count data was unavailable for five study participants and viral load data was not available for one individual. Peripheral Blood mononuclear Cells (PBMCs) were isolated using standard Ficoll-Hypaque density gradient centrifugation, and plasma isolated, as described in Chapter 2.

3.2.2. HIV-1 *gag* and *nef* sequencing

HIV RNA was extracted from plasma samples using the Magna-Pure Compact Nucleic Acid extractor (Roche). HIV *gag* cDNA was generated using the Invitrogen Thermoscript^{TM} RT-PCR system (Invitrogen). cDNA from the RT step was amplified in a first round polymerase chain reaction (PCR) using sequence-specific primers, Gag D forward HXB2 position 5"….3" 626-644 and Gag D reverse HXB2 position 5"….3" 2402-2382. Full length *gag* was amplified using Gag A forward 5"….3", HXB2 683-704, and Gag C reverse, HXB2 position 5"….3" 2356-2334 in a second round PCR. From the same extracted RNA, HIV *nef* cDNA was generated using Invitrogen Thermoscript^{TM} RT-PCR system (Invitrogen). The cDNA from the RT step was amplified in a first round using sequence specific primers namely Nef OR primer, HXB position 5"….3" 9608-9625 and SQ15FC, HXB position 5"….3" 8561-8578 and then in a second round using Nef F, HXB position, 5"….3" 8754-8776 and Nef R, HXB position 5"….3" 9443-9461 (Bredell *et al*., 2007). Primer sequences and cycling conditions are shown in Appendix B, B1 and B2 for Gag and Nef, respcetively. The amplified products were "bulk" sequenced (i.e. sequenced the major population) in both 5' and 3' directions on an automated DNA sequencer in six separate reactions for each study participant. The resulting sequences were assembled using ChromasPro or Sequencher, aligned using BioEdit and phylogenetic analyses performed using MEGA5.

3.2.3. Peptides

Group M consensus Gag and Nef 15-mer peptides overlapping by 11 amino acids from the NIH AIDS Reference Reagent Repository were used to test for immune reactivity in 66 HIV infected individuals. There were a total of 129 Gag and 53 Nef peptides. The peptides were arranged in protein specific and matrix pools, where a peptide appeared once in each protein-specific pool and once in each matrix pools, to allow for predictive single peptide mapping (Appendix D1). Peptides were arranged in 12 Gag pools, five Nef pools and 12 Matrix pools, as shown in the ELISpot plate layout sheet in Appendix D1. In some individuals, HIV Gag clade C_{Du422} peptides with were tested with consensus group M Gag and as well as group M Nef peptides using a different layout (Appendix D2). Peptide pools were tested in duplicate except in individuals where clade C_{Du422} peptides were included (Appendix D2, Table D2.4) and matrix pools in single wells, at a final concentration of 1μg/ml. Single peptides that were predicted to be reactive after deconvoluting the pool/matrix layout from the initial screen were tested at the same concentration in a confirmatory ELISpot assay. In 17 study participants, an additional set of Gag peptides based on subtype C_{Du422} from the NIH AIDS Reference Reagent Repository was used. These peptides were arranged into 12 pools and 10 matrix pools, and tested and confirmed as described above. The CEF peptide pool (obtained from the NIH AIDS Reference Reagent Repository) was used as a positive control for each sample, as well as for testing of a quality control sample on each plate. This peptide pool, which consisted of 32 8-11-mer CMV, EBV and Flu virus epitopes recognized by CD8+ T-cells (Currier *et al*., 2002), was used at a final concentration of 1μg/ml.

3.2.4. IFN-γ ELISpot assay

HIV-specific T-cell responses were assessed by an IFN-γ ELISpot assay, as previously described in Chapter 2 of this thesis.

3.2.5. Statistical analyses

Statistical analyses were performed using two-tailed distribution free or non-parametric methods using GraphPad Prism version 5.00 for Windows (GraphPad Software, San Diego

California USA). Total magnitude of responses to each individual reactive peptide in the Gag and Nef proteins were compared using the Mann-Whitney test for unpaired independent measures. The total number of peptides recognized by each study individual was compared between Gag and Nef proteins using the Wilcoxon signed rank test for matched pairs. The magnitude and number of peptides recognized by each subject were compared between consensus Gag and subtype C_{Du422} using the Wilcoxon signed rank test when matched, or Mann-Whitney test where data from chapter 2 where compared. The overall number of responses to the peptide sets used was tested for differences using the Kruskal-Wallis test followed by Dunn's post-test for multiple comparisons. The Mann-Whitney test was used to test for differences between ARV positive and negative subjects for each protein and within each protein. Correlation analyses were performed using the Spearman correlation method. In all analyses, $p<0.05$ was considered statistically significant.

3.3. RESULTS

3.3.1. Characteristics of study participants

In total, 66 participants were included in the study of which 44 ARV-naïve participants and 22 study individuals who were on ARV treatment were included (Table 3.1). The median viral load of untreated participants was 13000 copies/ml (range <50 to 490000), which was significantly higher than that of the individuals who were on ARV treatment (median <400 copies/ml, range <50-580000, p = 0.0019). Inclusion criteria for this study was a CD4 count ≥ 200, and there were no significant differences in the median CD4 counts between the two groups, with medians of 441 cells/µl (range 215-931) and 537 cells/µl (range 290-1306, p = 0.1553), respectively. Overall, there were 14 study individuals whose viral load was lower than detectable levels, 11 of whom were on treatment and three who were not.

All 66 study individuals were screened for HIV-specific T-cell responses using consensus group M Gag and Nef peptides using the IFN-g ELISPOT assay. Fifty-three participants had positive responses to at least one Gag or Nef pool, with the remainder not responding to any of the peptides tested. Out of the 53, single peptide confirmation was performed in 43 study individuals, 37 of whom were not on treatment. In a subset of 17 study individuals, responses to Gag clade C_{Du422} were assayed. This reagent was based on the subtype C_{Du422} sequence tested in Chapter 2 of this thesis.

3.3.2. Phylogenetic analysis of *gag* and *nef* sequences

To identify the infecting subtype, as well as to determine how these infecting viral sequences were related to the peptide reagents used in the ELISpot assays, full-length *gag* sequences were generated from 23 participants and *nef* sequence data from 19 participants. All study participants were infected with pure HIV-1 subtype C viruses in both the *gag* and *nef* regions (Figure 3.2A and B). There were four clusters observed in the *gag* region of two sequences each (boostrap values > 85%), namely NY 133 and NY 155; NY 032 and NY 021; NY 264 and NY 163; and NY 138 and NY 162. These clusters suggest a relatively recent common ancestor for each pair, which is possible since these individuals were recruited from the same neighbourhood. These significant clusters were not observed in the *nef* region (Figure 3.2B).

Table 3.1. Clinical characteristics of study participants

PID	CD4 count (cells/µl)	Viral load (copies/ml)	ARV use	PID	CD4 count (cells/µl)	Viral load (copies/ml)	ARV use
NY172	316	100000	no	**NY040**	322	7600	no
ZA006	618	10728	no	**NY123**	343	8700	no
NY166	217	11000	no	**ZA034**	469	ldl	no
NY155	441	110000	no	**NY140**	782	ldl	no
NY167	405	120000	no	**ZA011**	923	ldl	no
NY133	448	1500	no	**ZA029**	448	nd	no
NY234	680	15000	no	**ZA023**	547	nd	no
NY221	317	16000	no	**ZA028**	553	nd	no
NY210	333	160000	no	**ZA031**	635	nd	no
NY219	551	170000	no	**ZA030**	877	nd	no
NY145	nd	21000	no	**ZA022**	931	nd	no
NY032	534	21000	no	**NY248**	473	1000	yes
NY164	665	240000	no	**NY273**	768	20000	yes
ZA021	607	2544	no	**NY270**	650	230000	yes
NY233	nd	2600	no	**NY030**	705	2400	yes
NY138	401	27000	no	**NY062**	407	26000	yes
ZA035	394	2772	no	**NY028**	547	320	yes
NY163	388	330	no	**NY279**	316	530	yes
ZA026	517	3539	no	**NY263**	nd	580000	yes
NY130	322	39811	no	**NY240**	340	5900	yes
ZA025	614	40438	no	**NY066**	563	84000	yes
ZA033	379	460	no	**NY293**	nd	ldl	yes
NY294	nd	490000	no	**NY002**	290	ldl	yes
NY122	215	500	no	**NY074**	390	ldl	yes
NY157	225	50000	no	**NY292**	423	ldl	yes
NY055	335	510	no	**NY029**	439	ldl	yes
NY264	251	56000	no	**NY045**	530	ldl	yes
ZA002	352	5845	no	**NY267**	545	ldl	yes
ZA024	378	59696	no	**NY139**	554	ldl	yes
NY057	384	6400	no	**NY039**	582	ldl	yes
ZA001	492	64617	no	**NY076**	640	ldl	yes
NY165	258	69000	no	**NY044**	1306	ldl	yes
NY162	461	7100	no	**NY277**	395	nd	yes

nd- not done or data not available
ldl- lower than detectable level (< 50 copies/ml)

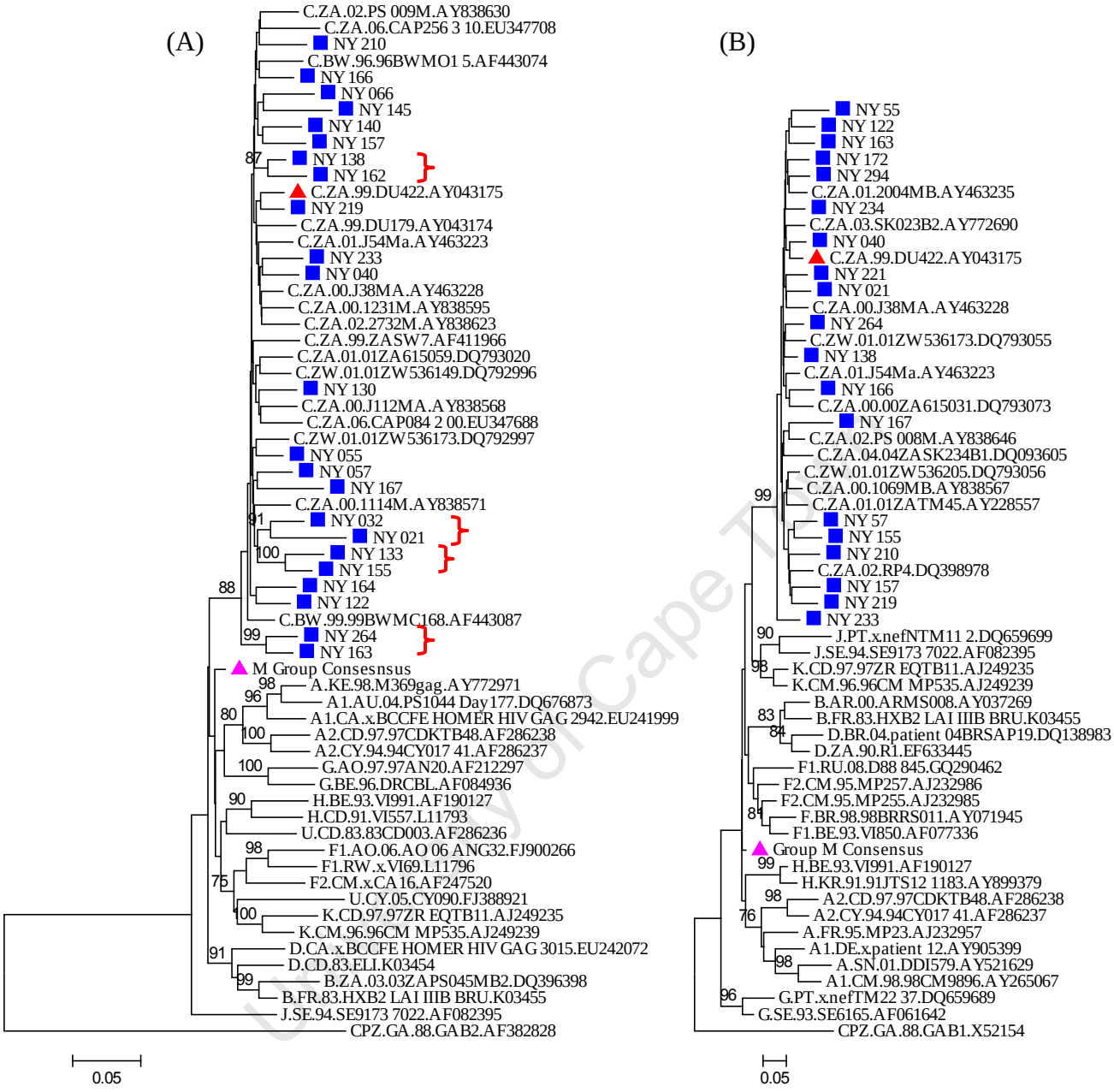


Figure 3.2. Phylogenetic relatedness of infecting viral sequences. (A) *Gag* gene sequences (~ 1500 bp). (B) *Nef* gene sequences (~700 bp) (next page). Phylogenetic analysis was performed on full-length *gag* and *nef* sequences and is based on the Neighbour-Joining method. The reliability of internal branches is based on 1000 bootstrap replicates. The scale bar indicates 10% genetic distance. Study participants infecting virus sequences are shown with blue squares. The Du422 sequence is shown with a red triangle while the consensus group M sequence with purple triangle. Sequences indicated by red right braces significantly clustered together with bootstrap values above 75%.

3.3.3. Amino acid distances between infecting viruses and ELISpot peptide reagents

Amino acid distances were generated to determine the relationship between sequences of infecting viruses and ELISpot peptide reagents. In the Nef region, amino acid distances between infecting viral sequences and consensus group M reagent were slightly higher than compared to subtype C_{Du422} reagent at a median of 15.5% (range 12.1-17%) and 14.4% (range 9.5-22.7%), respectively. However, this difference was not statistically significant ($p = 0.7239$, Figure 3.3A). However, amino acid distances in the Gag region were significantly higher between infecting viral sequences and consensus group M reagent when compared to the subtype C_{Du422} peptide reagent 11% (range 8.8-15%) and 6.7% (range 4.1-10.4%), respectively ($p < 0.0001$).

We then estimated the number of 9mers that were matched between the infecting virus and either Group M reagents or subtype C_{Du422} peptide reagents. We found, as expected, that epitope coverage was higher for Gag compared to Nef (Figure 3.3), and coverage was also higher for the subtype C_{Du422} reagent when compared to the consensus M peptide reagents. In the Gag protein, when considering exact matches, epitope coverage was 53.88% for the subtype C_{Du422} peptide reagent compared to 41.94% for the consensus group M peptide reagent, a 1.3-fold increase (Figure 3.3B). In the Nef protein, when considering exact matches, epitope coverage by C_{Du422} peptide reagent was 24.11%, and by the consensus M peptide reagent it was 18.75% (Figure 3.3B). Overall, the consensus M peptide reagents had lower predicted epitope coverage even when epitopes with 1 or 2 amino acid mismatches were taken into consideration, suggesting that a group M consensus immunogen would give poorer coverage compared to a subtype C sequence immunogen.

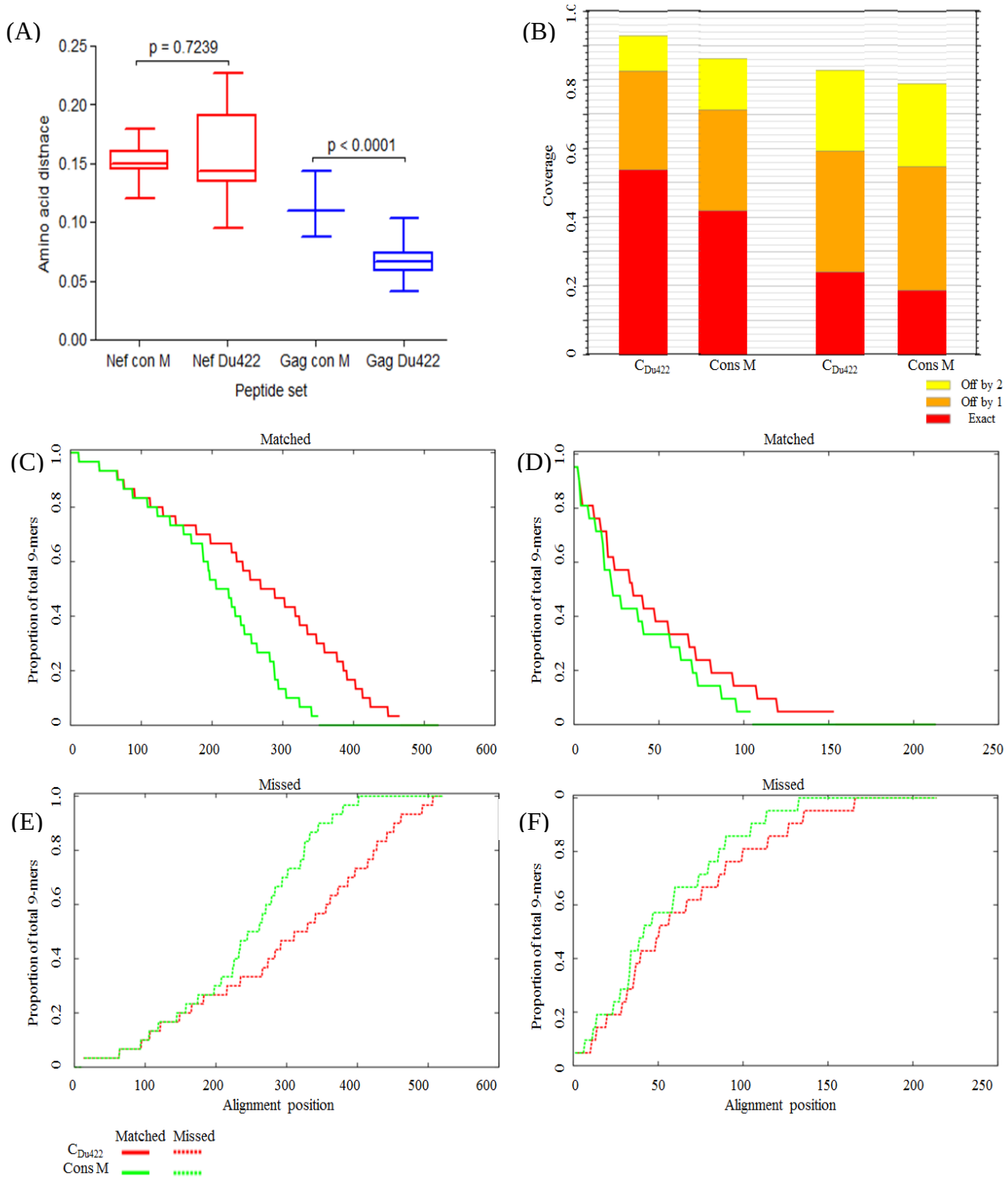


Figure 3.3. Characterization of infecting viral sequence and ELISpot peptide reagent sequences. (A) The amino acid distance between infecting viral sequences and ELISpot peptide reagent sequence. **(B)** Predicted epitope coverage of infecting viral sequences by ELISpot peptide reagent sequences. Epitope coverage was determined using the Epicover (www.hiv.lanl.gov) for Gag (N = 63, 40/70 sequences from Chapter 2) and Nef (N = 19). **(C-F)** The predicted positional epitope coverage of infecting viral sequences by ELISpot peptide reagent sequences. Matched epitope coverage ranked by coverage for Gag **(C)** and Nef **(D)**. Missed epitope coverage ranked by coverage for Gag **(E)** and Nef **(F)**. Reagent peptides were derived from C_{Du422}-South African subtype C strain Du422; and the consensus group M sequence (con M).

Next, the proportion of matched or mismatched epitopes was investigated using Posicover (www.hiv.lanl.gov) (Figure 3.3C to F). Consistent with the above data, the proportion of matched epitopes was higher for Gag compared to Nef (Figure 3.3C and D, respectively) and conversely, the proportion of mismatched epitopes was lower for Gag compared to Nef (Figure 3.3E and F, respectively). Furthermore, subtype C_{Du422} peptide reagent had a higher proportion of matched epitopes when compared to group M consensus peptide reagent and this was consistent for both Gag and Nef proteins (Figure 3.3C and D).

We then determined the variability of each pepide using the Shannon entropy score to determine which regions of the protein were more likely to be cross-reactive. Higher variability was observed for p17 (amino acids 1-145) and p15 (amino acids 379-520) Gag regions, and lower for the p24 (amino acids 146-378) region (Figure 3.4A). For the Nef protein, lower entropy scores (less variability) were observed in the central region (amino acids 80-164, Figure 3.4B).

3.3.4. Magnitude and breadth of Consensus M Gag- and Nef-specific T-cells

To assess the reactivity of HIV-1 group M consensus peptides in the 44 ARV untreated study participants, the frequency, magnitude and breadth of HIV-specific T-cells recognizing group M consensus Gag and Nef peptides were investigated. Study participants on treatment were excluded from this analysis as previous studies showed that HIV-specific T-cell responses are reduced in these individuals (Addo *et al*., 2003; Stranford *et al*., 2001), and also to allow for comparison with data from Chapter 2 of this thesis, where all study participants investigated were not on treatment.

Based on reactivity to the peptide pools, 81.8% (36/44) of study participants had a response to at least one peptide from group M consensus Gag or Nef, with 18.2% (8/44) individuals not recognizing any of the Group M peptides tested. Out of the 36 study individuals responding to consensus M Gag or Nef, 33 (91.7%) responded to at least one Gag peptide, while 29/36 (80.6%) responded to at least one Nef peptide. Confirmation of reactive peptides was performed in 29 responders. For these participants, 26/29 (90%) responded to Gag peptides, while 17/29 (59%) responded to Nef peptides. The magnitude and breadth of HIV-specific T-cell responses recognizing individual Gag and or Nef peptides in study participants are shown in Figure 3.5.

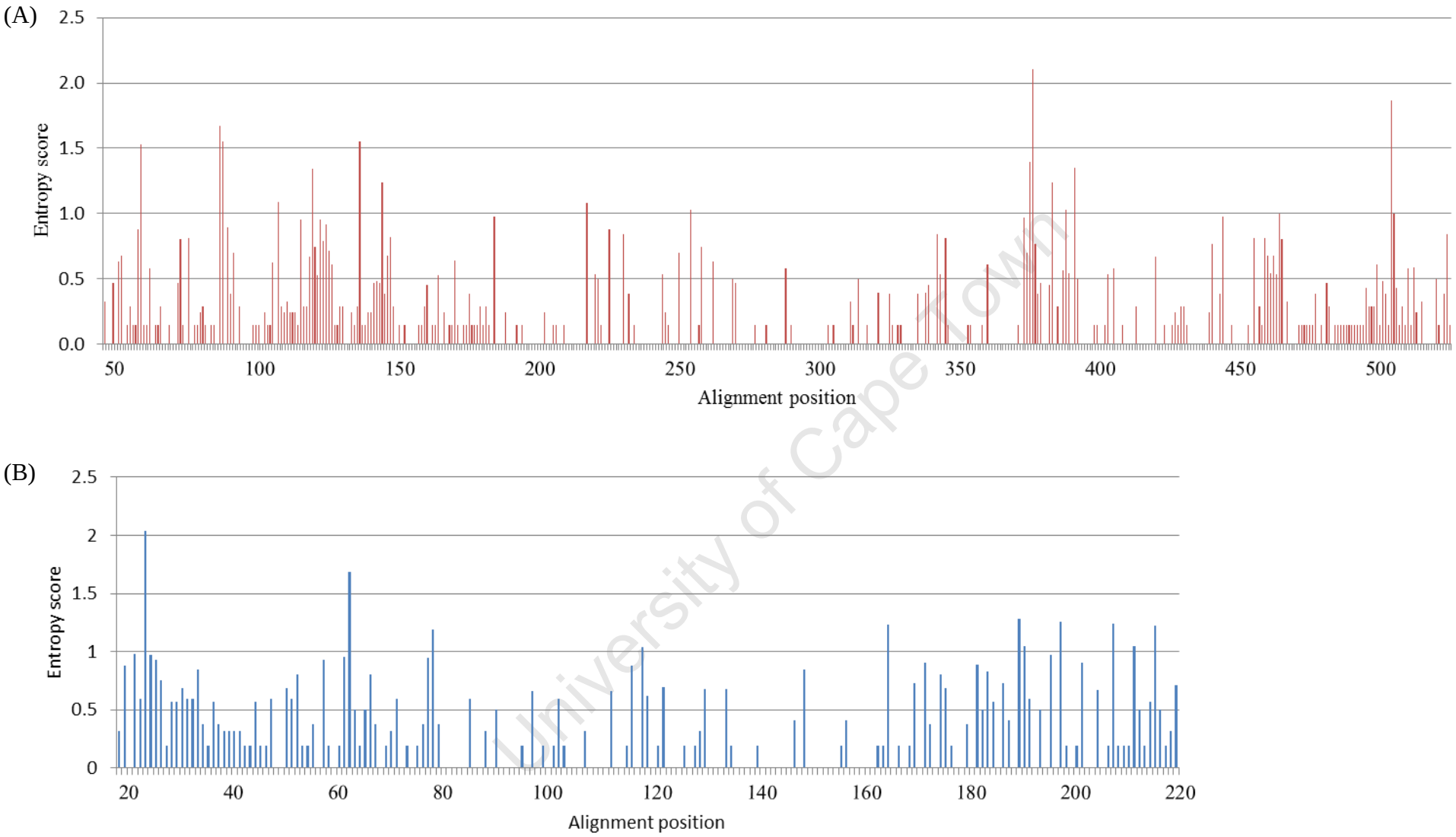


Figure 3.4. Entropy scores showing variation across protein sequences of study individuals' infecting viral sequences. (A) Gag protein, N = 23. **(B)** Nef protein, N= 19. The protein sequences of the study individuals' sequences were aligned and entropy score at each position in the alignment was determined.

The median magnitude of responses was significantly different between the two proteins, with Gag having a median of 1180 SFU/10^6 PBMC (range 0-7810) and Nef having a median of 195 SFU/10^6 PBMC (range 0-3730), respectively ($p = 0.0013$; Figure 3.5A). The median breadth of responses (the median number of epitopes targeted after taking into consideration consecutive peptides) was significantly higher for Gag, with a median of 2 (range 0-10) compared to the Nef protein, with a median of 1 (range 0-6) response ($p = 0.0017$; Figure 3.5B).

The two artificial proteins tested for reactivity in this study, HIV-1 group M consensus Gag and Nef, are characterized by a relatively conserved p24 and variable p17 and p15 regions for Gag, and a conserved central region in the Nef protein, with relatively variable outer regions. We assessed the relationship between immunodominance and location on the protein. Overlapping reactive peptides in the same individual were only counted as one response. A total of 50 different peptides were recognised in consensus M Gag (Figure 3.6A), out of 129 peptides making up Gag. There were nine peptides that were recognized in ≥3 study individuals for the Gag protein (18%), 14 in two individuals, and the majority, 27/50 peptides (54%) recognised in only one individual. The recognition of peptides was spread throughout the Gag protein, although the majority of the highly reactive peptides (7/9; 77.8%) were from the conserved p24 region of the protein, and 1/9 (11.1%) peptide each from the p17 and p15 regions of Gag. For remaining peptides that were recognised in only two individuals, 78.6% (11/14) were from p24, and for those recognised in only one individual, 70.4% (19/27) were from p24. The recognition of Nef was focused on 16 peptides (Figure 3.6B), out of 53 peptides spanning the protein, representing 30%. There were six Nef peptides that were recognized in ≥3 study participants (37.5%). Four out of these six peptides (66.7%) were from the conserved central region of the protein, while 2/6 (33.3%) peptides were from the amino terminus of the protein. There were four peptides that were recognized in only two individuals, 75% located in the central region of Nef, and six peptides recognized in only one study individual, 50% of which were located in the central conserved region of the protein. It was interesting to note that no peptide from the carboxyl terminus of the Nef protein was recognized in the study.

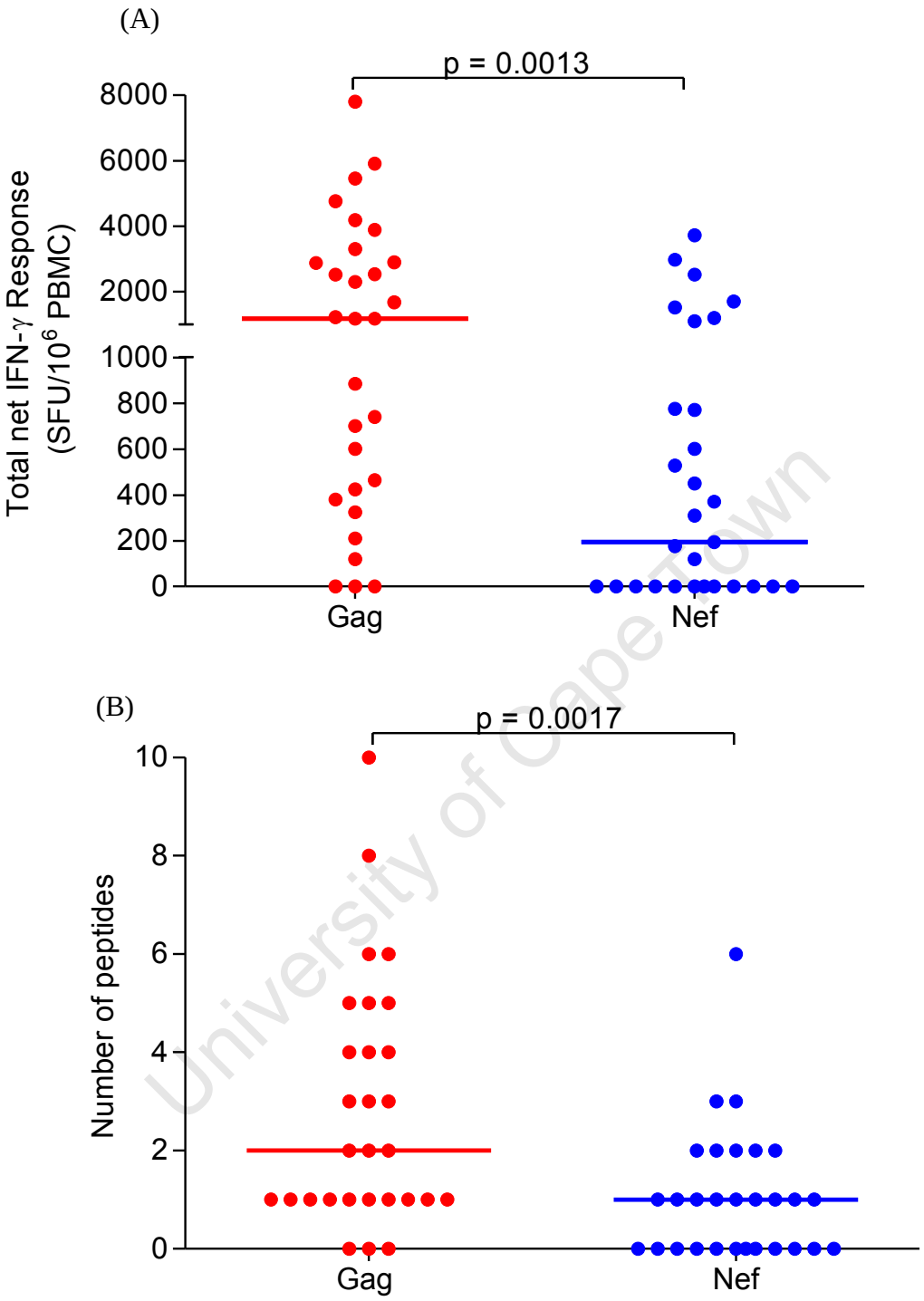
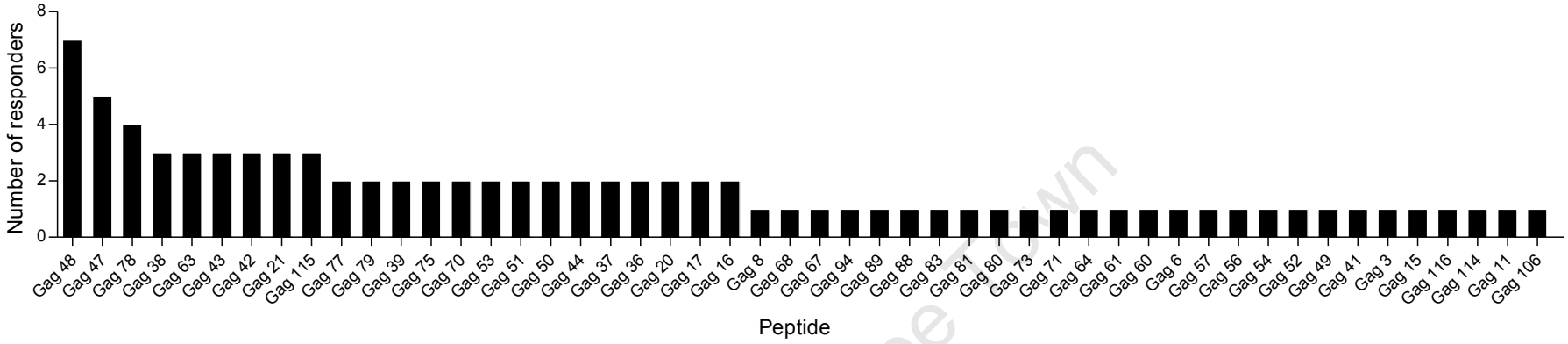


Figure 3.5. Total HIV-1-specific IFN-γ producing T-cells against group M consensus Gag and Nef proteins. **(A)** Total magnitude of responses to Gag and Nef individual peptides recognized in 29 study individuals. Each dot represents an individual. **(B)** Total number of peptides recognized in Gag or Nef proteins for 29 individuals. The number of responses is the minimum responses after taking into consideration consecutive peptides, as described in the Methods. Differences were compared using the Mann-Whitney test.

(A)



(B)

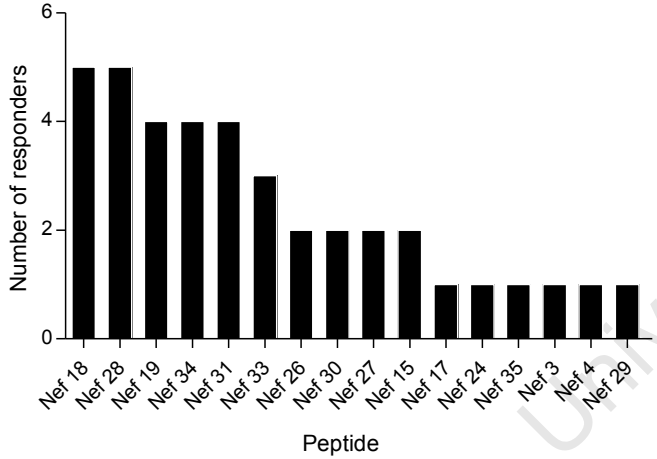


Figure 3.6. The immunodominance of HIV-1 specific T-cell responses to consensus group M peptides. (A) The number of individuals responding to each Gag, and **(B)** the number of individuals responding to each Nef peptide. The number of participants responding to each reactive Gag or Nef peptide for 29 study participants is indicated. In each individual two overlapping peptides were counted as one response and three as two responses, as described previously in the Methods.

3.3.5. Comparison of reactivity of Consensus M and subtype C peptides in a clade C population

To compare the reactivity of consensus group M Gag to clade C Gag, 17 responding study individuals were tested for T cell responses both to consensus M Gag and clade C_{Du422} Gag peptides. There was preferential recognition of the clade C Gag peptide reagents, with a median magnitude of 2005 SFU/10^6 PBMC (range, 0-10933) compared to consensus group M peptides, with a median of 1283 SFU/10^6 PBMC (range, 0-5911; Figure 3.7A). This difference was significant ($p = 0.0317$). In terms of the number of responses, clade C peptides were preferentially recognized in terms of breadth, with a median number of response of 3 (range, 0-11), however there was no significant difference in the median number of consensus group M peptides recognised, namely 2.5 (range, 0-10, $p = 0.1121$; Figure 3.7B) in these 17 study individuals.

To increase our power we performed additional comparisons using data generated from Chapter 2, in an unmatched analysis. Here we included Gag-specific T-cell responses to consensus M (n=29) and to clade C_{Du422} peptides (n=17) from this Chapter, and Gag-specific T-cell responses to clade C_{Du422} peptides from 39 individuals from Chapter 2. CD4 counts and viral loads were compared between individuals assessed for Group M reactivity (29) and those with clade C reactivity data (56), and no significant differences were found between these two groups of participants. Median CD4 counts were 547.0 cells/µl and 448.0 cells/µl, and viral loads 16000 copies/ml and 5845 copies/ml for group of the two groups of study participants, respectively ($p = 0.0798$ and $p = 0.3840$, Figure D3.1, Appendix D3). Consistent with genetic distance data and epitope coverage analysis, as well as the above analysis on the matchted group, the magnitude of responses was significantly higher (1.8-fold) for subtype C_{Du422} peptides compared to consensus group M-based peptides, with a median of 2078 SFU/10^6 PBMC (range 0-24550) and 1180 SFU/10^6 PBMC (range 0-7810) for subtype C_{Du422} Gag and consensus group M peptides, respectively in this larger cohort ($p = 0.0244$, Figure 3.8A). Using these increased numbers of participants we now found a significant difference in the median breadth to clade C and consensus M peptides. The breadth of responses was double for the clade C (C_{Du422}) peptide reagent, with a median of 4 responses (range 0-11), compared to consensus group M peptides, which was 2 (range 0-10, $p = 0.0048$; Figure 3.8B).

The trend in the magnitude or breadth of HIV-specific T-cells between clade C compared to group M consensus peptides was investigated and denoted by trend arrows joining matching cases (Figure 3.8A and B). In 75% (12/16) of the cases that had detectable Gag T-cell responses, the detected response was higher for the clade C peptide reagent compared to consensus group M reagent ($p = 0.0317$, Figure 3.8A). The remaining 25% had a lower magnitude of responses to the clade C peptides compared to consensus group M peptides. The breadth of responses followed a similar pattern, in which more reactive peptides were identified for the clade C peptide reagent compared to consensus group M petide reagent in 62.5% (10/16) of the cases (Figure 3.8B). The remaining 12.5% and 25% had equal and fewer reactive peptides detected by clade C reagent compared to group M reagent, respectively (Figure 3.8B). These data demonstrate that within a single individual, clade-matched HIV-Gag peptide reagents detect HIV-specific T-cells of higher magnitude and breadth than Gag reagents based on group M consensus sequences.

Altogether, these data suggest that although consensus group M based peptide reagents are recognized in HIV-1 clade C-infected individuals, they recognise fewer responses with a lower magnitude compared to clade-matched peptide reagent, in this case subtype C_{Du422}. This may have implications for the use of central sequences as vaccine immunogens.

In Chapter 2, it was demonstrated that the reagent matching the infecting clade is recognized with higher magnitude and breadth of responses when compared to clade-mismatched peptides reagents, namely consensus A, B and D. In this Chapter, peptides matching the infecting clade were recognized with higher frequency (36/39, 92% versus 33/44, 75%), magnitude and breadth when compared to consensus group M based peptide reagents. Therefore, data on inter-clade responses from Chapter 2 were compared with data from matched clade and consensus group M reagents. Only Gag reactivity was considered, since no Nef reactivity data was available for subtypes A, B and D.

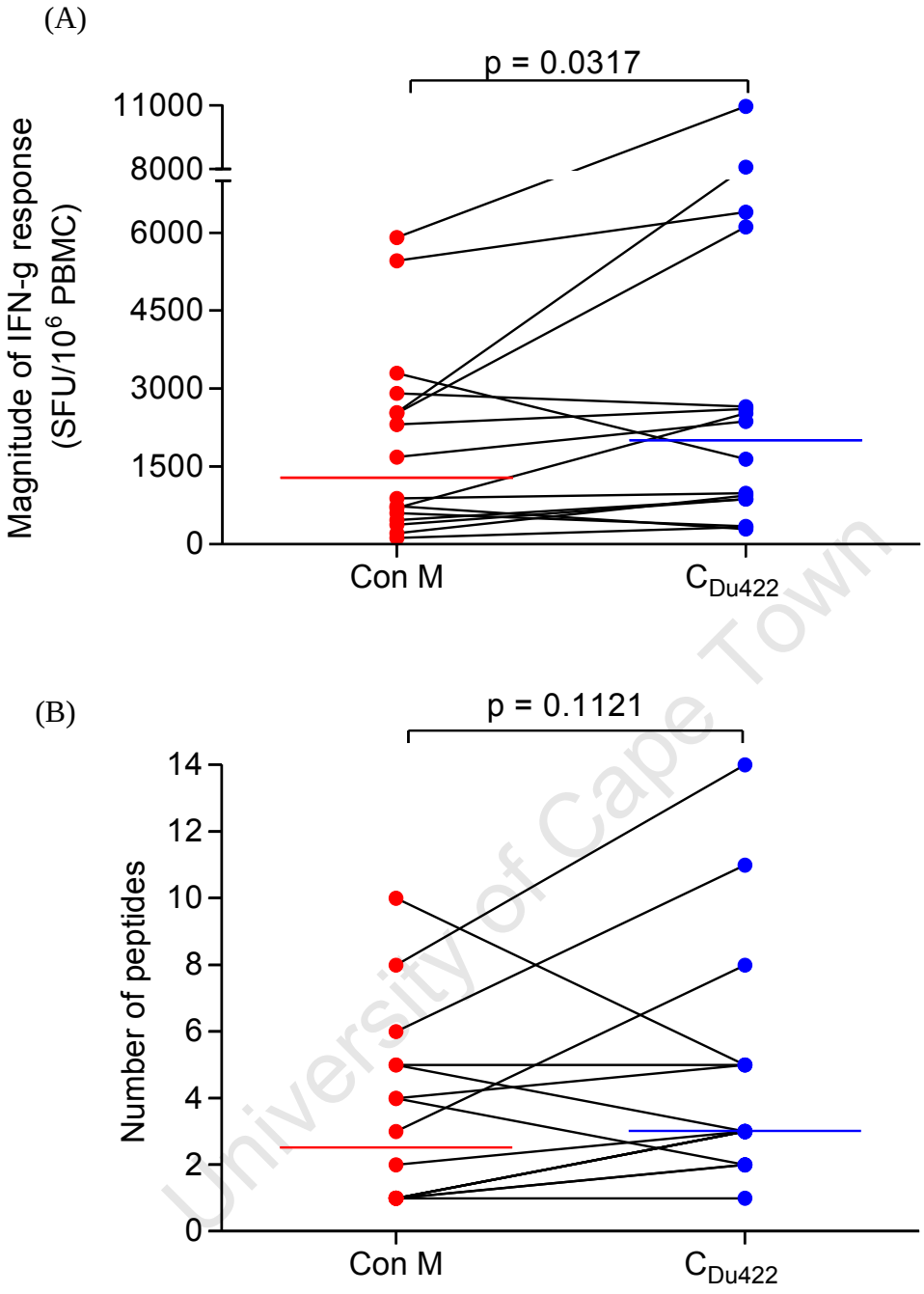


Figure 3.7. Matched comparison of magnitude and breadth of HIV Gag T-cell responses. (A) The total magnitude of HIV-specific T-cells recognizing consensus group M and clade C peptide reagents in clade C infected study participants. **(B)** The minimum number of consensus group M and clade C HIV Gag peptides recognized by HIV-specific T-cells in clade C infected study individuals. N=17. Each dot represents an individual. Differences were tested using the Wilcoxon signed rank test for matched data.

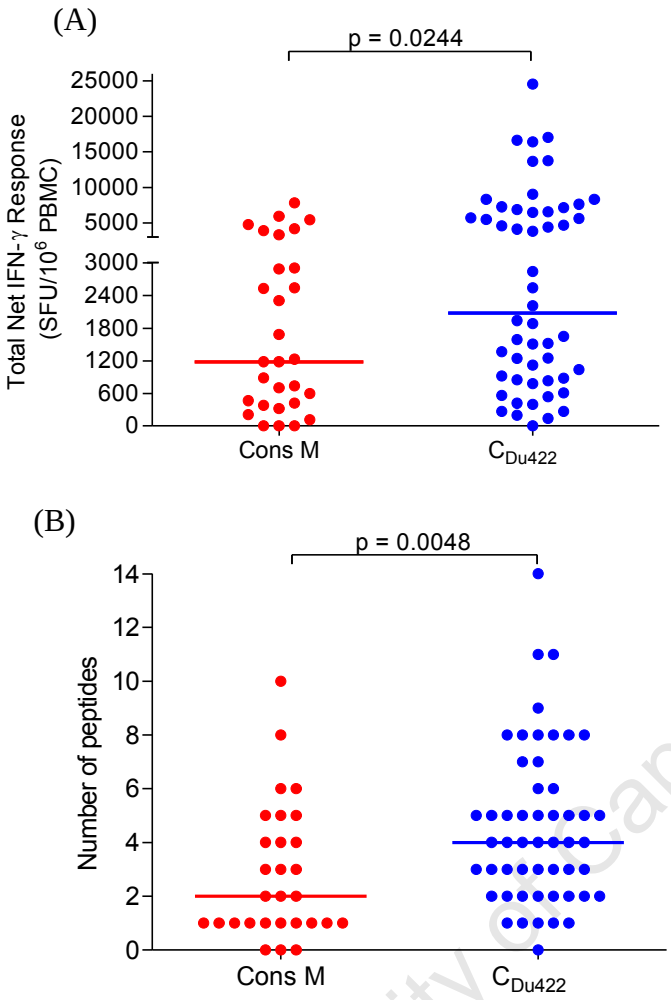


Figure 3.8. Comparison of total HIV-1 specific T-cells responding to group M consensus and subtype C Gag peptides. (A) Total magnitude of Gag responses. **(B)** Total number of peptides recognized. For C_{Du422}, individuals from Chapter 2 were included for analysis. Each dot represents an individual. Gag response data were available for 37 individuals for the Con M and for 56 (17 who had matched data plus 39 from Chapter 2) study individuals for the C_{Du422} responses. The Mann-Whitney test for unmatched data was used to test for differences.

Interestingly, the magnitude of responses were comparable between the consensus group M peptide reagents and the clade-mismatched reagents (Figure 3.9A), with a median of 1180 SFU/10^6 PBMC (range 0-7810) for consensus M, and inter-clade based consensus reagents based on subtypes A, B and D having medians of 750 SFU/10^6 PBMC (range 0-23817), 810 SFU/10^6 PBMC (range 0-11360) and 1390 SFU/10^6 PBMC (range 0-19474), respectively. Furthermore, the median number of responses did not differ significantly among the mismatched

peptide sets (Figure 3.9B, $p > 0.05$). In terms of frequency of responses, group M peptide reagent had detectable responses in 75% of individulas (33/44), and was comparable to that of mismatched clades of 76.9% (35/40), 71.7% (28/40) and 84.6% (33/40) for clades A, B and D, respectively. However, the frequency of recognition of Gag peptides was higher by 15.6% when clade-matched reagents were used compared to consensus group M Gag reagents, at 92.3% compared to 76.7% for clade C and consensus group M, respectively. These data suggest that, at least for Gag, consensus group M based peptide reagents can be used interchangeably with clade-mismatched reagents in the assessment of vaccine induced HIV-specific T-cell responses, and may imply that vaccine candidates based on consensus M spanning relatively conserved viral regions such as HIV-1 Gag, have no advantage over clade mismatched vaccines in a monoclade epidemic.

3.3.6. Comparison of the immune responses between untreated and treated individuals

Previous studies that have shown that HIV-specific T-cell responses are reduced in individuals on antiretroviral treatment, compared to those untreated. Since a portion of our study participants (22/66) were treated, the HIV-specific T-cell responses to Gag and Nef in study participants who were on treatment (n=16), compared to those who were not (n=29), were examined. The total magnitude of responses to Gag was significantly reduced in treated individuals compared to untreated individuals ($p = 0.0008$; Figure 3.10A). The responses ranged from 0-12585 and 5330 SFU/10^6 PBMC; with medians of 635 and 0 SFU/10^6 PBMC for untreated and treated individuals, respectively. Total HIV-specific T-cells against the Nef protein showed a similar trend, ranging from 0-4660 and 936 SFU/10^6 PBMC, with medians of 195 and 136 SFU/10^6 PBMC for untreated and treated study participants, respectively (Figure 3.10A). In contrast to the Gag responses, this difference was not statistically significant ($p = 0.3377$, Figure 3.10A). The median number of responses to the Gag protein was significantly higher for untreated study participants 2 (range 0-10) compared to treated study participants (1, range 0-6; $p = 0.0081$; Figure 3.10B). On the other hand, there was no significant difference in the median number of Nef responses between the treated and untreated study individuals ($p = 0.2834$, Figure 3.10). These data demonstrate that antigen load increases the magnitude and number of detectable HIV-specific T-cell responses. This effect was higher in Gag compared to Nef.

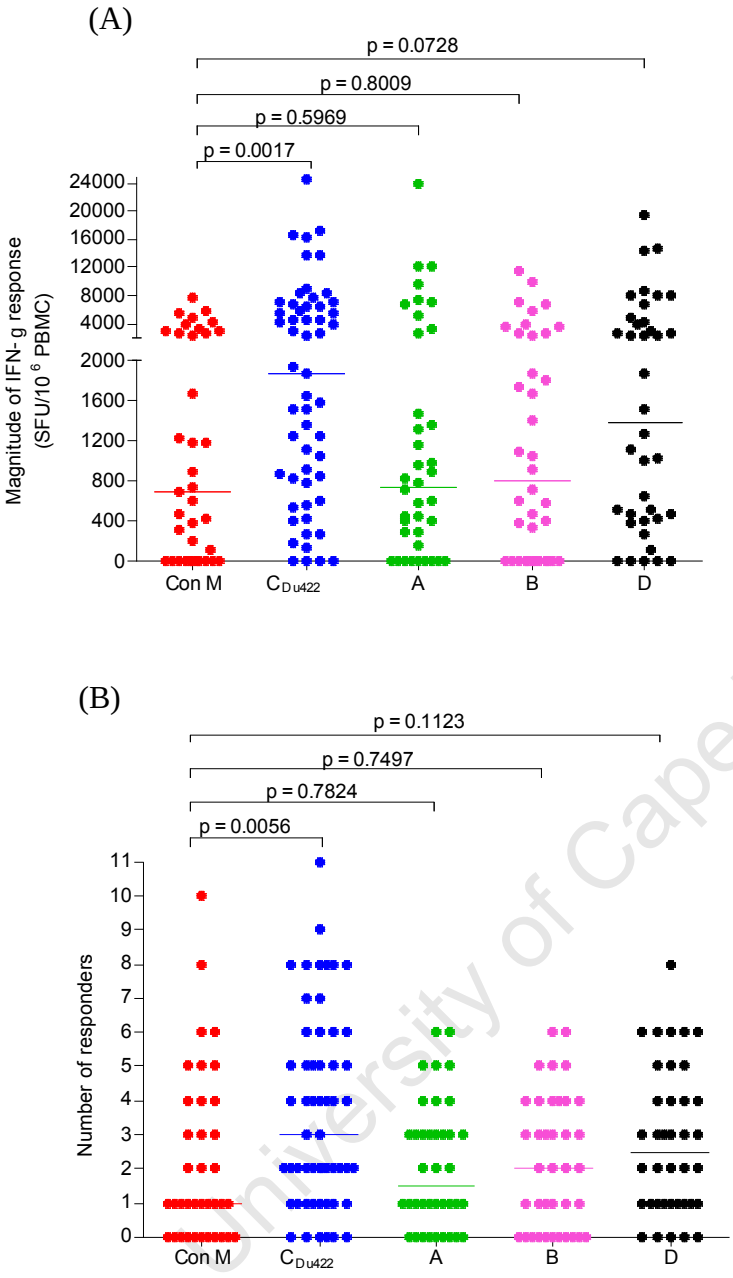
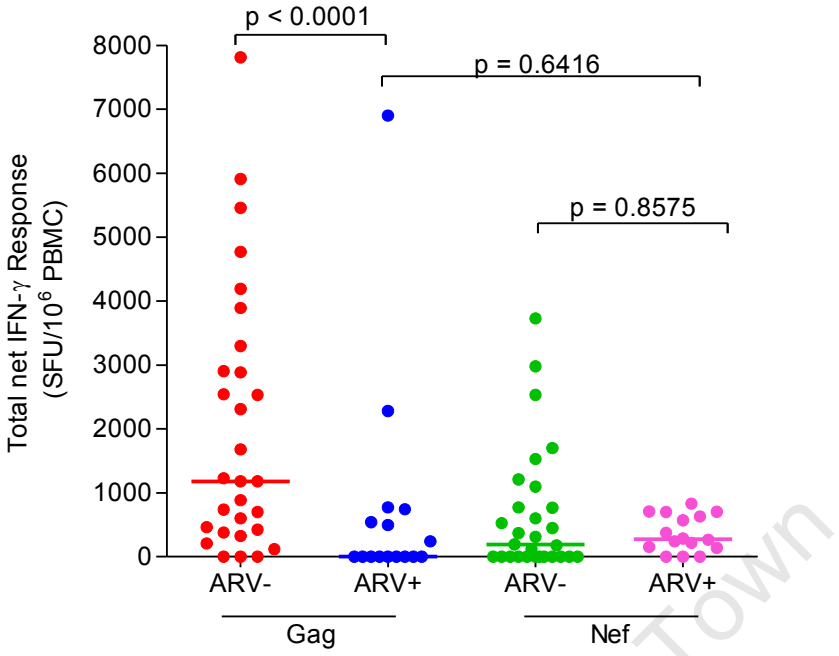


Figure 3. 9. Comparison of recognition of consensus group M peptides with subtype specific peptide reagents based on clades C_{Du422}, A, B and D. (A) Magnitude of responses. **(B)** Breadth of responses. Data from Chapter 2 were included in the analysis to compare consensus group M reagent to inter-clade consensus reagents namely A, B and D, as well as intra-clade based peptide reagent namely C_{Du422}. For C_{Du422} peptide set, total of 56 study individuals were included (17 from Chapter 3 and 39 from Chapter 2). For Con M (N=37), and for A, B and D (N=39) for each clade. Con M is the consensus group M peptide reagent. Mann-Whitney test for unmatched data was performed to test for differences.

(A)



(B)

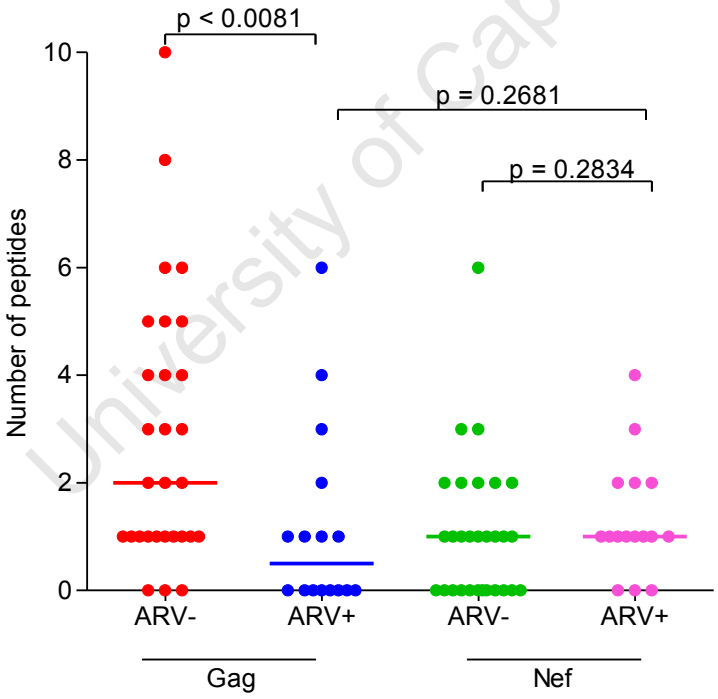


Figure 3. 10. Comparison of HIV-specific T-cell responses between participants on treatment (ARV+) and those not on treatments (ARV–). (A) The magnitude of HIV-specific T-cell responses to Gag and Nef proteins. **(B)** Comparison of the breadth of Gag- and Nef-specific T-cell responses between treated and untreated study participants.

3.3.7. Immunodominant peptides identified by consensus group M peptides

Peptides that were recognized in at least four study individuals were classified as immunodominant. In total, eight immunodominant peptides were identified in the study (Table 3.2). There were three such reactive peptides recognized within the Gag protein, and five within the Nef protein (Table 3.2). All (3/3) immunodominant reactive Gag peptides were located within the p24 region of the protein and the predescribed epitopes within these peptides were identical in sequence to their corresponding clade C variants. For Nef protein, all five immunodominant peptides were located in the conserved central region of the protein. All immunodominant peptides recognized contained previously described epitopes (Table 3.2). In addition, these peptides were also identified or reactive using the clade C peptide reagents. Sequence comparison between clade C and consensus M showed that previously defined epitopic regions were identical in 4/5 peptides (18, 19, 31 and 34) and different in 1/5 peptides. The dominant response to the TL9 epitope (peptide 47) could be due to the fact that this epitope is restricted by B*7 supertype (B*0702) which comprises about 15% of the South African population (Paximadis *et al*., 2011). HIV Gag peptides 48 and 78 (Table 3.2) were also identified as immunodominant in chapter 2 using clade-specific reagents.

Table 3.2. Immunodominant peptides recognized in the study

Peptide number	Sequence	Location	Number of Responders	Restricting HLA	Reference
Gag					
47	SEGATPQDLNTMLNT	p24	5	B*40:01 (B60)	Altfeld *et al*., 2000
48	TPQDLNTMLNTVGG	p24	7	B*0702 (B7)	Jin *et al*., 2000
78	RDYVDRFFKTLRAEQ	p24	4	Cw*03:04 (Cw10)	Honeyborne *et al*., 2010
Nef					
18	GFPVRPQVPLRPMT	Central	5	B*07:02 (B7)	Haas *et al*., 1996
19	RPQVPLRPMTYKAA	Central	4	B*35:01 (B35)	Culmann *et al*., 1991
28	KKRQEILDLWVYHT	Central	5	B*13:02 (B13)	Gray *et al*., 2009
31	HTQGYFPDWQNYTP	Central	4	B*15:01 (B62)	Culmann *et al*., 1989
34	TPGPGIRŶPLTFGW	Central	4	A*24:02 (A24)	Goulder *et al*., 1997

The underlined region denotes the pre-described epitope. The HLA in brackets denotes the supertype of the restricting HLA allele.

3.3.8. Reactive peptide maps of Gag and Nef

Reactive peptide maps were drawn to show differential epitope clustering across Gag and Nef protein regions with the peptides that were recognised using consensus M reagents (Figure 3.11). As mentioned, there was a higher peptide density in the relatively conserved p24 region of the Gag protein when compared to p17 or p15 (Figure 3.11A). Similarly, the more conserved central region of Nef protein had a higher reactive peptide density when compared to the outer regions of this protein (Figure 3.11B).

(A)

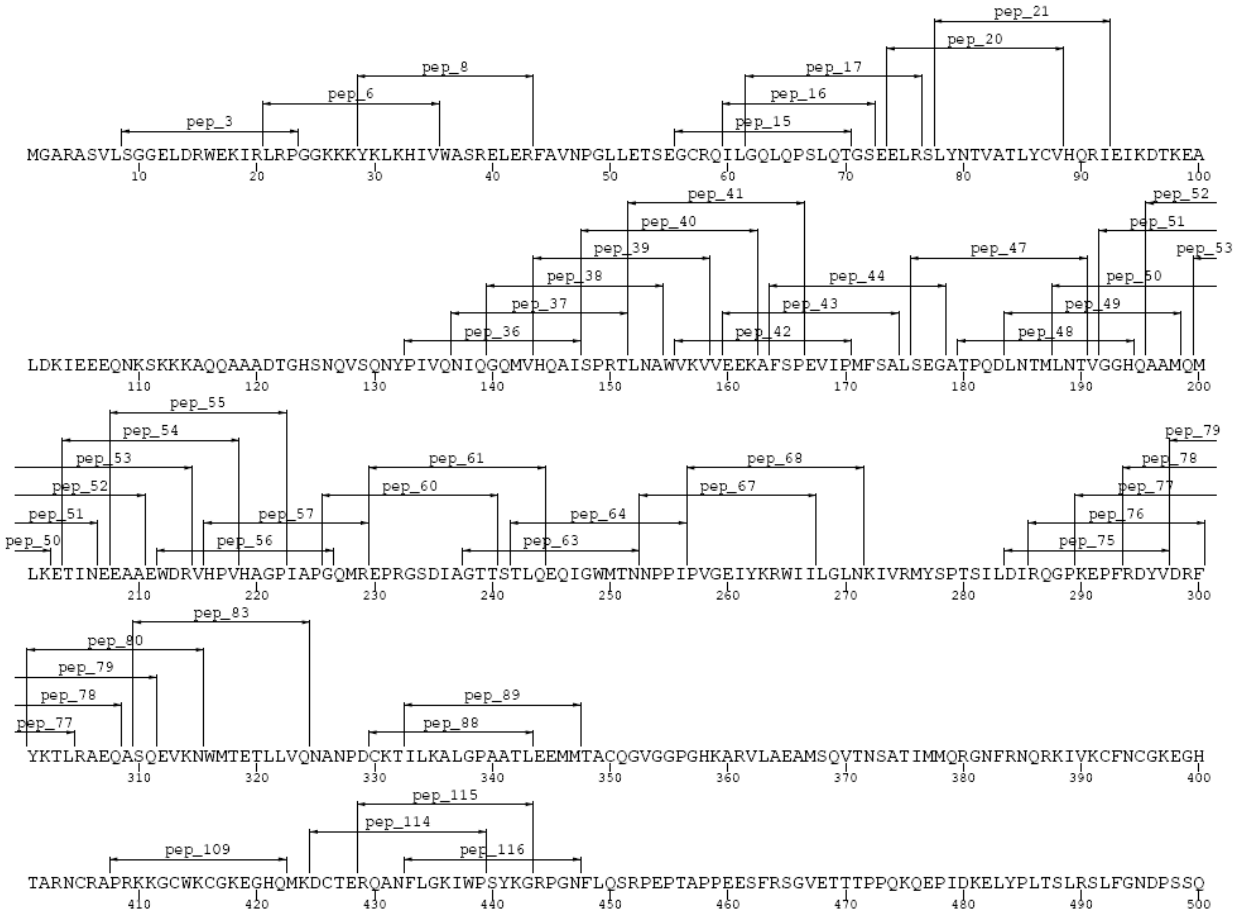


Figure 3.11. Reactive T-cell peptides mapped to within a region of 15 amino acids on the HIV proteins. (A) Reactive peptides identified using consensus group M Gag peptides. The locations of the reactive peptides identified by a positive IFN-γ ELISpot assay are indicated on the HXB2 protein sequences.

(B)

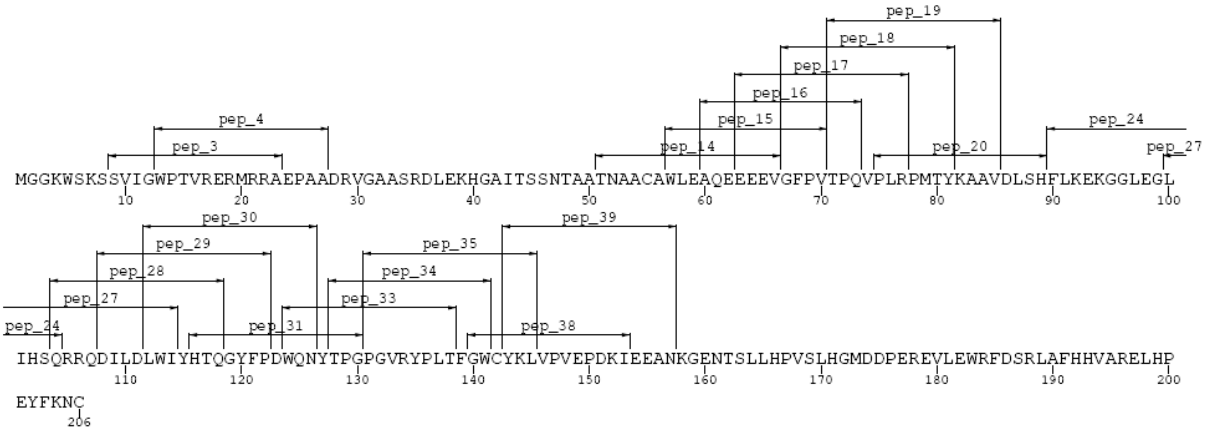


Figure 3.11. ***Continued.*** **(B)** Reactive peptide map showing peptides identified using consensus group M Nef peptides.

3.4. DISCUSSION

Given the high degree of diversity in Africa, it is not possible to make vaccines matched to all the subtypes circulating in different geographical regions. One way to overcome this problem is to design an immunogen based on a centralized sequence derived from the Group M HIV-1 viruses. However, this could come at a cost, as subtype coverage may be greater and this approach may result in a reduced number of responses at the subtype level compared to a clade-matched vaccine. In this chapter, we examined the reactivity of peptide sequences based on HIV-1 group M consensus sequences, and compared these to peptides based on subtype C (Du422 strain) in HIV-1 chronically infected individuals from a subtype C epidemic. This study focused on looking at cross-clade responses to Gag and Nef, as Gag has been associated with lower viral loads or better disease outcome (Edwards et al., 2002; Kiepiela et al., 2007), while Nef is the most immunodominant protein, at least in early subtype C infection (Gray *et al*., 2009).

As expected in this study, all the study individuals whose virus was characterized were infected with HIV-1 subtype C viruses. Previous studies have documented the South African HIV epidemic as being a monoclade epidemic, attibutable in >95% of cases to subtype C viruses (Iweriebor *et al*., 2011, Musyoki *et al*., 2011), although a minority of non-subtype C have recently been identified in South Africa (such as A1, B, F1, AD, AG and AC; Wilkinson and Engelbrecht, 2009). The median amino acid distance between the infecting viral sequences and the Gag subtype C peptide reagent compared to the Nef subtype C reagent was 6.7% and 14.4%, respectively, with the viruses being 1.6-fold more closely related to the subtype C_{Du422} sequence, a South African near-to C consensus isolate (Williamson *et al*., 2003), compared to the consensus group M sequence in the Gag protein (median amino acid difference of 6.7% and 11.0%, respectively). Furthermore, analysis of potential epitope coverage by these peptides predicted that subtype C_{Du422} reagents would detect 1.3-fold the breadth of consensus group M-specific T-cell responses for both Gag and Nef peptide reagents.

To determine how these differences would affect cross-clade reactivity, this study tested the ability of consensus group M and clade C peptides based on Gag and Nef to detect responses in a study population from South Africa, where the dominant circulating virus is clade C. We found that there was higher recognition of Gag than Nef. Particularly, we find that the median

magnitude and breadth of HIV-specific T-cell responses was 1.8-fold and 2-fold, respectively higher for the clade C reagent compared to the consensus group M Gag peptide reagent and higher than the 1.3-fold predicted by epitope coverage assessment tools. This suggests that group M-based peptide reagents can miss responses. Because the group M consensus sequence is central to all HIV-1 group M viruses, we would expect similar targeting of these proteins in the conserved regions across all group M viruses compared to the variable regions that have diversified in other subtypes. In this study we do find this, in that the immunodominant peptides were located in the more conserved regions of the proteins, namely the p24 for Gag and the central regions of Nef. Our findings are very similar to those generated from our partners in this multicentre study in Ugandan study (Serwanga *et al*., 2011). They found that 39% (50/129) and 34% (18/53) of clade A1 and D HIV-1 infected individuals recognised consensus group M Gag and Nef peptides respectively, compared to 39% (50/129) and 34% 18/53 in our study. Furthermore, they identified four peptides that were identified in this chapter as immunodominant, including two from Gag p24, TPQDLNTMLNTVGGH and RDYVDRFFKTLRAEQ, and two from the central core of Nef, PGIRYPLTFGWCFKL and KKRQEILDLWVYHTQ. However, unlike the study reported herein where the magnitude and breadth of responses differed significantly between Gag and Nef proteins, Serwanga *et al.* (2011) did not find significant differences in the magnitude and breadth of responses between the two proteins, possibly due to other factors such as HLA background and clinical characteristics of study participants since previous studies have shown that magnitude and breadth of HIV T-cell responses may be affected by antigen load (Geldmacher *et al*., 2007).

When the reactivity of consensus M Gag peptides was compared to that of inter-clade peptides, there was no significant difference observed in terms of magnitude and breadth. These data suggest that consensus group M Gag peptide reagents can be used inter-changeably with clade-mismatched reagents in the assessment of HIV-specific T-cell responses. However, they are less reactive in frequency, magnitude and breadth compared to clade-matched peptide reagents. This is in contrast to previous data from clade C-infected individuals in which consensus group M Gag and Nef peptide reagents performed better than inter-clade peptide reagents, and performed as well as the clade matching the infecting clade (Frahm *et al*., 2008). However, this could be due to the fact that the previous study had limited number of participants (only had 10 clade C-infected individuals) and possibly the different consensus sequences used.

In agreement with this study, and in contrast to our results, in another study (Bansal *et al*., 2006) peptides based on the clade matching the infecting clade (clade B) were shown to elicit responses of similar frequency (79%) compared to consensus group M (79%) and ancestral sequences (79%). However, although the frequency of recognition of clade B reagents was similar, in clade C infected participants they found that consensus sequences based on the clade not matching the infecting clade were less well targeted in terms of magnitude of responses compared to clade-matched, consensus and ancestral sequences. This was supported by a further study that compared the reactivity of HIV-specific T-cells among clade-specific reagents and centralised reagents, the frequency of recognition was equal (60%, 15/25 individuals) among the consensus B, centre of tree B, ancestor B, consensus M and consensus C peptide reagents (Malhotra *et al*., 2007). Only consensus A was less frequently recognised (48%, 12/25 individuals) when compared to the other peptide reagents, and this was attributable to lack of recognition of the TW10 epitope of the consensus A peptide set. The differences in results, together with loss of responses to specific epitopes, could to a certain extent be attributed to different HLA compositions in different populations, which have been shown to affect immunodominance (reviewed in Goulder and Watkins, 2008). Another possible reason for differences to two published studies that have looked at Gag responses to M consensus and compared them to infecting clades and mismatched clades is that these studies had limited sample sizes for clade C-infected individuals. Differences between their clade B infected individuals and our clade C infected individuals could be because M consensus reacts differently in these different epidemics, as pointed out by the study of (Frahm *et al*., 2008, where M consensus reactivity (magnitude and breadth) is higher in clade C infected individuals than in clade B infected individuals. In addition, differences in stages of infection, viral load and CD4 count could have also resulted in the observed differences in results between this study and previous studies. Our study has focused on Gag, which is a conserved protein, and Nef, and these results may not be applicable to more variable proteins such as Env. Overall, the protein being characterised and the diversity of the epidemic under study (*i.e.* which and how many virus clades are circulating) may influence the outcome of studies assessing the performance of group M consensus peptide reagents compared to other centralised peptide reagents such as centre-of-tree, ancestral and clade-specific peptide reagents.

The goal of this part of the multicentre African study was to assess the reactivity of group M Gag and Nef peptide reagents in a mono-clade epidemic that is predominantly clade C. The

results will be combined with those from other countries with different epidemics and HLA backgrounds within the populations.

The present study demonstrated that HIV-1 clade C-infected individuals can recognize peptides based on consensus group M Gag and Nef peptides. The magnitude and breadth of responses were higher for Gag compared to Nef. In addition, the pattern of recognition could be explained by the genetic relatedness of infecting viral sequences to peptide reagent sequences, in which Gag was more closely related to the infecting virus sequences compared to Nef. Furthermore, clade C (the clade-matched reagent) was significantly more highly recognized (in terms of magnitude and number of responses) compared to consensus group M peptide reagent.

Preliminary comparison with data from Uganda, another country involved in the multi-centre study, showed no significant differences when comparing the total magnitude and breadth of Gag and Nef responses; however the magnitude of responses may be related to time of infection, viral load and other factors. On the other hand, similarities were observed in terms of immunodominant peptides targeted between the two studies. This is also in line with unpublished results from Cameroon, a country demonstrating one of the greatest diversities of HIV-1, who is also participating in the multi-centre study (Tongo, Zembe, Burgers *et al*., manuscript in preparation). In that study 75% of individuals recognized at least one peptide from Gag and/or Nef and frequency of Nef recognition was slightly higher than Gag, at 82.5% compared to 79% for gag. Overall, there were no significant differences in the total magnitude of Gag and Nef responses. Interestingly, two Gag and four Nef peptides defined as immunodominant from Cameroon were also identified in our study as immunodominant (Marcel Tongo, personal communication). Overall, taking these three studies together, it is evident that individuals from diverse epidemics with different HLA backgrounds have T-cells that recognize certain regions of HIV that are conserved across group M viruses, as detected by consensus group M peptide reagents. This suggests the possibility of group M conserved regions as vaccine immunogens in populations of differing HLA background and HIV epidemics.

An important point to note, and a limitation of these types of studies, is that reactivity was attributable to a single immunological readout, namely IFN-γ production as measured by the ELISpot assay. HIV-specific T-cells have the ability to produce other cytokines and perform other functions, such as proliferation and cytotoxicity that may be important in their protective

ability. Also, it is not known whether the specific epitopes targeted would translate to antiviral activity *in vivo*. Therefore, further studies such as viral inhibition assays and multicolor flow cytometry assays to explore other functions and characterize the identity of T-cells specific for immunodominant, cross-reactive central sequences are important for assessing their true cross-reactive and antiviral ability. The present study is the first to assess reactivity of group M peptides reagents and compare them to clade-matched and mismatched reagents in a large cohort of clade C infection from South Africa, and comprehensive anaylsis of the combined results of the five countries participating in the multi-centre African study will identify important regions to target. These studies will assist in global immunogen design for HIV-1 vaccine development.

CHAPTER 4

INVESTIGATING THE FUNCTIONAL POTENTIAL OF CROSS-REACTIVE HIV EPITOPE VARIANTS

Table of Contents

4.0. INTRODUCTION

HIV-1 exhibits a high degree of sequence variability, and this poses a major challenge to the development of an effective vaccine against the virus. A range of studies have described cross-clade responses, that is, epitopes that differ between HIV clades but remain recognizable and able to stimulate HIV-specific memory T-cells in *in vitro* assays (Geels *et al*., 2005; Geldmacher *et al*., 2007; Gupta *et al*., 2006; McKinnon *et al*., 2005; Yu *et al*., 2005). Indeed, Chapters 2 and 3 of this thesis described matching peptides from different HIV clades with variant amino acid sequences that are recognized by T-cells despite variation in the epitopes. However, the vast majority of these studies, including those described in this thesis, measure cross-reactivity using the IFN-γ ELISpot assay, thus relying on measuring the reactivity of peptides by exogenously-added peptides (often in vast excess to physiological levels), and relying on one function of T-cells, namely cytokine production of IFN-γ.

In an attempt to understand immune correlates of favourable disease outcome or viral control in natural HIV infection, data from several studies comparing HIV-specific adaptive immune responses suggest that there are other aspects of the T-cell immune response to HIV that may require further elucidation apart from IFN-γ production. HIV-specific memory T-cells are capable of producing other cytokines such as IL-2 and TNF-α, chemokines for example MIP-1β, cytotoxicity through upregulation of perforin and/or granzyme B as well as degranulation through CD107a production (Betts *et al*., 2006; De Rosa *et al*., 2004; Roederer *et al*., 2004). On that note, previous studies have shown that polyfunctional CD8+ T-cells are found in significantly higher frequencies in LTNPs compared to chronic progressors (Betts *et al*., 2006). In that study, the quality of CD8+ T-cells were assessed by measuring five functions, namely degranulation, IFN-γ, MIP-1β, TNF-α and IL-2 simultaneously in chronically HIV-infected individuals and compared to elite controllers. The proportion and numbers of CD8+ T-cells with more functions inversely correlated with viral load in the progressors, showing that the quality of HIV-specific CD8+ T-cells is important in immune control of HIV infection (Betts *et al*., 2006). Additional studies have further assessed the function of CD8+ T-cells and identified the upregulation of a cytotoxic molecule, namely perforin, as a correlate of HIV elite control (Hesperger *et al*., 2010). In this cross-sectional study of chronically HIV infected individuals, CD8+ T-cells from elite controllers were able to

upregulate perforin consistently compared to viremic controllers, chronic progressors and viremic nonprogressors (Hesperger *et al*., 2010).

Other studies have shown that proliferation of T cells in response to HIV-1 antigens is associated with less progressive HIV-1 infection (Huang *et al*., 2008; Migueles *et al*., 2002; Younes *et al*., 2003). In one of these studies, it was the proliferative capacity of CD4+ T-cells before treatment interruption that correlated with CD4+ T-cell count decline during treatment interruption (Huang *et al*., 2008). In another study, CD4+ T-cell Gag- and Nef-proliferative responses were higher in aviremic compared to viremic study individuals (Younes *et al*., 2003). Furthermore, a separate study showed that the proliferative capacity of CD8+ T-cells of non-progressors is maintained compared to that of progressors (Migueles *et al*., 2002). Interestingly, this study also demonstrated that this was coupled to increased perforin production. This data may suggest that it is the combination of specific qualitative aspects of HIV-specific CD8+ T-cell functions that is associated with immune control of HIV infection.
A number of studies have suggested that in addition to polyfunctionality and proliferative capacity of HIV-specific T-cells being important in immune control of HIV, other aspects of HIV-specific T-cells, such as their functional avidity for the epitopes recognized, may be a major factor in the ability of these T-cells to produce more than one function and to suppress HIV replication in slow disease progressors (Almeida *et al*., 2009). These authors showed that CD8+ T-cells that had higher polyfunctionality and suppressive activity were highly sensitive to antigen, that is, they had higher avidity compared to less polyfunctional and less suppressive HIV-specific CD8+ T-cells. Because this study used purified CD8+ T-cell clones specific for one epitope restricted by HLA*B27:05, there were no artefacts due to heterogeneity in terms of different specificities and restriction by different HLA alleles of different T-cell clones, as might have been the case for previous studies that found contrasting results on the importance of antigen sensitivity on the functional profile of CD8+ T-cells (Ueno *et al*., 2004; Yang *et al*., 2003). Functional avidity has also been proposed to predict antiviral efficiency (Alexander-Miller *et al.,* 1996; Berger *et al*., 2011; Derby *et al*., 2001).

Current vaccine approaches promising to cater for the divergent forms of the virus involve inclusion of variant forms of the viral epitopes (‘mosaics’) to elicit broadly reactive T-cell immune responses (Barouch *et al*., 2010; Fischer *et al*., 2007; Santra *et al*., 2010). However,

little is known about the ability of HIV-specific T-cells recognizing these variant epitopes to mediate different T-cell functions, and the nature of responses at physiologically relevant concentrations of the epitopes. Recent published studies indicate that despite broad cross-reactivity to exogenously loaded HIV peptides in ELISpot assays, T-cell suppression of HIV-1 with the corresponding epitope was significantly impaired (Bennett *et al*., 2008). In addition, the functional profile of epitope variants differed with respect to levels of cytokine production (Malhotra *et al*., 2009), and importantly, ELISpot reactivity did not necessarily correlate with *in vitro* viral suppression (Valentine *et al*., 2008).

The hypothesis that was explored in this chapter was that HIV-specific T-cells recognizing corresponding epitope variants that differed in functional avidity would also differ in their functional profile and proliferative capacity. To test this, T-cells specific for variant HIV peptide pairs (*e.g.*, a peptide from subtype C and the corresponding one from Consensus group M that had amino acid mismatches but both detected HIV-specific responses in the ELISPOT assay) were assessed for their cytokine production profiles by measuring IFN-γ, IL-2, TNF-α and MIP-1β; their ability to up-regulate cytotoxic molecules, namely granzyme B, perforin and the degranulation marker CD107a; as well as their proliferative capacity, in a flow cytometry-based proliferation assay. The results from this study may shed more light on the true cross-reactive nature of HIV-specific CD8+ T-cells, which may be important for vaccine immunogen design.

4.1. MATERIALS AND METHODS

4.1.1. Study participants

Seventeen HIV-1 infected study participants were involved in this part of the study. These were all study individuals described in Chapters 2 and 3. The clinical characteristics of these individuals have been provided in these chapters. The individuals were first screened for HIV-1 consensus group M- and subtype C-specific T-cell responses using a pool and matrix IFN-γ ELISpot assay. Responses were mapped to individual reactive peptides and corresponding reactive peptides in each study individual were selected for subsequent analyses in this chapter.

4.1.2. Peptides

Peptides that gave a positive response in the IFN-γ ELISpot assay were selected and their corresponding variants that were available in the laboratory were tested for reactivity in a second IFN-γ ELISpot assay. Peptide variant pairs that were reactive with amino acid mismatches within and in flanking regions of pre-described epitopes were used to stimulate the same donor PBMC in peptide dilution ELISpot assays, for intracellular cytokines staining assays and proliferation assays. Apart from peptide dilution assays, where a range of concentrations was used, peptides were used at a final concentration of 1μg/ml in all assays.

4.1.3. Peptide dilution IFN-γ ELISpot assay

Predicted reactive peptides determined from the pool and matrix ELISpot screening assays were tested individually to confirm their reactivity. Limiting dilution ELISpot assays were performed on 16 study individuals with 17 peptide variant pairs using the IFN-γ ELISpot assay as described in Chapter 2 section 2.2.5, in order to investigate whether amino acid changes had an effect on functional avidity of the peptides, and whether functional avidity was related to the magnitude of response to the recognized peptides. The peptide variants were titrated in five ten-fold serial dilutions (from 10^{1} to 10^{-3}μg/ml) for each peptide variant,

and were performed in duplicate. A functional avidity of ≥2-fold between peptide variants was empirically considered different.

4.1.4. Antibodies and reagents

Costimulatory antibodies, anti-CD28 and anti-CD49d (BD Biosciences; San Diego, California) and Golgi transport inhibitor, Brefeldin A (Sigma-Aldrich; St. Louis, MO) were used at the beginning of each of the stimulation. For the cytokine staining panel, an antibody cocktail for surface staining included anti-CD4 FITC, and anti-CD8 PerCP Cy5.5 in a total of 50µl with 1% wash buffer per reaction. The intracellular staining (ICS) antibody cocktail included anti-CD3 APC Cy7, anti-IFN-γ Alexa 700, anti-IL-2 APC, anti-TNF-α PE Cy7 and anti-MIP-1β PE in a total of 50µl of 1 x Perm Wash buffer (BD Biosciences; San Jose, California).

For the cytotoxic panel, anti-CD107a FITC (BD Biosciences; San Jose, California) was included at the beginning of the stimulation in addition to costimulatory antibodies. The cocktail for surface staining included anti-CD4 PE Cy7, and anti-CD8 PerCP Cy5.5 in a total of 50µl 1% wash buffer (PBS containing 1% FCS; Fischer Scientific; Pittsburgh, Pennsylvania) and 0.1% sodium azide (Fischer Scientific; Pittsburgh, Pennsylvania) per reaction. The intracellular cytokine staining antibody cocktail included anti-CD3 APC Cy7, anti-IFN-γ Alexa Fluor 700, anti-Perforin PE (clone B-D48 that can recognize perforin in its various conformations and therefore can also detect newly formed perforin, Hesperger *et al*., 2008) and anti-Granzyme B Alexa 647 in a total of 50µl of 1 x Perm Wash buffer (BD Biosciences; San Jose, California) per reaction. All antibodies were titered to optimal concentrations for the assays.

Pairs of corresponding peptides with amino acid mismatches used in ELISpot peptide dilution assays were used for stimulations. The same concentration (1µg/ml) that gave a positive IFN-γ response during screening and confirmatory IFN-γ ELISpot assay was used for stimulations during ICS assays.

4.1.5. Stimulation

The stimulation mix included 4.5ml Rosewell Park Memorial Institute (RPMI) medium with 10% FCS, anti-CD28 and anti-CD49d at final concentrations of 0.004µg/ml and 0.001µg/ml respectively, and DNase (0.2mg/ml). After resting for at least 4 hours, PBMC were centrifuged for 10 minutes at 1200rpm and the pellet re-suspended the stimulation mix and divided into half, for staining with a cytokine panel a cytotoxic panel. For the cytokine panel, approximately 10^6 cells (in 195µl) were plated into v-bottomed wells and 5µl of peptide added into each stimulation at a final concentration of 1µg/ml, and set up in duplicate. An equal volume of RPMI containing 10% FCS and co-stimulatory antibodies was added to the negative, unstimulated well. For the cytotoxic panel, 193µl of cells were plated into each well and the same peptides added as for the cytokine panel. Anti-CD107a FITC was added to each well, as this needs to be included during stimulation. Cells were mixed and the plate was incubated at 37^0C and 5% CO_2 for 16 hours.

4.1.6. Intracellular cytokine staining

After 16 hours of stimulation, cells were washed twice with PBS by centrifugation at 1000g for 3 minutes at 4^0C. Pre-titered LIVE/DEAD Fixable Violet Dead Cell Stain (Vivid, Applied Biosystems, Invitrogen, USA), a viability marker was added in a total volume of 50µl and cells incubated in the dark for 20 minutes, after which they were washed twice with FACS wash buffer at 1000g for 3 minutes at 4^0C. Surface antibody cocktail, containing anti-CD4 FITC and anti-CD8 PerCP Cy5.5 in a total of 50µl FACS wash buffer for the cytokine panel and containing anti-CD4 PE Cy7, anti-CD8 PerCP Cy5.5 in a total of 50µl FACS wash buffer for the cytotoxic panel, were added to each well and incubated in the dark for 20 minutes. Cells were washed twice with FACS wash buffer at 1000g for 3 minutes and 4^0C. Cytofix/Cytoperm (100µl, BD Biosciences; San Jose, California) was added and cells incubated for 20 minutes in the dark, after which they were washed twice with Permwash buffer (BD Biosciences; San Jose, California) and stained with 50µl of intracellular cytokine antibody cocktail for 20 minutes in the dark. For the cytokine panel, the antibody cocktail included anti-CD3 APC Cy7, anti-IFN-γ Alexa Fluor 700, anti-IL-2 APC, anti-TNF-α PE Cy7, anti-MIP-1β in a total of 50µl of FACS wash buffer. For the cytotoxic panel, the

cocktail included anti-CD3 APC Cy7 (stained intracellularly due to internalization upon stimulation by antigen), anti-IFN-γ Alexa Fluor 700, anti-Perforin PE and anti-Granzyme B Alexa 647 in a total of 50µl of FACS wash buffer per well. Cells were washed twice with Permwash buffer at 1000g for 3 minutes at 4^0C and were resuspended in 150µl Cell Fix (BD Biosciences; San Jose, California) and then acquired on an LSRII flow cytometer (BD Biosciences; San Jose, California).

4.1.7. Oregon Green proliferation assay

After resting for at least 6 hours, PBMC were washed twice with PBS. CellTrace Oregon Green 488 (10µg/ml, Molecular Probes, Invitrogen, USA) was added to 1 x 10^7 cells and incubated for 4 minutes in the dark. Cells were mixed and incubated for another 3 minutes in the dark, after which they were vortexed at high speed for 10 seconds and an equal volume of PBS was added, followed by a further incubation for 3 minutes in the dark. PBS was added up to 15 ml and the cells were centrifuged at 571g for 10 minutes. The cells were resuspended in RH10 (RPMI with 10% human AB serum, Sigma-Aldrich; St. Louis, MO) such that 195µl of cells could be plated per well with 200 000 cells/well, with four wells for each peptide variant. Corresponding peptide pairs with amino acid mismatches were used to stimulate cells at a final concentration of 1µg/ml in a 200µl reaction volume. A negative control well was included, to which RH10 was added in the place of peptide. Cells were incubated at 37^0C, 5% CO_2 for 6 days. On day 6, four wells were combined for each peptide stimulation and centrifuged at 857g for 5 minutes. Cells were stained for surface markers namely CD4 and CD8, as well as for CD3 intracellularly, as described in the previous section. Samples were then acquired on an LSRII flow cytometer (BD Biosciences; San Jose, California)

4.1.7. Data acquisition and statistical analyses

Approximately 500 000 events were acquired for each stimulation on a three-laser LSRII flow cytometer using BD FACS DIVA software (BD Biosciences, San Jose, California). Data were analysed using FlowJo version 9.3 (TreeStar, Ashland, Oregon) and Spice version 5.2 (Dr Mario Roederer, NIH, Bethesda, Maryland). For the upregulated perforin and granzyme

B, the perforin or granzyme gates were copied into either the CD107a gate or IFN-γ gate, so that perforin or granzyme produced together with either CD107a or IFN-γ could be identified. Results reported are background subtracted. A response was considered positive if it was twice above background and above 0.025% after background subtraction, and with more than 10 events above background.

4.2. RESULTS

4.2.1. Functional avidity of corresponding peptide variants

Although amino acid mutations in corresponding epitope variants have been suggested to be key factors leading to loss of recognition of these epitope variants, the effect of these mutations in corresponding peptide variants that remain recognized by T-cells has not been well characterized. Firstly, the effect of amino acid mismatches on the functional avidity of HIV-specific T-cells was investigated. Functional avidity was defined as the peptide concentration that gave rise to half the maximum IFN-γ response.

Functional avidity, shown in Figure 4.1-4.3, was classified into three categories. Firstly, 8/17 (47%) of corresponding peptide pairs with amino acid mismatches had differences in the magnitude of HIV-specific IFN-γ T-cell responses, and subsequently a more than 2-fold difference in their functional avidity (Figure 4.1). Three out of these seven (38%) peptide pairs had amino acid mismatches involving non-conservative mutations, in participants ZA023, ZA030 and ZA033 (Figure 4.1). There were 5/8 (63%) of these whose functional avidities were different despite conservative amino acid mutations between peptide pairs, that is, the amino acid mismatch involved substitution by an amino acid with a closely related side chains (ZA031, ZA035, ZA034, TM16 and ZA029, Figure 4.1).

In the second category were 3/17 (18%) of peptide pairs that had differences in the magnitude of IFN-γ T-cell responses but did not show differences in their functional avidity (Figure 4.2). Further assessment of these four peptide pairs showed that 2/3 (67%) of them had amino acid mismatches involving very closely related amino acids (TM16 and NM07), while in 1/3 (33%) peptide pairs the mismatch involved distantly related amino acids (ZA006, Figure 4.2). It was interesting to note that NM07 recognized the variant with higher magnitude than the infecting sequence (Figure 4.2). This could possiblly be due to the fact that the sequence that is now a variant sequence in the individual represents the original sequence that was in the transmitting host and had an epitope that is more avid than the infecting sequence in NM07 or simply the mutation resulted in an epitope that binds stronger than the wild type sequence.

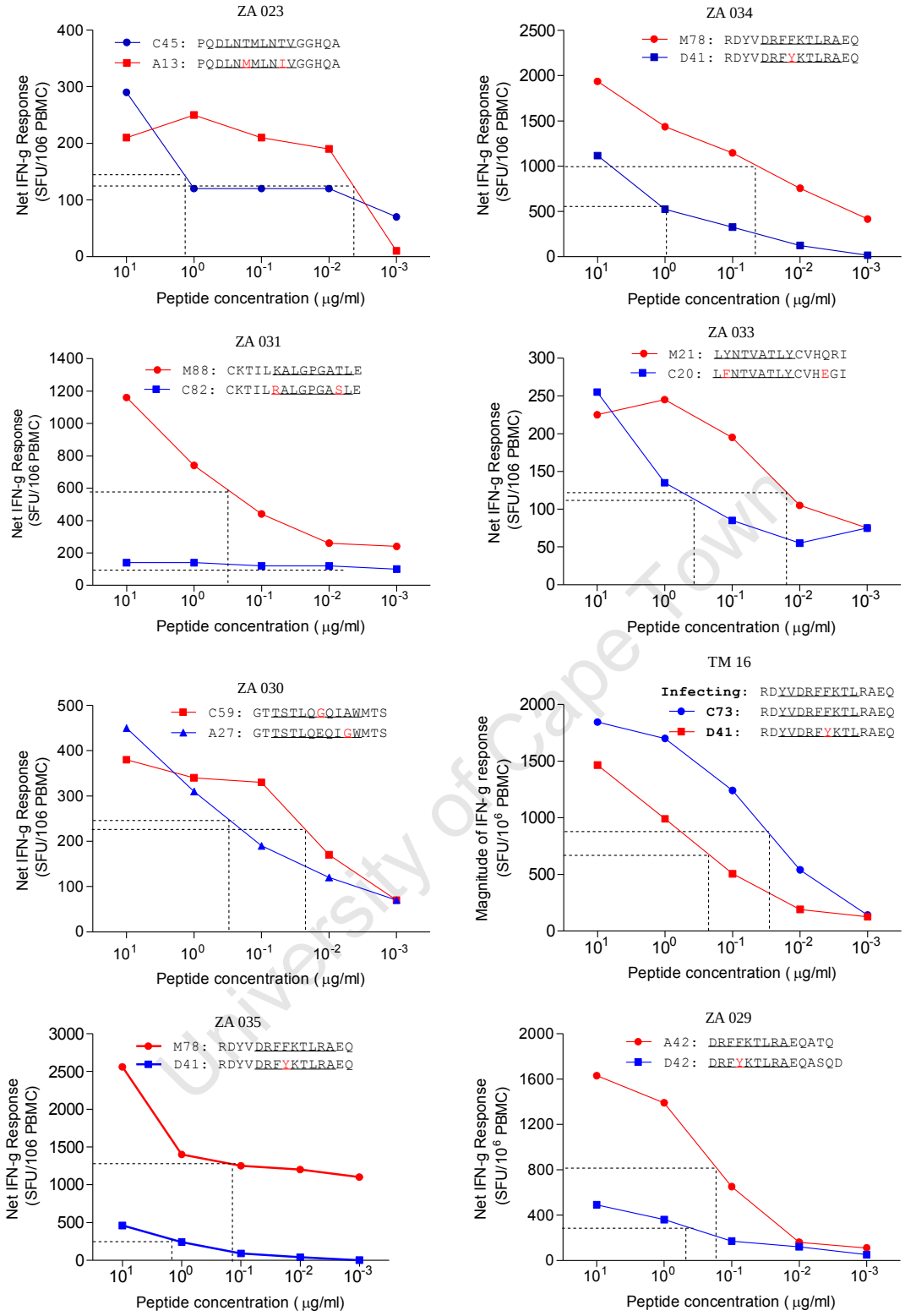


Figure 4.1. Peptide dilution assay for corresponding peptide pairs that differed in magnitude and functional avidity. IFN-γ responses were determined by ELISPOT assay at five peptide concentrations. The number at the top of each figure shows the sample tested. Functional avidity was defined as the peptide concentration giving rise to half the maximum IFN-γ response, and is denoted by the dotted lines. A difference in functional avidity was defined as >2-fold. The underlined region of the peptide (indicated by the peptide number shown before the sequence) denotes the predicted epitope. The red letters denote amino acid mismatch between corresponding peptide variants. In individuals where the infecting virus was sequenced, the sequence is shown.

In the third category, 6/17 (35%) of tested peptide variant pairs had a similar magnitude of response at almost all peptide concentrations tested (Figure 4.3) and no difference in functional avidity. Only 2/6 (33%) of these peptide pairs had amino acid mismatches involving non-conservative mutations (MM15 and ZA011) while the majority involved closely related amino acids (ZA001, ZA028, ZA035, ZA023, Figure 4.3).

Interestingly, in one case where the same peptide variant pairs were recognised in different individuals (C45 and A13 in ZA006 and ZA023), these peptides behaved distinctly, in one case showing similar avidity, and in the other very different avidity. This may suggest that different HLA alleles were restricting the same peptides in these two people, and the amino acid changes affected them differently, or the T cell receptor usage was different between them.

In summary, these results show that corresponding peptides that are both recognized and even have similar magnitude of T-cell responses in an ELISpot assay at the standard, high concentrations typically employed in these assays, may have different functional avidities, and suggests that some of these peptides may not be recognized *in vivo*, for example, in ZA035 the M78 variant continues to stimulate a response at low peptide concentrations while the corresponding variant D41 ceases to be detectable (Figure 4.1). A summary of the results for the participants that had samples available for further study is shown in Table 4.1, and these were further investigated for functional attributes in the next sections.

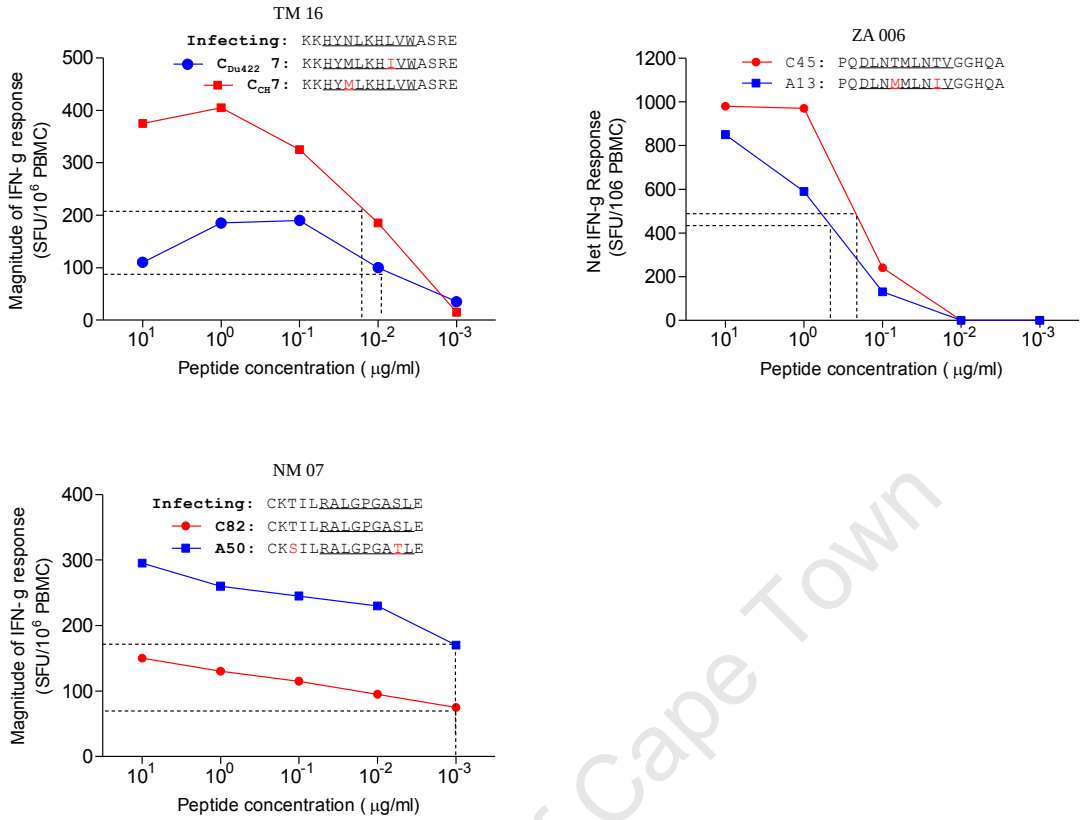


Figure 4.2. Functional avidity of corresponding peptide pairs that demonstrated similar functional avidity despite differences in total magnitude of IFN-γ response. Similar functional avidity was defined as that which is <2-fold between corresponding peptide pairs. Functional avidity is denoted by dotted lines. The infecting virus sequence was available for participants TM16 and NM07. The underlined region denotes the predicted epitope. The red letters denote amino acid mismatch between corresponding peptide variants.

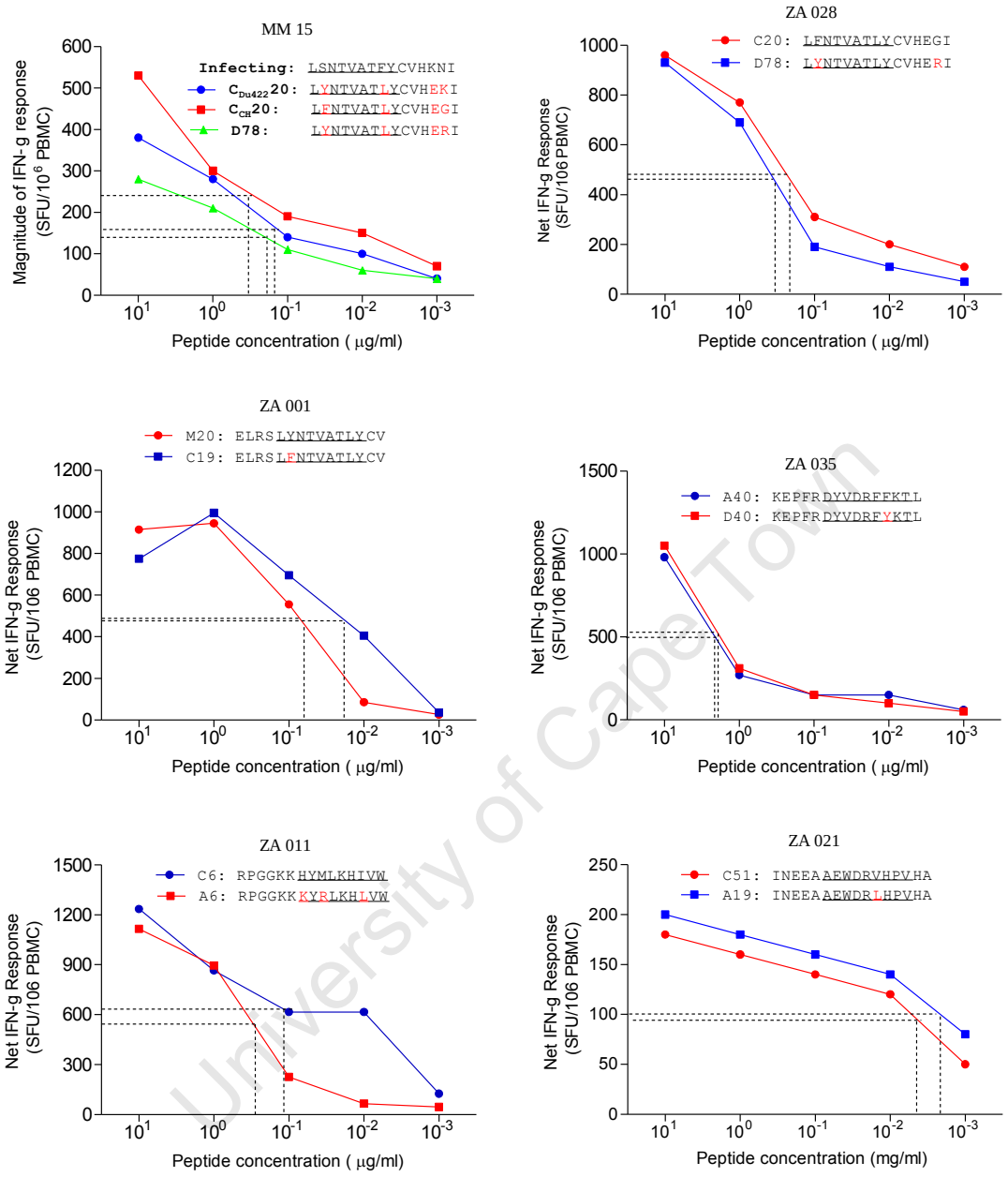


Figure 4.3. Functional avidity of peptide variants which demonstrated similar functional avidity and magnitude of IFN-γ response. Dotted lines denote functional avidities, which is the peptide concentration yielding half maximum IFN-γ response. Amino acid mismatches are denoted by red letters and the underlined region denotes the pre-described epitopic region within the tested peptides. The infecting virus sequence was available for participant MM15.

Table 4.1. Summary of functional avidities of peptide pairs investigated in the study

PID	Peptide	Sequence	Location	HLA	Avidity (EC_{50}, µg/ml)	Fold difference in avidity	Reference
ZA023	C45	PQDLNTMLNTVGGHQA	p24	B*14:02	1.259	**356**	Gillespie *et al.*, 2002
	A13	PQDLNMMLNIVGGHQA			0.004		
ZA029	A42	DRFFKTLRAEQATQE	p24	B*14:02	0.141	**3**	Reche *et al.*, 2006
	D42	DRFYKTLRAEQASQD			0.447		
ZA030	A27	GTTSTLQEQIGWMTS	p24	B*58:01	0.158	**10**	Bhattacharya et al., 2001
	C59	GTTSTLQGQIAWMTS			0.016		
ZA033	M21	LYNTVATLYCVHQRI	p17	A*29:02	0.014	**23**	Boutwell *et al.*, 2007
	C20	LFNTVATLYCVHEGI			0.316		
ZA006	A13	PQDLNMMLNIVGGHQA	p24	B*14:02	0.398	1.8	Gillespie *et al.*, 2002
	C45	PQDLNTMLNTVGGHQA			0.224		
ZA021	C51	INEEAAEWDRVHPVHA	p24	B*40:06	0.004	1.33	Thakar *et al.*, 2005
	A19	INEEAAEWDRLHPVHA			0.003		
ZA028	D78	LYNTVATLYCVHERI	p17	A*29:02	4.467	1.41	Boutwell *et al.*, 2007
	C20	LFNTVATLYCVHEGI			3.162		
ZA011	C06	RPGGKKHYMLKHIVW	p17	B*24:02	0.102	1.95	Llano *et al.*, 2009
	A06	RPGGKKKYRLKHLVW			0.199		
ZA035	M78	RDYVDRFFKTLRAEQ	P24	B*14:02	O.126	**10**	Reche et al., 2006
	D41	RDYVDRFYKTLRAEQ			1.259		

The letters in red denote the amino acid mismatch between corresponding peptide pairs. The underlined region shows the predicted epitope region based on data in Los Alamos National Laboratory (LANL) immunology database (http://www.hiv.lanl.gov/content/immunology). Fold difference in avidity was obtained by dividing the larger EC_{50} by the smaller EC_{50} for each peptide pair, where EC50 is the peptide concentration yielding a half maximum peptide-specific IFN-γ response in a peptide dilution ELISpot assay. The cited reference is one of the studies in which the underlined epitope was previously described. Avidity differences ≥2-fold are bolded.

4.2.2. Cytokine production by HIV-specific T-cells

Having observed that amino acid mismatches in peptide variants had a differential effect on the magnitude of IFN-γ responses and functional avidity, peptide variant pairs were selected (Table 4.1) and the cytokine profiles of their specific T-cells were investigated. Samples were selected based on PBMC availability. The cytokine profile of these peptide variants was assessed using intracellular cytokine staining and multiparameter flow cytometry, measuring four cytokines, namely IFN-γ, IL-2, MIP-1β and TNF-α. The gating strategy is shown in Figure 4.4 and a layout for each study participant is shown in Appendix E1, Figures E1.1-E1.7.

Only HIV-specific responses by CD8+ T cells were detected, with no CD4 responses detected in the samples and peptides that were studied. Responses were not detected in one study individual (ZA011), possibly due the fact that the responses were below the level of detection of ICS assays, leaving seven individuals for analysis. Overall, HIV-specific CD8+ T-cell responses were highest for MIP-β with a median of 0.27% (range 0.09-2.04%), followed by IFN-γ, with a median of 0.23% (range 0.08-1.35%; data not shown). IL-2 and TNF-α production were the weakest responses in the study, both with a median of 0.05% (range 0-0.83% and 0-0.06%, respectively).

We next determined whether there were differences in the cytokines produced in response to each peptide variant. Peptide pairs demonstrated discordant effects of amino acid mismatches on the cytokine profile of HIV-specific T-cells recognizing them. In 4/7 study individuals, namely ZA006, ZA021, ZA028 and ZA033 (Figure 4.5), similarities were observed in the profile of cytokines produced in response to each peptide variant. Interestingly, 3/4 of these study individuals demonstrated similar functional avidities of HIV-specific T-cells recognizing their respective peptides pairs, with only ZA033 exhibiting more than a log difference in functional avidity (Table 4.1).

On the other hand, three study individuals demonstrated differences in the profile of cytokines produced by their HIV-specific T-cells upon stimulation by peptide variants. These individuals, namely ZA023, ZA029 and ZA030, had HIV-specific T-cells producing different combinations of cytokines in response to their respective peptide variants. In two cases (ZA029 and ZA030), the peptide variant with higher avidity detected an additional TNF-α

response that was not produced in response to the lower avidity variant (Figure 4.5). ZA023 was an interesting case, where the higher avidity peptide A13 detected higher MIP-1β and IFN-γ responses, but the C45 peptide was able to detect additional TNF-α and IL-2 responses (Figure 4.5).

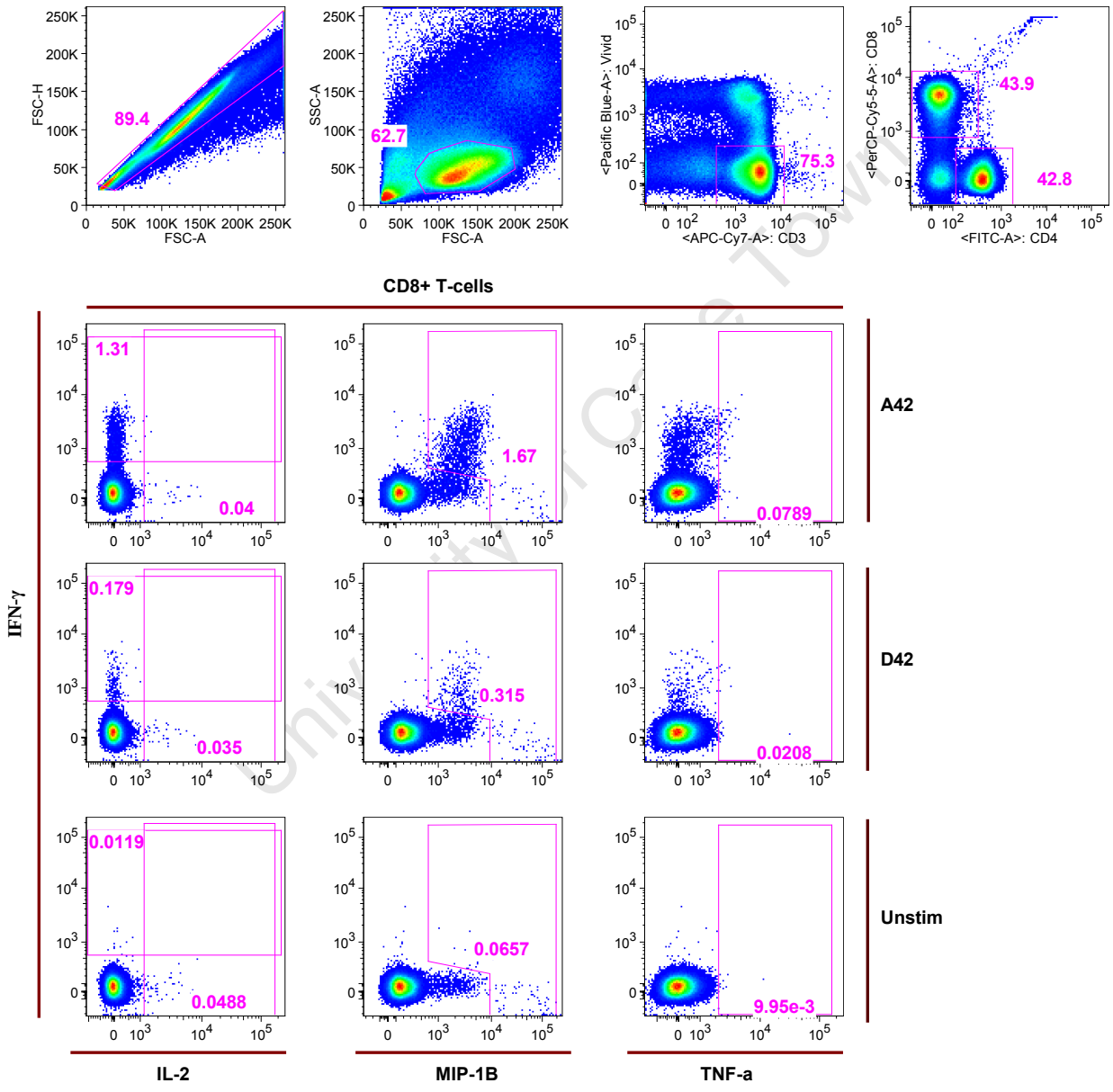


Figure 4. 4. Cytokine profiles of HIV-specific CD8+ T-cells in response to corresponding peptide variants with amino acid mismatches. HIV-specific T-cells were assessed for their cytokine production profiles in response to peptide variants. The upper row of flow plots demonstrates the gating strategy. This is a representative figure for study individual ZA029 after 16 hour stimulation with two corresponding Gag peptide variants A42 and D42. The proportion of cytokine producing CD8+ T-cells is indicated in each plot. The last row of the bottom panel shows unstimulated cells (Unstim). This result is representative of two independent experiments for each peptide variant. Gates were adjusted to exclude high MIP-1β background (middle column).

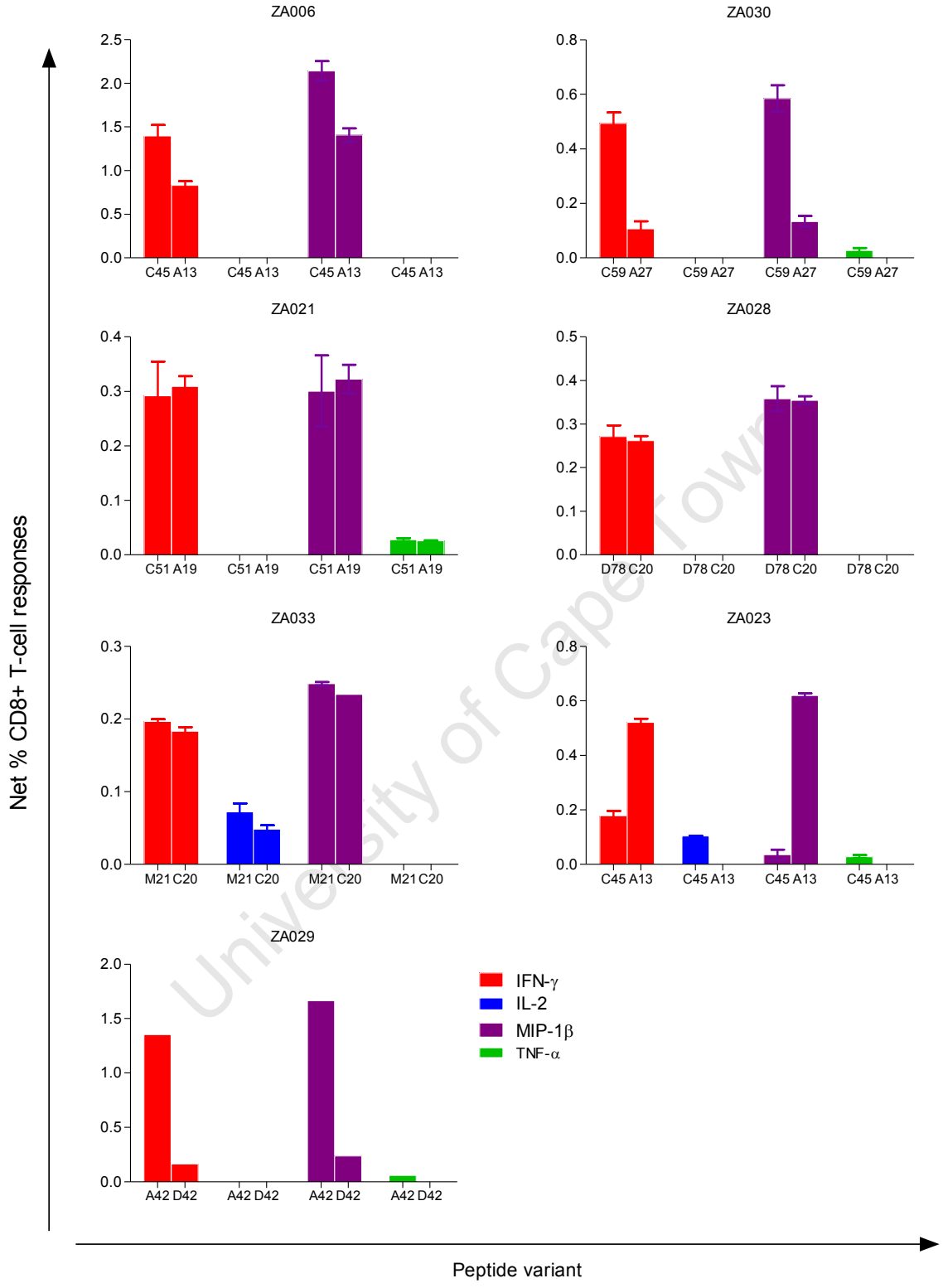


Figure 4.5. The cytokine profile of HIV-specific T-cells in response to stimulation with peptide pairs with amino acid mismatches. The participant number is indicated on top of each graph.

Having observed discordant effects of avidity and amino acid mismatches on the cytokine profile of HIV-specific T-cells recognizing these peptides, the median fluorescence intensity (MFI) of intracellular cytokines was compared (Figure 4.6). The MFI indicates the amount of cytokine produced by each responding cell. There were no significant differences in the per-cell cytokine producing capacity for the peptide pairs for all the cytokines investigated, namely IFN-γ ($p = 0.2969$), IL-2 (not applicable), MIP-1β ($p = 0.4688$) and TNF-α ($p = 0.3125$).

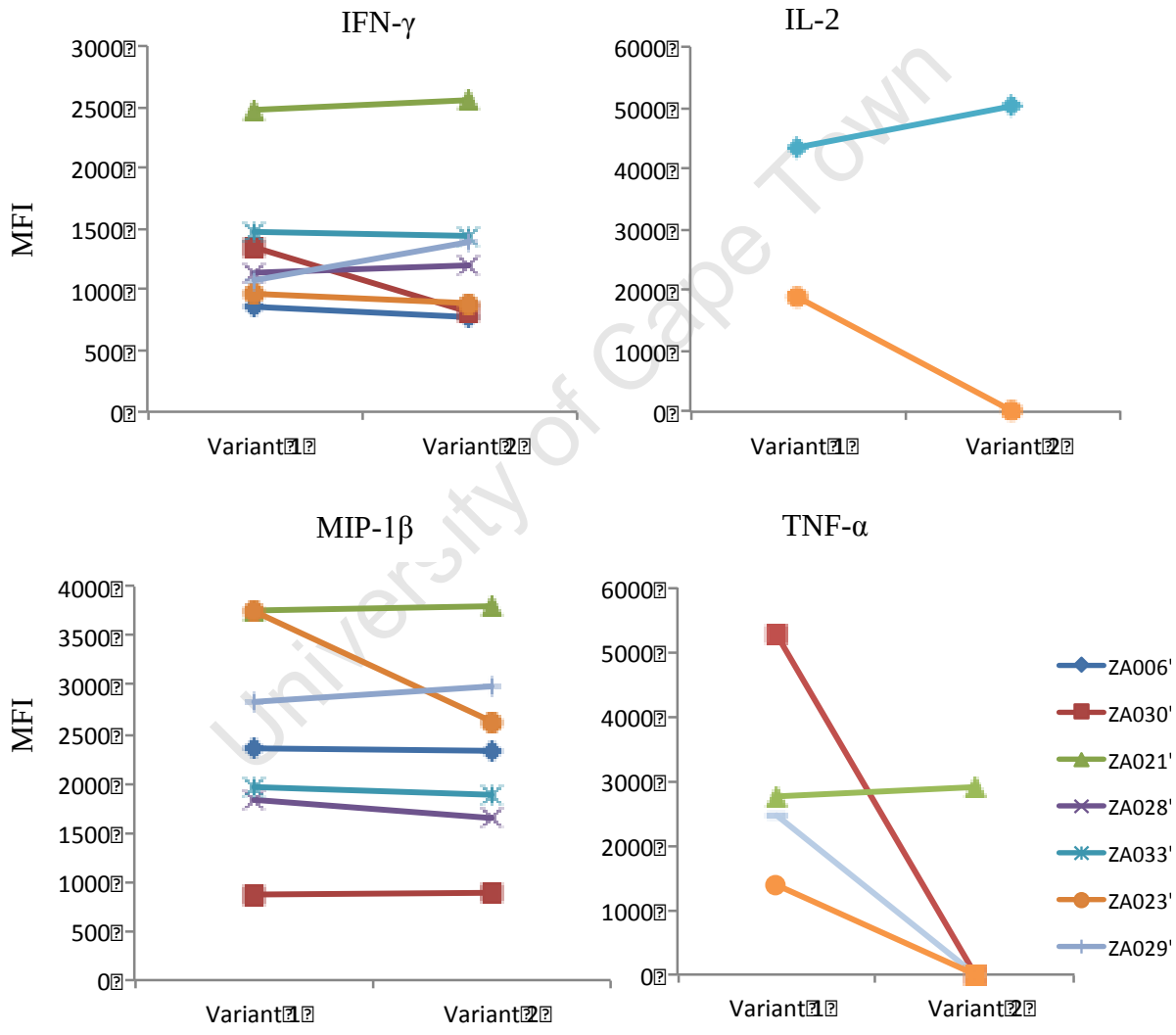


Figure 4.6. Comparison of MFI values of cytokines produced in response to corresponding peptide variants in seven study individuals. Each study individual is colour coded and for each study individual, variant 1 is the peptide variant that had a higher functional avidity, while variant 2 is the corresponding variant that had a lower functional avidity. Responses that did not qualify for positive criteria had their MFI set to zero if the other variant was positive. When both variants were negative for a particular cytokine, that cytokine was not plotted for that individual.

4.2.3. Polyfunctionality of cross-reactive HIV-specific T-cells

The next level of characterization of the cytokine profile of HIV-specific T-cells recognizing epitope variants was the investigation of the polyfunctionality of these cells, or the ability to produce multiple or combinations of cytokines simultaneously. Only 1/7 study individuals (ZA023) had HIV-specific T cells that produced all the three cytokines and one chemokine (four functions) tested in response to one of the peptide variants (Appendix E1). Four out of seven individuals had HIV-specific T cells producing three functions. These functions were predominantly (in 3/4) IFN-γ, MIP-1β and TNF-α, while in 1/4 individuals, IFN-γ, IL-2 and MIP-1β were produced. There were 2/7 individuals who demonstrated two functions from their HIV-specific T cells, namely IFN-γ and MIP-1β production. Overall, all study individuals (7/7) produced IFN-γ (as expected from the ELISPOT assay), as well as MIP-1β (Appendix E1).

There were 4/7 study individuals who had similar polyfunctionality between peptide variants, for example, ZA006 (Figure 4.7), where both peptide variants (C45 and A13) elicited similar proportions of CD8+ T-cells producing MIP-1β and IFN-γ or MIP-1β only (Figure 4.7B). On the other hand, 3/7 study individuals demonstrated differences in the proportions of cytokine combinations produced in response to peptide variants, as shown for ZA023 (Figure 4.8). Here, HIV-specific CD8+ T-cells recognizing the C45 peptide variant produced IL-2, which was absent for the A13 peptide variant. In addition, the C45 peptide variant elicited single cytokine producing cells for IFN-γ, TNF-α and IL-2. The only single cytokine producing cells in the A13 variant were MIP-1β-producing cells, which were not present in the C45 variant, where all MIP-1β-producing cells were dual cytokine producing cells with MIP-1β and IFN-γ (Figure 4.8B). Taken together, the data suggest that amino acid mismatches between peptides variants may lead to differences in the profile of cytokines produced by the specific T cell clones that recognise them.

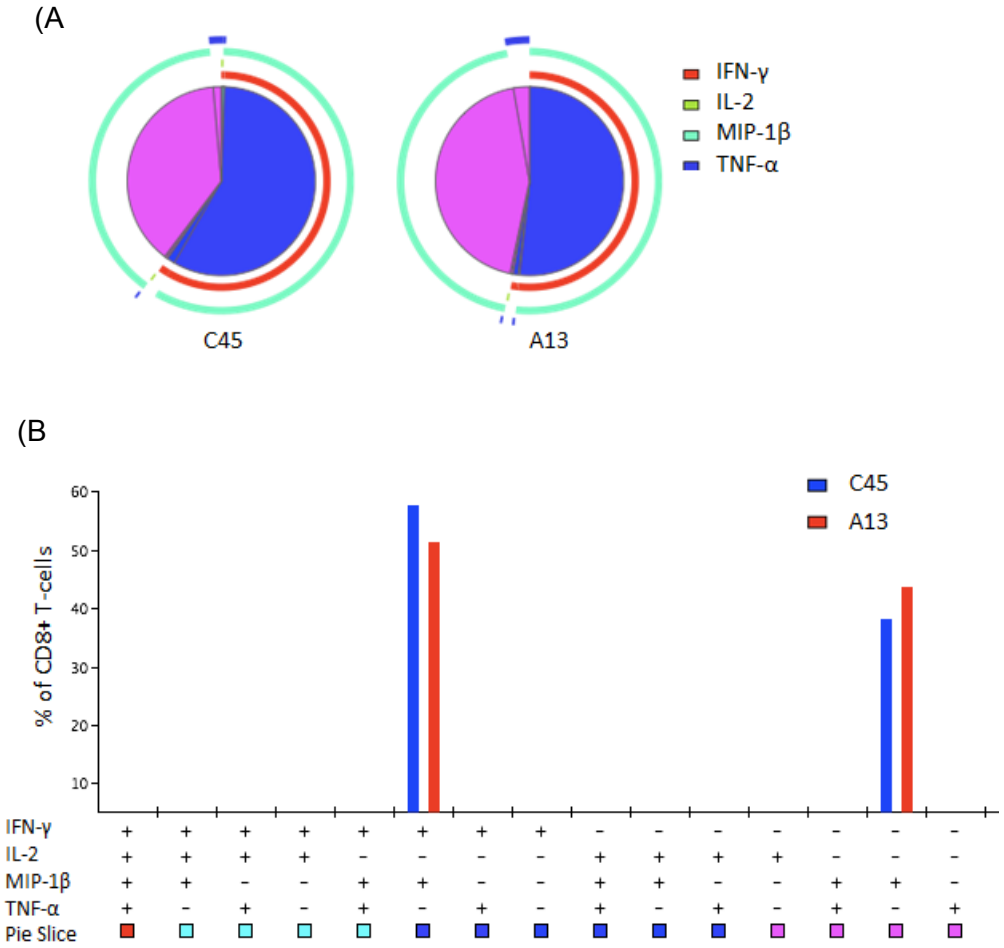


Figure 4.7. A representative figure of an individual showing similar polyfunctionality of HIV-1 specific T-cells recognizing HIV-1 peptide variants (ZA006). (A) Pie charts showing the proportion of HIV-specific T-cells producing specific cytokines. **(B)** Bar graph showing the proportion of CD8+ T-cells producing the different combinations of cytokines. Each arc shows the proportion of CD8+ T-cells producing the colour-coded cytokine as a fraction of total responding CD8+ T-cells.

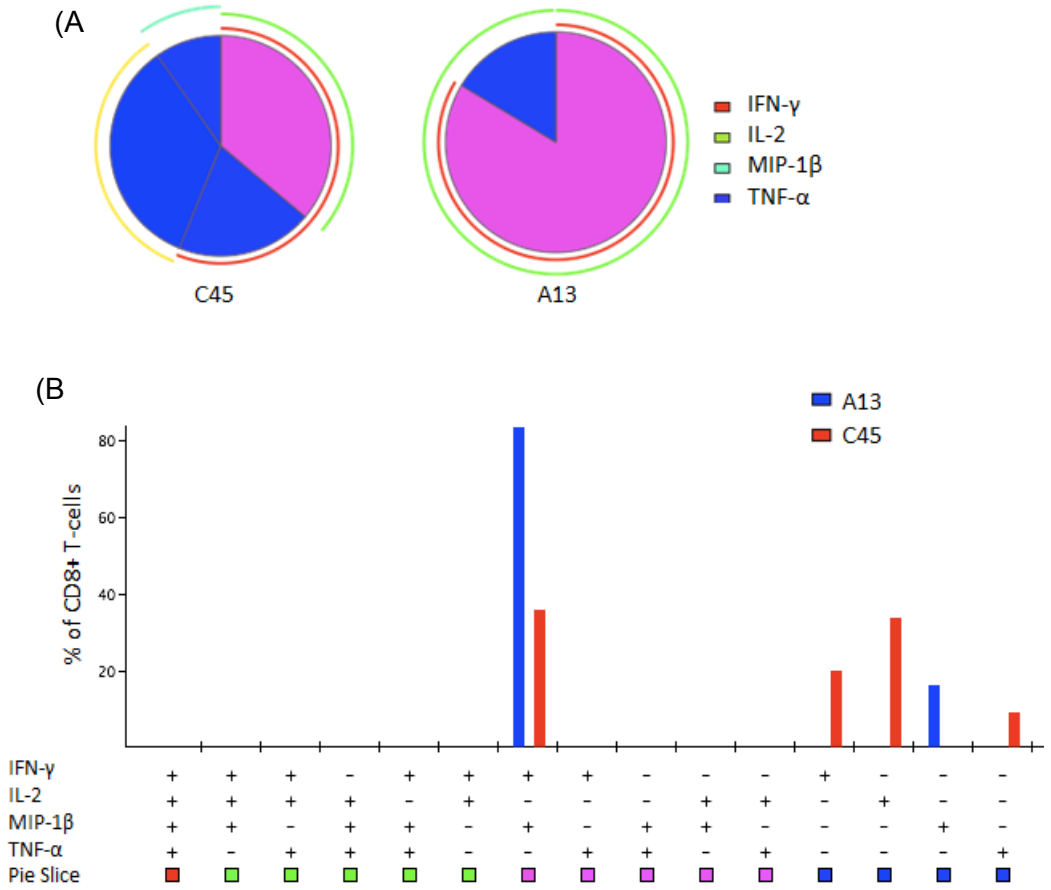


Figure 4.8. Representative figure of an individual (ZA023) showing differences in polyfunctionality between HIV-1 specific T-cells recognizing HIV peptide variants. (A) Pie charts showing the proportion of HIV-1specific T-cells producing specific cytokines. **(B)** Bar graph showing the proportion of CD8+ T-cells producing the different combinations of cytokines. Each arc shows the proportion of CD8+ T-cells producing the colour-coded cytokine as a fraction of total responding CD8+ T-cells.

4.2.4. Cytotoxic potential of cross-reactive HIV-specific T-cells

HIV-specific T-cells can perform other functions apart from cytokine production, including degranulation (CD107a expression), have specific cytotoxic potential (measured by perforin and/or granzyme B). Therefore, the effect of amino acid mismatches in peptide variants on the cytotoxic potential of T-cells recognizing these peptides was investigated, by measuring CD107a, IFN-γ, perforin and granzyme B by flow cytometry. To determine the proportions of HIV-specfic cells expressing perforin and granzyme, since there are large intracellular stores of these proteins, perforin and granzyme were gated together with IFN-γ and/or CD107a and layouts for each study individual are shown in Appendix E2, Figures E2.1-E2.7.

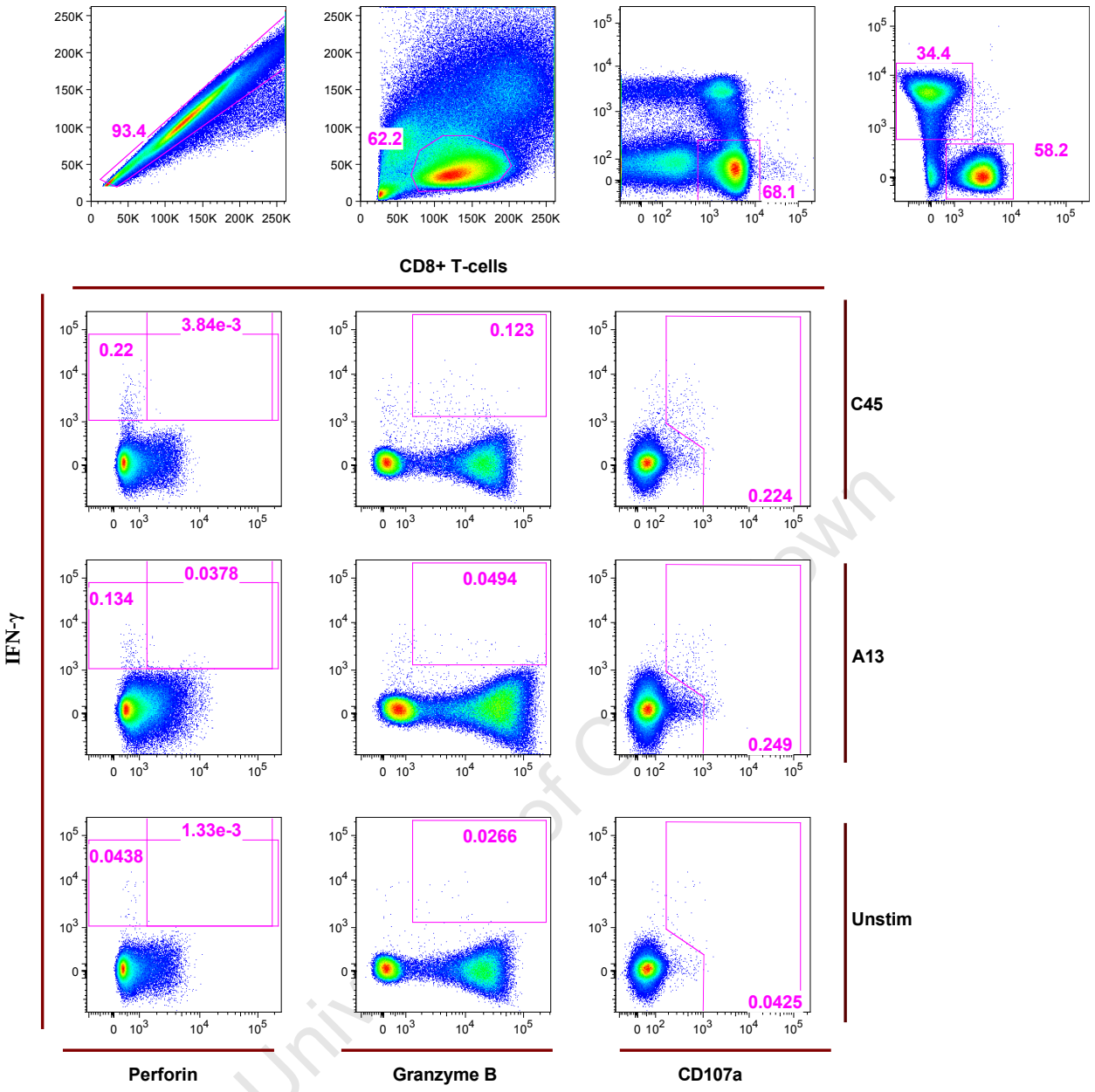


Figure 4.9. Cytotoxic potential of HIV-specific CD8+ T-cells stimulated with corresponding peptide variants with amino acid mismatches. A representative figure from study individual ZA023 showing the expression of IFN-γ, perforin, granzyme B and CD107a after 16-hour stimulation with two corresponding Gag peptide variants, A13 and C45. The upper row flow plots are showing the gating strategy. The bottom panel shows thew cytotoxic profile for each peptide variant. Granzyme B and Perforin were co-gated with IFN-γ to exclude non-antigen-specific cells with endogenous stores of these proteins. The last row of the bottom panel shows the unstimulated sample. The proportion of CD8+ T-cells producing each cytotoxic molecule is shown in each plot. The data is representative of two independent experiments.

Firstly, the cross-reactive cytotoxic potential of these peptide variants was investigated by determining the quality or profile of cytotoxic potential as the ability to degranulate or

express perforin or granzyme B in HIV-specific cells. Specific perforin or granzyme B were defined as the sum of that which is constitutively produced with IFN-γ and/or CD107a. A representative example of the gating strategy is shown for study individual ZA023 (Figure 4.9). In this study individual, both perforin and granzyme B were upregulated to at least one peptide variant tested. In addition, CD107a was expressed (Figure 4.9).

Figure 4.10 summarise the results of all seven individuals tested. An uneven pattern of the quality and quantity of CD107a production and perforin and granzyme B in HIV-sepcfic cells was observed. Peptide-specific CD8+ T cells in ZA028 and ZA033 express perforin to both rvariants demonstrating a similar quality of cytotoxic potential. .In ZA006, CD107a and granzyme B were both expressed at identical frequencies. However, in this individual, the level of perforin upregulation was different between peptide variants, with A13 upregulating perforin in 50% and C45 in 30% (1.7-fold) of CD107a and/ IFN-γ producing CD8+ T-cells. In ZA021 and ZA029, the cytotoxic potential was similar between the tested peptide variants. However, different frequencies of T cells producing or expressing the different molecules were observed. For example, in ZA021, the A19 variant detected more granzyme B positive cells than the C51 variant (1.8-fold more, Figure 4.10). Similary, in ZA029, A42 detected consistently higher frequencies of CD8+ T-cells producing CD107a and granzyme B than D42. On the other hand, ZA023 and ZA030 demonstrated differences in the quality of cytotoxic potential between peptide variants. In ZA023, when CD107a and granzyme B were both expressed, the proportions were higher in CD107a, 34% for A13 and 20% for C45 variant (1.8-fold), and comparable in granzyme B.

When combinations of cytotoxic molecules were assessed, polycytotoxic potential of HIV-specific CD8+ T-cells recognizing the different peptides variants differed in some individuals and not in the others. For example, in study individual ZA021, the quality of polycytotoxic potential was similar for the peptide variants A19 and C51 (Figure 4.11). On the other hand, ZA023 demonstrated differences in the quality of polycytotoxic potential of CD8+ T-cells recognizing peptide variants A13 and C45 (Figure 4.12). In the response detected to the one peptide variant, A13, perforin was up regulated, but this was not the case for the other peptide variant. Interestingly, in this individual, using the cytokine panel, it was the C45 peptide variant that detected T cells that could produce additional cytokines produced compared to the A13 variant (Figure 4.12).

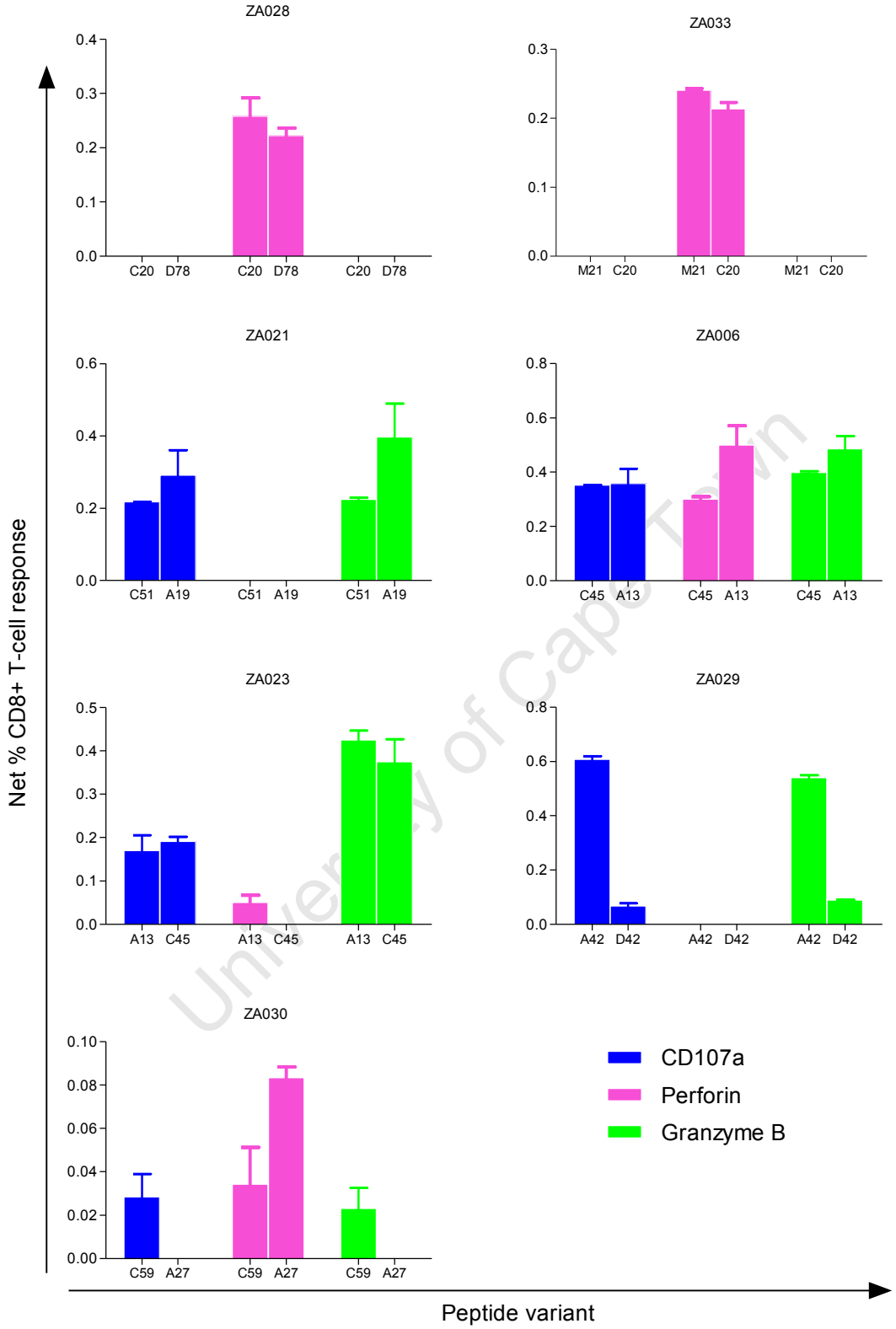


Figure 4.10. The profile of cytotoxicity potential of HIV-1 specific T-cells recognizing HIV peptide variants. Responses are corrected for background. Perforin and granzyme B responses refer to the upregulation of the molecules, that is, produced in cells also expressing IFN-γ and/or CD107a.

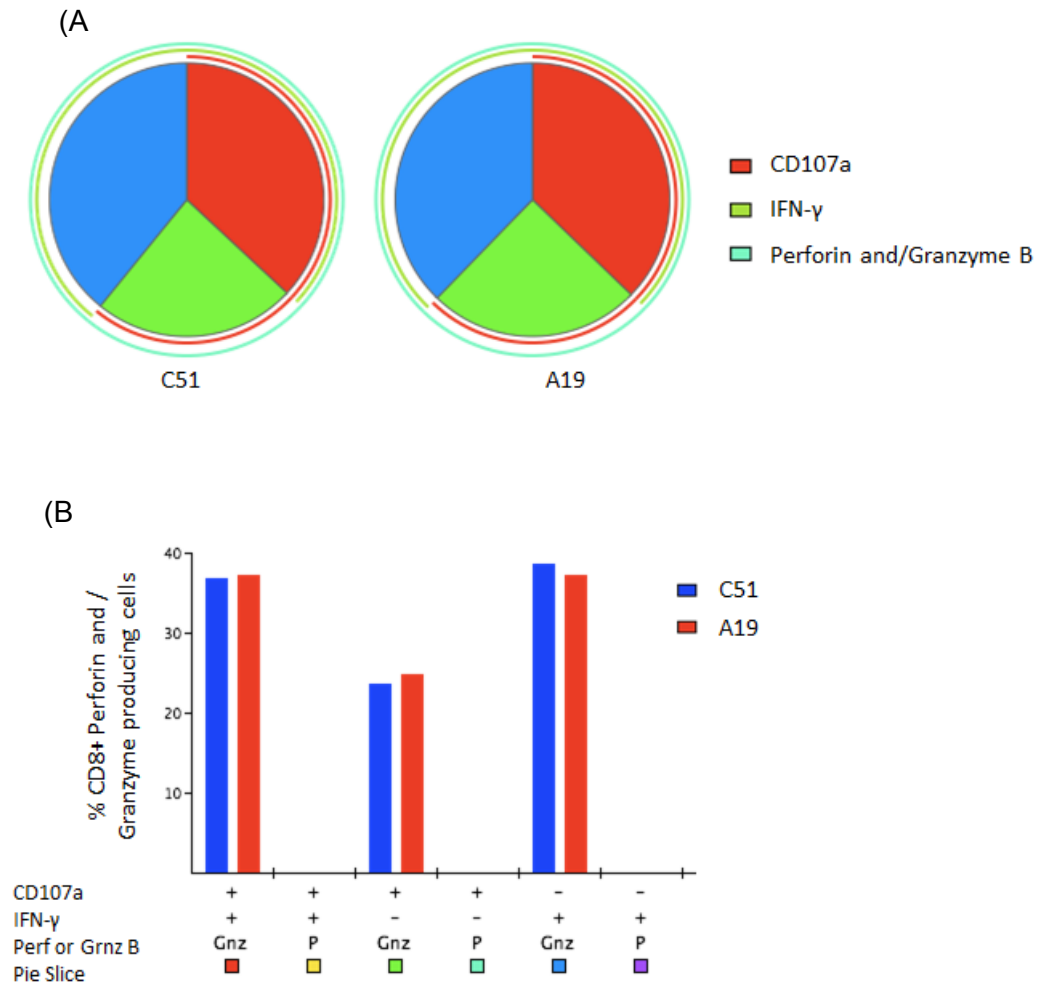


Figure 4.11. A representative example of an individual (ZA021) showing similarities in cytotoxic potential of HIV-1 specific T-cells recognizing HIV peptide variants. (A) Pie chart showing the proportion of HIV-1specific T-cells producing specific cytotoxic molecules. **(B)** Bar graph showing the proportion of CD8+ T-cells producing the different combinations of cytotoxic molecules. The pie arcs shows the proportion of perforin and/or granzyme B positive CD8+ T-cells producing IFN-γ and/or CD107a, that is CD8+ T-cells upregulating perforin and or granzyme B.

Overall, 2/7 study individuals (ZA023 and ZA030) demonstrated differences in the cytotoxic potential profile (Appendix E2). In these two study individuals, the peptide variant with greater avidity upregulated perforin, namely A13 (for ZA023) and C59 (for ZA030), which was not observed with the peptide with lower avidity (Appendix E2). On the other hand, 5/7 individuals demonstrated similarities in the profile or quality of cytotoxic potential (Appendix E2). Overall, 5/7 individuals had HIV-specific T-cell degranulation (as measured by CD107a expression) in response to stimulation with corresponding peptide variants

(Appendix E2). However, it was interesting to note that 2/7 individuals upregulated perforin but did not demonstrate degranulation (ZA028 and ZA033). Furthermore, not all HIV-specific T-cells upregulated both cytotoxic molecules upon stimulation. There were 5/7 individuals who up regulated perforin and an equal number of subjects who upregulated granzyme B. There were 3/7 individuals who upregulated both perforin and granzyme B together (Appendix E2).

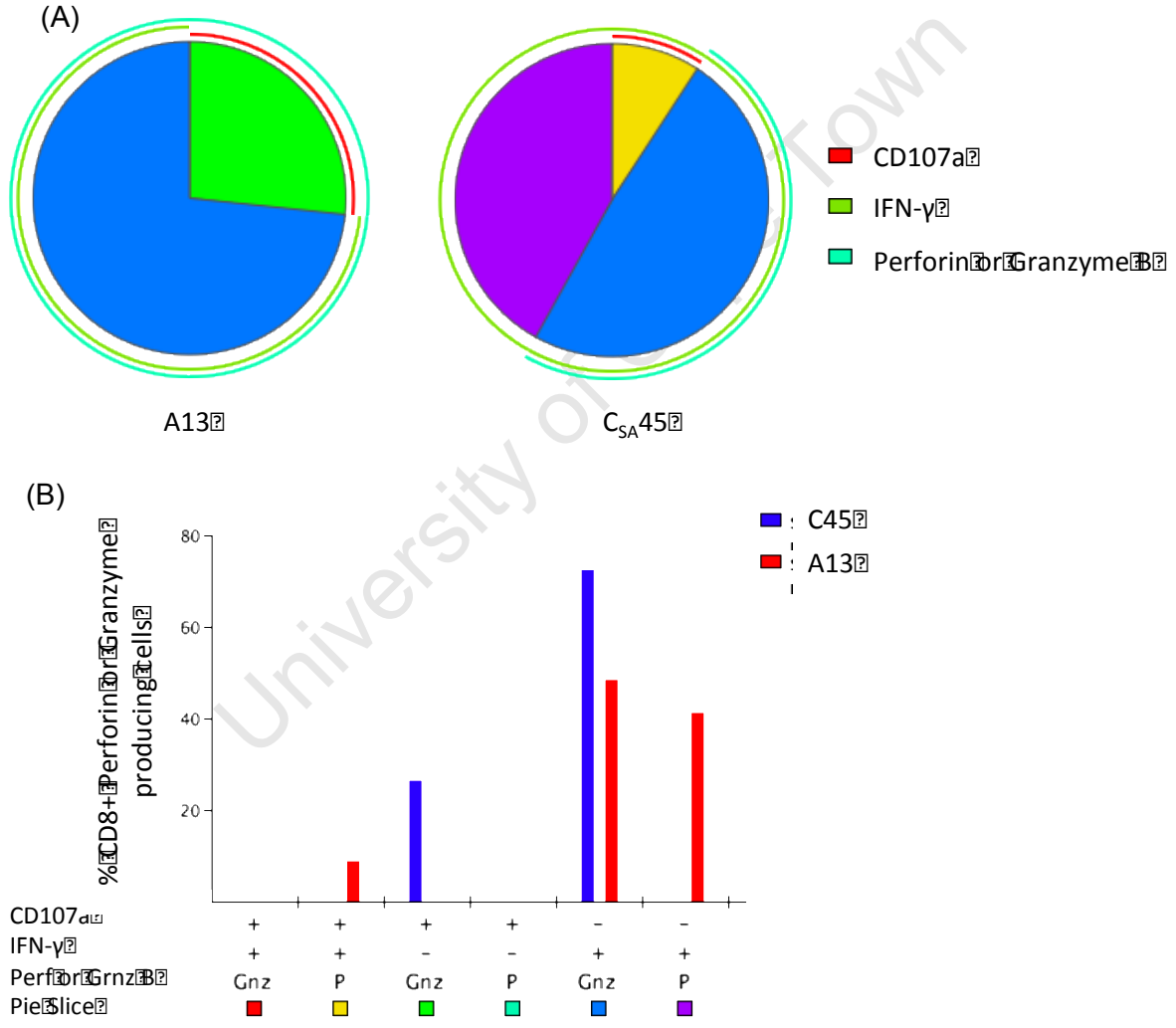


Figure 4.12. A representative example of an individual (ZA023) showing differences in cytotoxic potential of HIV-1 specific T-cells recognizing HIV peptide variants. (A) Pie chart showing the proportion of HIV-1specific T-cells producing specific cytotoxic molecules. **(B)** Bar graph showing the proportion of CD8+ T-cells producing the different combinations of cytotoxic molecules. The pie arcs shows the proportion of perforin and/or granzyme B positive CD8+ T-cells producing IFN-γ and/or CD107a, that is CD8+ T-cells upregulating perforin and or granzyme B.

4.2.5. Proliferative capacity of cross-reactive HIV-specific T-cells

The next level of characterization of the functional cross-reactive potential of HIV epitope variants was investigation of the proliferative capacity of HIV-specific T-cells recognizing corresponding peptide variants. This was performed using a 6 day Oregon Green proliferation assay in four study individuals due to limited sample availability. Figure 4.13 shows the gating strategy for detecting proliferating CD8+ T cells in response to the peptide variants tested, and the four individuals studied.

Two participants with reactive peptide variants with different functional avidity (ZA030 and ZA035) and two participants with reactive peptide variants with similar functional avidity (ZA021 and ZA028, less that 1.5 times difference in avidity between variants) were studied, as shown in Figure 4.13. 'Variant 1' indicates the peptide with the greater avidity for ZA030 and ZA035, which were both 10 fold greater than Variant 2. Net proliferative responses are shown in Fig 4.14A. Peptide variants demonstrated differences in the effects of amino acid mismatches on the proliferative capacity of HIV-specific T-cells recognizing them. Participants ZA030 and ZA035 demonstrated higher proliferative responses for the reactive peptides with greater functional avidity (C59 and M78). Interestingly, the A27 (variant 2) reactive peptide did not stimulate any proliferation in ZA30, indicating that it might be reactive with T cell clones that could not proliferate, despite being reactive in the ELISPOT assay and secreting multiple cytokines. Strikingly, for ZA028 and ZA021, despite the peptides having similar avidity, giving very similar ELISPOT responses and producing the same repertoire of cytokines, their proliferative capacity differed considerably (Figure 4.14A). The differences in proliferative capacity were not statistically significant, most likely due to the small numbers of individuals investigated (Figure 4.14B, $p = 0.0571$).

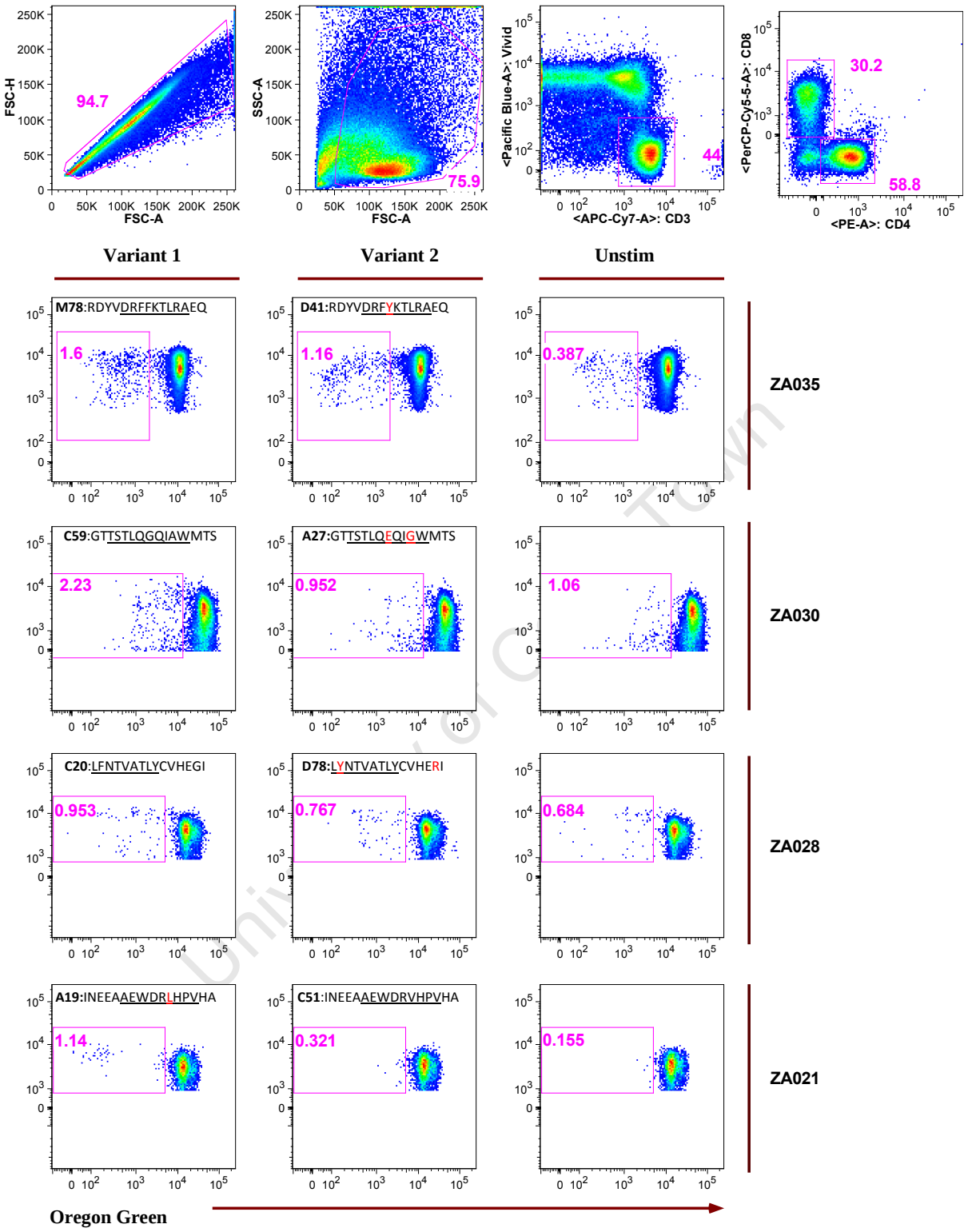


Figure 4.13. The proliferative capacity of HIV-specific T-cells in response to stimulation by corresponding peptide variants with amino acid mismatches. The top row shows the gating strategy and the remainder of the panels show results of the four study individuals tested, showing proliferation of CD8+ T cells in response to the peptide pair with variant amino acids. The peptides used for stimulation are indicated, and the right column shows the background proliferation in the wells in which no peptide stimulation was added (Unstim).

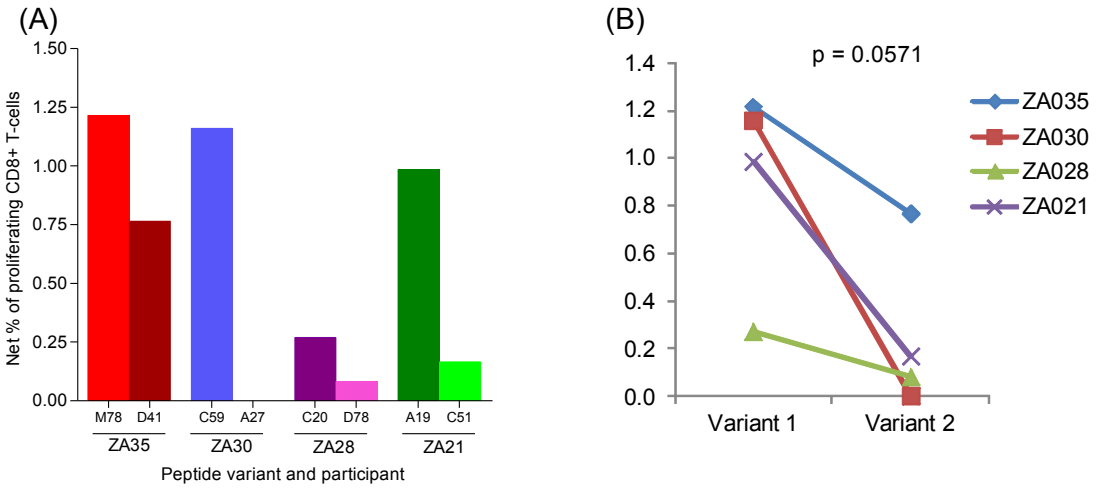


Figure 4.14. HIV-specific T-cell proliferative capacity in response to peptide variants with amino acid mismatches. (A) The net percentage of CD8+ T-cells proliferating in response to peptide variants in four study individuals. **(B)** Matched pairs comparison of the frequency of CD8+ T-cells proliferating in response to stimulation with peptide variants. PBMC were labelled with Oregon Green and stimulated with corresponding peptide variants for 6 days after which proliferation was analysed by flow cytometry. The p value was obtained using the non-parametric Wilcoxon matched T test.

Results from the four assays, namely functional avidity, cytokine production, cytotoxic potential and proliferative capacity, are combined into Table 4.2. It was noted that the majority (3/4) of individuals that demonstrated functional avidity differences that also had cytokine profile data available had differences in their cytokine profile in response to stimulation with peptide variants. In 2/3 of the individuals (ZA029 and ZA030), the peptide variant with higher avidity resulted in more cytokines being produced compared to the peptide variant with higher functional avidity, and this difference in TNF-α production for ZA029, and in the case of ZA030, CD107a and granzyme B production (Table 4.2). The third individual with cross-reactive peptides with vastly differential avidity (ZA023, 356-fold difference) had some functions of peptide-specific cells in common, namely IFN-γ, MIP-1β, CD107a, and granzyme B, whilst the higher avidity peptide additionally detected HIV-speific cells expressing perforin, and the low avidity peptide stimulated IL-2 and TNF-α production. There was one case where peptide variants that demonstrated functional avidity differences did not show differences in cytokine profile, in individual ZA033. On the other hand, three study individuals that demonstrated little or no functional avidity differences between peptide variants, also demonstrated no differences in the cytokine profile or cytotoxic potential of

CD8+ T-cell they detected (ZA006, ZA021 and ZA028). Interestingly, for the four individuals that had CD8+ proliferation data, 2/4 individuals (ZA028 and ZA021) that demonstrated no differences in functional avidities, cytokine and cytotoxic profiles had marked differences in proliferative capacity of HIV-specific T-cells recognizing the peptide variants (3-fold and 6-fold, respectively). For the remaining two individuals (ZA030 and ZA035), differences in functional avidity translated into differences in proliferative capaicy, where the peptide with the higher avidity stimulated more cells to proliferate.

Together, the data suggest that using one immunological readout, such as the IFN-γ ELISPOT assay, may not reveal functional differences exhibited by HIV-specific T-cells recognizing epitope variants, and in the context of a vaccine, amino acid mismatches between peptide variants may result in differences in functional avidity and/or discordant effects on the functional profile of HIV-specific T-cells they elicit.

Seventy-five per cent (3/4) pairs of peptides recognized by HIV-specific T-cells with similar functional profiles had conservative amino acid substitution between their corresponding variants. In one case (ZA006), a similar profile was observed but with non-conservative amino acid substitutions between the peptide variants (Table 4.2). On the other hand, 67% (2/3) peptide pairs recognized by HIV-specific T-cells exhibiting different functional profiles between the variants had non-conservative amino acid substitutions. Only one case (1/3) was a different functional profile observed but with a conservative amino acid substitution between the peptide variants (Table 4.2). Overall, there is a trend towards different functional profiles of HIV-specific T-cells recognizing HIV epitope variants with non-conservative amino acid substitutions, and a trend towards similarities in functional profile if the peptide pairs differed in conservative amino acid substitutions or closely related side chains. This suggests that the different HLA alleles and TCR contact sites in different individuals restricting these epitopes have differing degrees of tolerating amino acid changes resulting in discordant effects of these changes on functional profiles of HIV-specific T-cells recognizing them.

Table 4.2. Summary of functional avidity, cytokine profile, cytotoxicity potential and proliferative capacity of HIV-1specific T-cells

PID	Peptide number	Peptide sequence	Amino acid change	Functional avidity (µg/ml)	Cytokine profile				Cytotoxicity potential			Proliferative capacity (%)
					IFN-γ	IL-2	MIP-1β	TNF-α	CD107a	Perforin	Granzyme B	
ZA06	A13 C45	PQDLNMMLNIVGGHQA PQDLNTMLNTVGGHQA	Non-conservative	0.398 0.224								nd
ZA30	C59 A27	GTTSTLQGQIAWMTS GTTSTLQEQIAWMTS	Non-conservative	0.016 0.158								1.160 0.000
ZA23	C45 A13	PQDLNTMLNTVGGHQA PQDLNMMLNIVGGHQA	Non-conservative	1.259 0.004								nd
ZA21	C51 A19	INEEAAEWDRVHPVHA INEEAAEWDRLHPVHA	Conservative	0.004 0.003								0.985 0.166
ZA28	C20 D78	LFNTVATLYCVHEGI LYNTVATLYCVHERI	Conservative	4.467 3.162								0.270 0.083
ZA33	M21 C20	LYNTVATLYCVHQRI LFNTVATLYCVHEGI	Conservative	0.014 0.316								nd
ZA29	A42 D42	DRFFKTLRAEQATQ DRFYKTLRAEQASQD	Conservative	0.447 0.141								nd
ZA11	C06 A06	RPGGKKHYMLKHIVW RPGGKKKYRLKHLVW	Conservative	0.102 0.199	Responses not detected by flow cytometry							nd
ZA35	M78 D41	RDYVDRFFKTLRAEQ RDYVDRFYKTLRAEQ	Conservative	0.126 1.259	nd							1.213 0.773

The underlined region denotes the pre-described epitope. Amino acid mismatches are denoted in red. Conservative and non-conservative classification of amino acid changes means amino acid substitution by an amino acid with a related or un-related side chain, respectively. Functional avidity was defined as the peptide concentration giving half maximum T-cell response by the IFN-γ ELISpot assay. Results are based on CD8+ T-cell responses detected. Where positive CD8+ T-cell responses were detected, the different functions are color-coded. Blank cells represent where the function was not detected as positive for that particular peptide variant. Nd denotes not done. For ZA11, no responses were detectable by flow cytometry and for ZA35, cytokine and cytotoxic production experiments were not performed. Perforin and granzyme B are based on upregulation of these two cytotoxic molecules in combination with CD107a and/or IFN-γ. Bolded functional avidities were considered different (≥3-fold difference).

4.3. DISCUSSION

The importance of HIV-specific T-cells in the control of virus replication during natural HIV-1 infection is well documented (Borrow *et al*., 1994; Koup *et al*., 1994). Yet there is a need to better understand the HIV-specific T-cell functional correlates of cell-mediated control of HIV. Moreover, the role played by the huge genetic diversity of the virus requires deeper understanding, as these immune correlates of cell-mediated control of HIV are unravelled. One way of doing this is to functionally characterize HIV-specific T-cells with the ability to cross-recognize epitope variants from the virus.

There are previous studies that have assessed the ability of HIV-specific T-cells to cross-recognize different HIV-1 clades both in regions of single (Geldmacher *et al*., 2007; McKinnon *et al*., 2005) and multiple circulating clade epidemics (Gudmundsdotter *et al*., 2008; Gupta *et al*., 2006). These studies reached a consensus that HIV-specific T-cells have the ability to cross-recognize HIV variants or clades but there is preference for the recognition of peptides matching the infecting clade. Furthermore, different HIV epitopes demonstrated different degrees of cross-reactivity (Yu *et al*., 2005). However, these studies relied on the use of a single immunological readout, namely the production of IFN-γ in an ELISpot assay, to measure 'cross-reactivity'. A characteristic feature of this assay is that it relies on the use of exogenously loaded HIV peptides that are in excess of physiologically relevant concentrations of these peptides *in vivo*. Therefore, whether the T-cells identified as cross-reactive in these studies are truly cross-reactive *in vivo* remains elusive. Additionally, whether the cross-reactive responses exhibit antiviral suppressive activity still needs further investigation. In one previous study, cross-clade activity of HIV-specific CD8+ T-cells identified in an ELISpot assay was compared to results from an *in vitro* viral replication assay (Bennett *et al*., 2008). CD8+ T-cells that demonstrated substantial cross-reactivity in the peptide-based ELISpot assay measuring IFN-γ production only had antiviral activity to the autologous clade B, and not other clades tested, in the viral inhibition assay. Similar finding were observed in SIV-infected macaques in which a significant impairment of CD8+ T-cells to kill infected cells was observed, despite broad cross-clade responses by the exogenous peptide based IFN-γ ELISpot assay (Valentine *et al*., 2008). Another recent study demonstrated that vaccine induced CD8+ T-cells equally recognize wild type and variant SIV epitopes using tetramer based assays (Hulot *et al*., 2011). However, functional studies

demonstrated that these T-cells elicit suboptimal cytokine production on recognition of variant epitopes compared to wild type epitopes, suggesting limited cross-reactivity when other T-cell functions are taken into consideration.

This study was undertaken in an attempt to investigate further the cross-reactivity of cross-reactive epitopes identified in the studies in this thesis. The first stage of this investigation involved testing whether corresponding peptide variants with amino acid mismatches that were cross-recognized at high concentrations of exogenously loaded peptides, were still cross-recognized at low or physiologically relevant concentrations of the peptide variants, using the IFN-γ ELISpot peptide dilution assay. This allowed us to test whether amino acid mismatches between cross-recognized corresponding peptide variants affected the functional avidity of the T-cells recognizing them.

Functional avidity is a measure of T-cell responsiveness to peptide (Bennett *et al*., 2007). Therefore a peptide variant with higher avidity will be one that requires less peptide to elicit a response equal in magnitude to its corresponding variant or counterpart. In this study, the results of the relationship between magnitude of IFN-γ response and functional avidity could be categorized into those HIV-specific T-cells that differed in magnitude and functional avidities, those that demonstrated similar functional avidities despite observed differences in magnitude of IFN-γ T-cell responses, and those that had similar functional avidity and similar magnitude of responses. The differences observed in some peptide pairs in their magnitude and functional avidity could be as a result of differential strength of binding to TCR or their restricting HLA alleles. Interestingly, further characterization of the nature of amino acid mismatches showed that the majority of peptide pairs in this category had amino acid mismatches between distantly related amino acids (non-conservative amino acid substitutions), implying that in natural HIV infection, these amino acids may be involved in either peptide processing, binding to restricting HLA or TCR contact sites during the recognition of these peptides by HIV-specific T-cells. This result is in line with data from previous studies suggesting that even single amino acid mutations between peptide variants was sufficient to alter their functional avidity, illustrating the importance of variation in epitopes on T-cell recognition (Bennett *et al*., 2007).

The fact that the nature of amino acid substitution might be having an effect on other key players involved in the presentation of these peptides to T-cells was further supported by the result that those peptide variants with conservative substitutions between variants had a similar magnitude of IFN-γ T-cell responses and similar functional avidities. Non-conservative substitutions that did not result in differences in magnitude or functional avidities among different individuals suggest that HLA alleles or TCR contact sites in these individuals have different degrees of sensitivity to amino acid changes, or that these changes did not affect either. However, the effect of other factors such as stability of these peptides in the cytosol and physiological factors cannot be neglected (Lazaro *et al*., 2011). Therefore, the recognition of both peptide pairs at lower peptide concentrations suggests that cross-reactivity observed at high peptide concentration in a standard ELISpot assay measuring IFN-γ production may also be identified *in vivo*. The loss of recognition of one variant and not the other for other peptide pairs suggest that for some epitopes, cross-reactivity is lost at physiological concentrations of these peptides and possibly *in vivo*.

The functional cross-reactive potential of these HIV peptide variants was investigated by characterizing cytokine profiles measuring IFN-γ, IL-2, MIP-1β and TNF-α, cytotoxic potential measuring upregulation of perforin and granzyme B and the degranulation marker CD107a, as well as proliferative capacity of HIV-specific T-cells recognizing these peptide variants using flow cytometry. This allowed us to test the hypothesis that amino acid mismatches and avidity differences between corresponding peptide variants would result in differences in the functional profile of CD8+ T-cells recognizing them.

Highly discordant effects of amino acid substitutions on these measured parameters were observed, and few patterns emerged, with no clear relationship between functional avidity and multiple cytokines secreted or a greater cytotoxic or proliferative potential. Relatively few peptide pairs tested limited our study, and a larger study may reveal consistent differences. The identification of a subset of CD8+ T-cells degranulating but without producing IFN-γ suggest the need to include more parameters in the measurement of HIV-specific T-cells, as some responses are missed by the use of one marker such as IFN-γ. Furthermore, the data suggest that there are qualitative differences observed in the CD8+ T-cells recognizing HIV-1 epitope variants when more immunological readouts are used that

may not be detected by using a limited number of parameters, as has been suggested previously (Richmond *et al*., 2011).

Antigen sensitivity or functional avidity of HIV-specific T-cells has been shown to influence the functional profile of T-cells (Almeida *et al*., 2009). Polyfunctionality of HIV-specific T-cells decreased with decreasing peptide concentration and decreasing functional avidity of the T-cells recognizing the peptides tested. In the study presented in this chapter, although polyfunctionality was assessed at only a single peptide concentration due to sample availability, and in limited number of individuals certain peptides that had similar functional avidities at that particular concentration still demonstrated differences in their profile of cytokine production and cytotoxic potential, future studies with a larger dataset would go beyond descriptive value and allow for a more rigorous analysis of avidity and its relationship with T cell function. Furthermore, some peptide pairs that had differences in functional avidities also had differences in their profile of cytokines, cytotoxic potential and proliferative capacity. This is in line with previous findings that demonstrated that cross-reactivity of T-cells differs between overnight IFN-γ assays and longer proliferation assays (McKinnon *et al*., 2007). These results again highlight of the fact that HIV-specific T-cells recognizing epitope variants cannot be fully understood by using a single technique or by measuring a single function. In other words, avidity predicts function but only in a subset of individuals. Although this was not a correlates study, there are previous studies that have highlighted the importance of other functions of HIV-specific T-cells (Betts *et al*., 2006; De Rosa *et al*., 2004; Hesperger *et al*., 2010; Roederer *et al*., 2004), hence the necessity to characterize them to fully understand whether particular epitope variants might still mount effective responses in the context of vaccines.

It is important to note that this study made the assumption that cytotoxic potential may be a surrogate for potential viral suppression *in vivo*, which in fact can only be shown using viral suppression assays. Although in a previous study (Bennett *et al*., 2008) it was only IFN-γ production by ELISpot assay that did not translate to actual viral suppression, it is still possible that it might be the case with other functions that were measured in this study. In addition, the TCR repertoire of responding cells was not characterized in this study to determine whether there were differences in T-cell clonotypes and TCR usage of the responses that were detected for different peptide variants as many T-cells have public

clonotypes. Therefore, additional studies looking at TCR usage and viral replication inhibition assays (Saez-Cirion *et al*., 2007) using CD8+ T-cells specific for the peptide variants studied and larger datasets may be important in drawing conclusions about the effect of amino acid mismatches on the true cross-reactive nature of HIV-specific T-cells and drawing clear-cut overall correlation analysis in these datasets. Also, the differences in assay sensitivity among various assay methods may affect the classification of responses being discordant or not. The fact that continual antigen exposure has an effect on the expression of perforin, proliferative capacity and polyfunctionality makes it difficult to make conclusive interpretation of this data. Because of this, it is possible that VL, CD4 counts and period of infection might have affected the predicted cross-reactivity. One way to overcome this could be to perform longitudinal studies in which study individuals are followed up from acute to chronic infection and their polyfunctional cross-reactive potential assessed and analysed together with clinical data and period of infection to predict the best correlate of cross-reactivity. However, there is no evidence as to whether increased cross-reactivity prevents or delys disease progression. Further studies will be required that characterize cross-reactivity in controllers and progressors.

CHAPTER 5

DISCUSSION

The development of a globally effective HIV vaccine is a major health priority. Developing a vaccine that can elicit responses that will protect against the diversity of HIV-1 strains responsible for the HIV-1 pandemic remains a major challenge in the field. It is generally accepted that an effective vaccine will need to elicit both broadly cross-neutralizing antibody responses, and well as effective CD8+ T cell responses. While the recently discovered broadly cross-neutralizing antibodies are effective against approximately 80% (Pejchal et al., 2010; Walker et al., 2010) of viruses from different clades, there is some evidence in natural infection of clade specificity (Li *et al*., 2011; Pejchal *et al*., 2011; Seaman *et al*., 2010).

There is a wealth of evidence to support the role of CD8+ T cells in controlling HIV replication. Studies have shown that the emergence of virus-specific CD8+ T-cells during primary HIV infection coincides with initial drop in peak viraemia (Borrow *et al*., 1994; Goonetilleke *et al*., 2009; Koup *et al*., 1994). Furthermore, studies in macaques have shown CD8+ T-cell depletion results in an increase in viraemia in SIV infected macaques (Friedrich *et al*., 2007; Jin *et al*., 1999; Lifson *et al*., 2001; Schmitz *et al*., 1999). Strong evidence of the role of CD8+ T cells was provided by genome-wide association studies that demonstrated that polymorphisms in the HLA class I gene loci were associated with differential effects on disease outcome (Fellay *et al*., 2007; Pereyra *et al*., 2010). Therefore, there has been a major focus on HIV vaccine immunogen designed to elicit cellular immunity.

HIV exhibits an enormous genetic variability and challenges faced due to the huge genetic diversity of the virus cannot be neglected. Although there have been some recent vaccine trials testing the mosaic immunogens which are designed to minimize the distance between vaccine immunogen and the circulating viruses (phase 1 clinical trial IAVI B003/IPCAVD-004 by Barouch *et al* began in Boston, USA, http://www.iavi.org/archives/2010), vaccines going forward into Phase III trials in South Africa still contain wild type virus sequences (Glenda Gray; AIDS Vaccine 2011, Thialand; AIDS Vaccine Clinical Trials Update). Thus, evaluating the potential effectiveness of these types of vaccine immunogens against the global strains remains relevant to understand whether they will induce virus-specific CD8+ T-cells which can recognize diverse strains, and whether these cross-reactive responses will

have potent antiviral activity. Evaluation of cross-clade efficacy can only be directly addressed by performing human clinical HIV vaccine trials in different populations with different HLA genetic backgrounds, and where people are likely to encounter different HIV-1 clades. However this process is very costly and laborious and insight into these issues may be obtained by assessing the ability of cells from individuals infected with a particular HIV-1 clade to recognize peptides using laboratory based assays. It is most relevant to test this using sequences representing vaccines that are likely to be tested in South Africa.

The goal of the first study described in this thesis was to investigate inter- and intra-clade cross-reactivity of HIV-specific CD8+ T-cells using peptides based on current candidate HIV vaccine inserts. This study provides more relevant data than previous studies where cross-reactivity of HIV-specific T-cells relied on the use of peptides that were not based on candidate vaccine immunogen sequences. Our data corroborated previous findings that HIV-infected individuals can recognize HIV peptide variants based on other clades not matching the ones infecting them (Geels *et al*., 2005; Geldmacher *et al*., 2007; Gudmundsdotter *et al*., 2008; Gupta *et al*., 2006; McKinnon *et al*., 2005), and also provides novel evidence of responses in South African subtype C-infected individuals that both exclusively, and commonly, target clades A, B, C (China), and D peptides. The substantial cross-reactivity obtained with clades A, B and D suggests that vaccines based on these clades would provide some protection in a clade C epidemic. However, we found preferential recognition of clade C peptide reagents suggesting that there is an advantage of matching a vaccine to the infecting clade. Furthermore, there was no increased recognition of South African clade C peptides compared to Chinese clade C peptides suggesting no further advantage of matching vaccines to viruses circulating in the region.

We found that cross-reactive peptides were commonly located in highly conserved, functionally important regions, particularly the p24 region of the Gag protein. Studies have suggested that there are segments in p24 that are highly targeted by virus-specific CD8+ T-cells which the virus is not able to mutate and therefore evade immune pressure int hese regions due to functional constraints (Dahirel *et al*., 2011). These highly targeted regions that show extensive cross reactivity would serve as important regions to include in a vaccine immunogen.

An important consideration for vaccine development is not only whether regions included in vaccines will be cross-reactive to a range of clades globally, but also whether the immune responses they generate will provide protection against HIV. Further insights into CD8+ T-cell responses that may be beneficiary in slowing disesase progression and may be subsequently lower HIV transmission rates can be predicted by calculating a protective ratio (PR; Mothe *et al*. 2011). In this study, the PRs served as a surrogate for the ability of each peptide to elicit protective T-cell responses and were calculated by dividing median viral loads of responders by that of non-responders. PRs were classified as beneficial if the PR >1. We compared all beneficial Gag peptides identified by Mothe *et al*. (2011) from a clade B cohort with peptides reactive in the study reported in this thesis, and found that all of them (10/10) were reactive in the individuals studied in this thesis, with 3/10 (30%) being classified as immunodominant peptides using the near to consensus clade C_{Du422} vaccine insert. In addition, of the 18 beneficial peptides that Mothe *et al*. (2011) identified in the clade C infected individuals, 17/18 (94 %) were recognized in our study, with 7/18 (39%) being identified as immunodominant. Furthermore, these beneficial peptides showed similar clustering patterns to the cross-reactive peptides we identified (Figure D4.1A and B, Appendix D4), further supporting the protective potential of these T-cell responses.

After establishing that HIV-1 subtype C infected individuals can mount cross-reactive HIV-specific CD8+ T-cell responses to the four major circulating HIV clades, the reactivity of centralized sequences based on consensus group M Gag and Nef peptides was assessed. As centralized sequences have been suggested as potential future candidate vaccine immunogens for a globally relevant vaccine, as well as their use as reagents to assess vaccine induced responses in individuals infected by different HIV clades, it was important to assess their performance in a mono-clade epidemic of subtype C viruses. Although the overall goal of this part of the study was to compare the reactivity of these reagents to that obtained in regions where multiple clades circulate such as Cameroon and Uganda, for the purposes of this thesis, the reactivity of these peptides was compared to those obtained for the four clades described above. We found that Group M peptides detected fewer responses compared to peptides matched to the infecting clade, and were comparable to clade-mismatched peptide. Poorer recognition of centralized sequences compared to clade-matched peptides suggests that they are not an ideal reagent for assessing HIV-specific T-cell responses in a mono-clade epidemic. Whether this is the same case with multi-clade epidemic will require further

analysis of combined data. However, data from the Ugandan study where clade A, D and recombinant viruses predominate, suggest that responses to these peptides reagents are focused towards a few immunodominant epitopes (Serwanga *et al*., 2011).

Previous cross-clade studies, including the ones reported in this thesis, have relied on using exogenously loaded peptides at high concentrations (possibly exceeding physiological levels) to measure the ability of HIV-specific T-cells to recognize them. Although this approach will identify CD8+ T-cells recognizing the peptide variants, whether these are detected at physiological concentrations of the antigens in vivo remains to be answered. Furthermore, the IFN-γ ELISPOT assay, while showing that CD8+ T-cells can recognize the peptides, does not prove that these T-cells have antiviral activity *in vivo.* One way to predict the efficiency of the interaction between CD8+ T-cells and virus infected target cells is by measuring the functional avidity of the T-cells (Almeida et al., 2009). Here we report on seventeen HLA class I-restricted peptide pairs that were assessed for functional avidity using a peptide dilution assay. This study allowed us to test whether amino acid mismatches between corresponding peptide variants resulted in differences in functional avidities. Interestingly, most peptides tested (23/34) were still recognizable at low peptides concentrations, with the recognition of other variants (11/34) being lost at physiologically relevant concentrations. Therefore, amino acid mismatches may have discordant effects on the 'true' cross-reactivity of peptides variants. Although it is known that in general, CD8+ T-cell populations exhibit several clonotypes that are specific for the same antigenic complex (Wilson *et al*., 1998), a recent study showed no preference in use of a particular clonotype in recognizing KK10, an HLA-B*27-restricted epitope from HIV Gag (Almeida *et al*., 2007). We did not perform molecular analysis of TCRB gene usage, to assess if use of certain CDR3 motifs resulted in higher functional avidity in one of the tested peptide variants or amino acid mismatches resulted in differential modes of binding to the same motif.

HIV-specific T-cells able to mediate multiple functions ('polyfunctional' cells) have been found at higher frequencies in individuals who control HIV (Betts *et al*., 2006). Similarly, polyfunctional HIV-specific CD8+ T-cells have been shown to demonstrate a more potent HIV suppressive activity than those that exhibit less polyfunctionality (Saez-Cirion *et al*., 2007). We investigated whether amino acid mismatches in cross-reactive peptide pairs identified in our studies would result in differences in the functional profile of CD8+ T-cells

recognizing them, beyond IFN-γ secretion identified in the ELISpot assay. We found that amino acid mismatches between peptides had discordant effects on the profile of cytokines and chemokines produced their cytotoxic potential as well as proliferative capacity. Peptide variants that had similar functional avidity also demonstrated similar functional profiles, while some that had differences in functional avidity exhibited both different and similar functional profiles. These results suggest that the development of 'mosaic' immunogens that include multiple variants of single epitopes to elicit not only broad coverage of HIV-1, but also broad functional responses is a good approach, as proposed by others (Malhotra *et al*., 2009). These results also emphasise that we need to define 'cross-reactive' responses in a more detailed, qualitative manner, rather than just quantity of IFN-γ responses.

There were several limitations to our study and to similar studies in general that could be addressed in future studies. Although the infecting viruses may represent future viruses the individuals may encounter, testing vaccine peptides in HIV infected invidivuals is quite different from using the same vaccine sequences to induce vaccine responses in HIV uninfected individuals who will encounter HIV. For example it is not know whether cross-reactivity will be affected by level of CD4 count or presence or absence of CD4 help, immune activation and viral load. In addition, in an HIV infected person, continual antigen exposure has a huge impact on perforin expression, proliferative capacity as T-cell approach exhaustion and polyfunctionality. The advantage of vaccine iduced responses as compared to testing vaccine antigens in people already infected is that vaccine induced T-cells precede HIV exposure. Furthermore, functional studies were performed at a single peptide concentration, and therefore the effect of antigen concentration on the profile of functions could not be characterized. In addition, IFN-γ ELISPOT readout was used as the criteria for reactive peptides. It is possible that some peptides might not produce IFN-γ, but other cytokines or chemokines. Initial screening of reactive peptides using other readouts, such as a perforin ELISPOT or high throughput proliferation assays, could be performed in future studies. Also, clonal analysis of HIV-specific CD8+ T-cells was not performed, since it was possible that clones with different TCRB gene usage resulted in differences in functions between peptide variants. We did not measure viral suppression, that is, measuring the ultimate function of CD8+ T-cells to eliminate virus-infected cells using a viral suppression assay (Saez-Cirion *et al*., 2010), which may provide more insight into whether the cross-reactivity demonstrated by the ELISpot assay and additional functional aspects of these cells

translated to antiviral cross-reactivity. However, a limitation of these types of studies performed in HIV-infected humans is the difficult of knowing whether a 'good' functional profile is the cause or consequence of viral control. These types of studies need to be performed in vaccine trials of candidate HIV vaccines

In conclusion, in a monoclade subtype C epidemic, we found that HIV-infected individuals mounted robust inter- and intra-clade cross-reactive HIV-specific T-cell responses, suggesting that vaccines based on these clades may work equally well. However, preferential of recognition of clade C-based peptide reagents suggest some advantage of matching a vaccine to the infecting clade. Moreover, despite the central nature of consensus group M Gag and Nef peptide reagents, their use to evaluate HIV-specific T-cell responses or as immunogens may be limited due to mismatches with particular clades. Overall, the cross-reactive qualitative and quantitative aspects of the HIV-specific T-cell response is a complex phenomenon to de-convolute that is related to the functional avidity of the CD8+ T-cell specific for a particular HLA class I-restricted epitope, and is affected by amino acid mismatches between peptide variants. These data may advance the development of an effective cross-reactive HIV vaccine immunogen, an urgent and critical need both in South Africa and globally.

REFERENCES

1. **Adamson, C. S., and E. O. Freed.** 2007. Human immunodeficiency virus type 1 assembly, release, and maturation. Adv Pharmacol **55:**347-87.

2. **Addo, M. M., X. G. Yu, A. Rathod, D. Cohen, R. L. Eldridge, D. Strick, M. N. Johnston, C. Corcoran, A. G. Wurcel, C. A. Fitzpatrick, M. E. Feeney, W. R. Rodriguez, N. Basgoz, R. Draenert, D. R. Stone, C. Brander, P. J. Goulder, E. S. Rosenberg, M. Altfeld, and B. D. Walker.** 2003. Comprehensive epitope analysis of human immunodeficiency virus type 1 (HIV-1)-specific T-cell responses directed against the entire expressed HIV-1 genome demonstrate broadly directed responses, but no correlation to viral load. J Virol **77:**2081-92.

3. **Aidoo, M., S. Sawadogo, E. C. Bile, C. Yang, J. N. Nkengasong, and J. M. McNicholl.** 2008. Viral, HLA and T cell elements in cross-reactive immune responses to HIV-1 subtype A, CRF01_AE and CRF02_AG vaccine sequence in Ivorian blood donors. Vaccine **26:**4830-9.

4. **Alexander, L., P. O. Illyinskii, S. M. Lang, R. E. Means, J. Lifson, K. Mansfield, and R. C. Desrosiers.** 2003. Determinants of increased replicative capacity of serially passaged simian immunodeficiency virus with nef deleted in rhesus monkeys. J Virol **77:**6823-35.

5. **Alexander-Miller, M. A., G. R. Leggatt, and J. A. Berzofsky.** 1996. Selective expansion of high- or low-avidity cytotoxic T lymphocytes and efficacy for adoptive immunotherapy. Proc Natl Acad Sci U S A **93:**4102-7.

6. **Allen, T. M., M. Altfeld, S. C. Geer, E. T. Kalife, C. Moore, M. O'Sullivan K, I. Desouza, M. E. Feeney, R. L. Eldridge, E. L. Maier, D. E. Kaufmann, M. P. Lahaie, L. Reyor, G. Tanzi, M. N. Johnston, C. Brander, R. Draenert, J. K. Rockstroh, H. Jessen, E. S. Rosenberg, S. A. Mallal, and B. D. Walker.** 2005. Selective escape from CD8+ T-cell responses represents a major driving force of human immunodeficiency virus type 1 (HIV-1) sequence diversity and reveals constraints on HIV-1 evolution. J Virol **79:**13239-49.

7. **Allen, T. M., M. Altfeld, X. G. Yu, K. M. O'Sullivan, M. Lichterfeld, S. Le Gall, M. John, B. R. Mothe, P. K. Lee, E. T. Kalife, D. E. Cohen, K. A. Freedberg, D. A. Strick, M. N. Johnston, A. Sette, E. S. Rosenberg, S. A. Mallal, P. J. Goulder, C. Brander, and B. D. Walker.** 2004. Selection, transmission, and reversion of an antigen-processing cytotoxic T-lymphocyte escape mutation in human immunodeficiency virus type 1 infection. J Virol **78:**7069-78.

8. **Allen, T. M., X. G. Yu, E. T. Kalife, L. L. Reyor, M. Lichterfeld, M. John, M. Cheng, R. L. Allgaier, S. Mui, N. Frahm, G. Alter, N. V. Brown, M. N. Johnston, E. S. Rosenberg, S. A. Mallal, C. Brander, B. D. Walker, and M. Altfeld.** 2005. De novo generation of escape variant-specific CD8+ T-cell responses following cytotoxic T-lymphocyte escape in chronic human immunodeficiency virus type 1 infection. J Virol **79:**12952-60.

9. **Almeida, J. R., D. A. Price, L. Papagno, Z. A. Arkoub, D. Sauce, E. Bornstein, T. E. Asher, A. Samri, A. Schnuriger, I. Theodorou, D. Costagliola, C. Rouzioux, H. Agut, A. G. Marcelin, D. Douek, B. Autran, and V. Appay.** 2007. Superior control of HIV-1 replication by CD8+ T cells is reflected by their avidity, polyfunctionality, and clonal turnover. J Exp Med **204:**2473-85.

10. **Almeida, J. R., D. Sauce, D. A. Price, L. Papagno, S. Y. Shin, A. Moris, M. Larsen, G. Pancino, D. C. Douek, B. Autran, A. Saez-Cirion, and V. Appay.** 2009. Antigen sensitivity is a major determinant of CD8+ T-cell polyfunctionality and HIV-suppressive activity. Blood **113:**6351-60.

11. **Altfeld, M., M. M. Addo, E. S. Rosenberg, F. M. Hecht, P. K. Lee, M. Vogel, X. G. Yu, R. Draenert, M. N. Johnston, D. Strick, T. M. Allen, M. E. Feeney, J. O. Kahn, R. P. Sekaly, J. A. Levy, J. K. Rockstroh, P. J. Goulder, and B. D. Walker.** 2003. Influence of HLA-B57 on clinical presentation and viral control during acute HIV-1 infection. AIDS **17:**2581-91.

12. **Altfeld, M., M. M. Addo, R. Shankarappa, P. K. Lee, T. M. Allen, X. G. Yu, A. Rathod, J. Harlow, K. O'Sullivan, M. N. Johnston, P. J. Goulder, J. I. Mullins, E. S. Rosenberg, C. Brander, B. Korber, and B. D. Walker.** 2003. Enhanced detection of human immunodeficiency virus type 1-specific T-cell responses to highly variable regions by using peptides based on autologous virus sequences. J Virol **77:**7330-40.

13. **Altfeld, M., E. T. Kalife, Y. Qi, H. Streeck, M. Lichterfeld, M. N. Johnston, N. Burgett, M. E. Swartz, A. Yang, G. Alter, X. G. Yu, A. Meier, J. K. Rockstroh, T. M. Allen, H. Jessen, E. S. Rosenberg, M. Carrington, and B. D. Walker.** 2006. HLA Alleles Associated with Delayed Progression to AIDS Contribute Strongly to the Initial CD8(+) T Cell Response against HIV-1. PLoS Med **3:**e403.

14. **Ammaranond, P., D. J. van Bockel, K. Petoumenos, M. McMurchie, R. Finlayson, M. G. Middleton, M. P. Davenport, V. Venturi, K. Suzuki, L. Gelgor, J. M. Kaldor, D. A. Cooper, and A. D. Kelleher.** 2011. HIV immune escape at an immunodominant epitope in HLA-B*27-positive individuals predicts viral load outcome. J Immunol **186:**479-88.

15. **Arenas, M., and D. Posada.** 2010. Computational design of centralized HIV-1 genes. Curr HIV Res **8:**613-21.

16. **Ayouba, A., S. Souquieres, B. Njinku, P. M. Martin, M. C. Muller-Trutwin, P. Roques, F. Barre-Sinoussi, P. Mauclere, F. Simon, and E. Nerrienet.** 2000. HIV-1 group N among HIV-1-seropositive individuals in Cameroon. AIDS **14:**2623-5.

17. **Bachand, F., X. J. Yao, M. Hrimech, N. Rougeau, and E. A. Cohen.** 1999. Incorporation of Vpr into human immunodeficiency virus type 1 requires a direct interaction with the p6 domain of the p55 gag precursor. J Biol Chem **274:**9083-91.

18. **Bailey, J. R., T. M. Williams, R. F. Siliciano, and J. N. Blankson.** 2006. Maintenance of viral suppression in HIV-1-infected HLA-B*57+ elite suppressors despite CTL escape mutations. J Exp Med **203:**1357-69.

19. **Bansal, A., E. Gough, D. Ritter, C. Wilson, J. Mulenga, S. Allen, and P. A. Goepfert.** 2006. Group M-based HIV-1 Gag peptides are frequently targeted by T cells in chronically infected US and Zambian patients. AIDS **20:**353-60.

20. **Bansal, A., S. Sabbaj, B. H. Edwards, D. Ritter, C. Perkins, J. Tang, J. J. Szinger, H. Weiss, P. A. Goepfert, B. Korber, C. M. Wilson, R. A. Kaslow, and M. J. Mulligan.** 2003. T cell responses in HIV type 1-infected adolescent minorities share similar epitope specificities with whites despite significant differences in HLA class I alleles. AIDS Res Hum Retroviruses **19:**1017-26.

21. **Barouch, D. H., K. L. O'Brien, N. L. Simmons, S. L. King, P. Abbink, L. F. Maxfield, Y. H. Sun, A. La Porte, A. M. Riggs, D. M. Lynch, S. L. Clark, K. Backus, J. R. Perry, M. S. Seaman, A. Carville, K. G. Mansfield, J. J. Szinger, W. Fischer, M. Muldoon, and B. Korber.** 2010. Mosaic HIV-1 vaccines expand the breadth and depth of cellular immune responses in rhesus monkeys. Nat Med **16:**319-23.

22. **Bennett, M. S., H. L. Ng, A. Ali, and O. O. Yang.** 2008. Cross-clade detection of HIV-1-specific cytotoxic T lymphocytes does not reflect cross-clade antiviral activity. J Infect Dis **197:**390-7.

23. **Berger, C. T., N. Frahm, D. A. Price, B. Mothe, M. Ghebremichael, K. L. Hartman, L. M. Henry, J. M. Brenchley, L. E. Ruff, V. Venturi, F. Pereyra, J. Sidney, A. Sette, D. C. Douek, B. D. Walker, D. E. Kaufmann, and C. Brander.** 2011. High-functional-avidity cytotoxic T lymphocyte responses to HLA-B-restricted Gag-derived epitopes associated with relative HIV control. J Virol **85:**9334-45.

24. **Berkowitz, R., J. Fisher, and S. P. Goff.** 1996. RNA packaging. Curr Top Microbiol Immunol **214:**177-218.

25. **Betts, M. R., D. R. Ambrozak, D. C. Douek, S. Bonhoeffer, J. M. Brenchley, J. P. Casazza, R. A. Koup, and L. J. Picker.** 2001. Analysis of total human immunodeficiency virus (HIV)-specific CD4(+) and CD8(+) T-cell responses: relationship to viral load in untreated HIV infection. J Virol **75:**11983-91.

26. **Betts, M. R., J. Krowka, C. Santamaria, K. Balsamo, F. Gao, G. Mulundu, C. Luo, N. N'Gandu, H. Sheppard, B. H. Hahn, S. Allen, and J. A. Frelinger.** 1997. Cross-clade human immunodeficiency virus (HIV)-specific cytotoxic T-lymphocyte responses in HIV-infected Zambians. J Virol **71:**8908-11.

27. **Betts, M. R., M. C. Nason, S. M. West, S. C. De Rosa, S. A. Migueles, J. Abraham, M. M. Lederman, J. M. Benito, P. A. Goepfert, M. Connors, M. Roederer, and R. A. Koup.** 2006. HIV nonprogressors preferentially maintain highly functional HIV-specific CD8+ T cells. Blood **107:**4781-9.

28. **Bhattacharya, T., M. Daniels, D. Heckerman, B. Foley, N. Frahm, C. Kadie, J. Carlson, K. Yusim, B. McMahon, B. Gaschen, S. Mallal, J. I. Mullins, D. C. Nickle, J. Herbeck, C. Rousseau, G. H. Learn, T. Miura, C. Brander, B. Walker, and B. Korber.** 2007. Founder effects in the assessment of HIV polymorphisms and HLA allele associations. Science **315:**1583-6.

29. **Bobkov, A. F., E. V. Kazennova, A. L. Sukhanova, M. R. Bobkova, V. V. Pokrovsky, V. V. Zeman, N. G. Kovtunenko, and I. B. Erasilova.** 2004. An HIV type 1 subtype A outbreak among injecting drug users in Kazakhstan. AIDS Res Hum Retroviruses **20:**1134-6.

30. **Borrow, P., H. Lewicki, B. H. Hahn, G. M. Shaw, and M. B. Oldstone.** 1994. Virus-specific CD8+ cytotoxic T-lymphocyte activity associated with control of viremia in primary human immunodeficiency virus type 1 infection. J Virol **68:**6103-10.

31. **Borrow, P., H. Lewicki, X. Wei, M. S. Horwitz, N. Peffer, H. Meyers, J. A. Nelson, J. E. Gairin, B. H. Hahn, M. B. Oldstone, and G. M. Shaw.** 1997. Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. Nat Med **3:**205-11.

32. **Braaten, D., and J. Luban.** 2001. Cyclophilin A regulates HIV-1 infectivity, as demonstrated by gene targeting in human T cells. EMBO J **20:**1300-9.

33. **Brambilla, A., L. Turchetto, A. Gatti, C. Bovolenta, F. Veglia, E. Santagostino, A. Gringeri, M. Clementi, G. Poli, P. Bagnarelli, and E. Vicenzi.** 1999. Defective nef alleles in a cohort of hemophiliacs with progressing and nonprogressing HIV-1 infection. Virology **259:**349-68.

34. **Brennan, F., and A. Foey.** 2002. Cytokine regulation in RA synovial tissue: role of T cell/macrophage contact-dependent interactions. Arthritis Res **4 Suppl 3:**S177-82.

35. **Briggs, J. A., M. N. Simon, I. Gross, H. G. Krausslich, S. D. Fuller, V. M. Vogt, and M. C. Johnson.** 2004. The stoichiometry of Gag protein in HIV-1. Nat Struct Mol Biol **11:**672-5.

36. **Brockman, M. A., A. Schneidewind, M. Lahaie, A. Schmidt, T. Miura, I. Desouza, F. Ryvkin, C. A. Derdeyn, S. Allen, E. Hunter, J. Mulenga, P. A. Goepfert, B. D. Walker, and T. M. Allen.** 2007. Escape and compensation from early HLA-B57-mediated cytotoxic T-lymphocyte pressure on human immunodeficiency virus type 1 Gag alter capsid interactions with cyclophilin A. J Virol **81:**12608-18.

37. **Browning, M., and P. Krausa.** 1996. Genetic diversity of HLA-A2: evolutionary and functional significance. Immunol Today **17:**165-70.

38. **Brumme, Z. L., C. J. Brumme, D. Heckerman, B. T. Korber, M. Daniels, J. Carlson, C. Kadie, T. Bhattacharya, C. Chui, J. Szinger, T. Mo, R. S. Hogg, J. S. Montaner, N. Frahm, C. Brander, B. D. Walker, and P. R. Harrigan.** 2007. Evidence of differential HLA class I-mediated viral evolution in functional and accessory/regulatory genes of HIV-1. PLoS Pathog **3:**e94.

39. **Brumme, Z. L., M. John, J. M. Carlson, C. J. Brumme, D. Chan, M. A. Brockman, L. C. Swenson, I. Tao, S. Szeto, P. Rosato, J. Sela, C. M. Kadie, N. Frahm, C. Brander, D. W. Haas, S. A. Riddler, R. Haubrich, B. D. Walker, P. R.**

Harrigan, D. Heckerman, and S. Mallal. 2009. HLA-associated immune escape pathways in HIV-1 subtype B Gag, Pol and Nef proteins. PLoS One **4:**e6687.

40. **Buchbinder, S. P., D. V. Mehrotra, A. Duerr, D. W. Fitzgerald, R. Mogg, D. Li, P. B. Gilbert, J. R. Lama, M. Marmor, C. Del Rio, M. J. McElrath, D. R. Casimiro, K. M. Gottesdiener, J. A. Chodakewitz, L. Corey, M. N. Robertson, and T. Step Study Protocol.** 2008. Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. Lancet **372:**1881-93.

41. **Bukovsky, A., and H. Gottlinger.** 1996. Lack of integrase can markedly affect human immunodeficiency virus type 1 particle production in the presence of an active viral protease. J Virol **70:**6820-5.

42. **Bunnik, E. M., Z. Euler, M. R. Welkers, B. D. Boeser-Nunnink, M. L. Grijsen, J. M. Prins, and H. Schuitemaker.** 2010. Adaptation of HIV-1 envelope gp120 to humoral immunity at a population level. Nat Med **16:**995-7.

43. **Buonaguro, L., E. Del Guadio, M. Monaco, D. Greco, P. Corti, E. Beth-Giraldo, F. M. Buonaguro, and G. Giraldo.** 1995. Heteroduplex mobility assay and phylogenetic analysis of V3 region sequences of human immunodeficiency virus type 1 isolates from Gulu, northern Uganda. The Italian-Ugandan Cooperation AIDS Program. J Virol **69:**7971-81.

44. **Burgers, W. A., G. K. Chege, T. L. Muller, J. H. van Harmelen, G. Khoury, E. G. Shephard, C. M. Gray, C. Williamson, and A. L. Williamson.** 2009. Broad, high-magnitude and multifunctional CD4+ and CD8+ T-cell responses elicited by a DNA and modified vaccinia Ankara vaccine containing human immunodeficiency virus type 1 subtype C genes in baboons. J Gen Virol **90:**468-80.

45. **Burke, D. S.** 1997. Recombination in HIV: an important viral evolutionary strategy. Emerg Infect Dis **3:**253-9.

46. **Burton, D. R., R. C. Desrosiers, R. W. Doms, W. C. Koff, P. D. Kwong, J. P. Moore, G. J. Nabel, J. Sodroski, I. A. Wilson, and R. T. Wyatt.** 2004. HIV vaccine design and the neutralizing antibody problem. Nat Immunol **5:**233-6.

47. **Buseyne, F., M. L. Chaix, B. Fleury, O. Manigart, M. Burgard, S. Blanche, C. Rouzioux, and Y. Riviere.** 1998. Cross-clade-specific cytotoxic T lymphocytes in HIV-1-infected children. Virology **250:**316-24.

48. **Buseyne, F., and Y. Riviere.** 2001. The flexibility of the TCR allows recognition of a large set of naturally occurring epitope variants by HIV-specific cytotoxic T lymphocytes. Int Immunol **13:**941-50.

49. **Buseyne, F., D. Scott-Algara, F. Porrot, B. Corre, N. Bellal, M. Burgard, C. Rouzioux, S. Blanche, and Y. Riviere.** 2002. Frequencies of ex vivo-activated human immunodeficiency virus type 1-specific gamma-interferon-producing CD8+ T cells in infected children correlate positively with plasma viral load. J Virol **76:**12414-22.

50. **Calarota, S. A., A. Foli, R. Maserati, F. Baldanti, S. Paolucci, M. A. Young, C. M. Tsoukas, J. Lisziewicz, and F. Lori.** 2008. HIV-1-specific T cell precursors with high proliferative capacity correlate with low viremia and high CD4 counts in untreated individuals. J Immunol **180:**5907-15.

51. **Cale, E. M., H. S. Bazick, T. A. Rianprakaisang, S. M. Alam, and N. L. Letvin.** 2011. Mutations in a dominant Nef epitope of simian immunodeficiency virus diminish TCR:epitope peptide affinity but not epitope peptide:MHC class I binding. J Immunol **187:**3300-13.

52. **Camerini, D., and B. Seed.** 1990. A CD4 domain important for HIV-mediated syncytium formation lies outside the virus binding site. Cell **60:**747-54.

53. **Cao, H., I. Mani, R. Vincent, R. Mugerwa, P. Mugyenyi, P. Kanki, J. Ellner, and B. D. Walker.** 2000. Cellular immunity to human immunodeficiency virus type 1 (HIV-1) clades: relevance to HIV-1 vaccine trials in Uganda. J Infect Dis **182:**1350-6.

54. **Cao, K., A. M. Moormann, K. E. Lyke, C. Masaberg, O. P. Sumba, O. K. Doumbo, D. Koech, A. Lancaster, M. Nelson, D. Meyer, R. Single, R. J. Hartzman, C. V. Plowe, J. Kazura, D. L. Mann, M. B. Sztein, G. Thomson, and M. A. Fernandez-Vina.** 2004. Differentiation between African populations is evidenced by the diversity of alleles and haplotypes of HLA class I loci. Tissue Antigens **63:**293-325.

55. **Carlson, J. M., Z. L. Brumme, C. M. Rousseau, C. J. Brumme, P. Matthews, C. Kadie, J. I. Mullins, B. D. Walker, P. R. Harrigan, P. J. Goulder, and D. Heckerman.** 2008. Phylogenetic dependency networks: inferring patterns of CTL escape and codon covariation in HIV-1 Gag. PLoS Comput Biol **4:**e1000225.

56. **Carr, A., and D. A. Cooper.** 1998. HIV and AIDS. Gap between biology and reality in AIDS. Lancet **352 Suppl 4:**SIV16.

57. **Carrington, M., G. W. Nelson, M. P. Martin, T. Kissner, D. Vlahov, J. J. Goedert, R. Kaslow, S. Buchbinder, K. Hoots, and S. J. O'Brien.** 1999. HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. Science **283:**1748-52.

58. **Charpentier, C., T. Nora, O. Tenaillon, F. Clavel, and A. J. Hance.** 2006. Extensive recombination among human immunodeficiency virus type 1 quasispecies makes an important contribution to viral diversity in individual patients. J Virol **80:**2472-82.

59. **Choisy, M., C. H. Woelk, J. F. Guegan, and D. L. Robertson.** 2004. Comparative study of adaptive molecular evolution in different human immunodeficiency virus groups and subtypes. J Virol **78:**1962-70.

60. **Coffin, J. M., S. H. Hughes, and H. E. Varmus.** 1997. The Interactions of Retroviruses and their Hosts.

61. **Coplan, P. M., S. B. Gupta, S. A. Dubey, P. Pitisuttithum, A. Nikas, B. Mbewe, E. Vardas, M. Schechter, E. G. Kallas, D. C. Freed, T. M. Fu, C. T. Mast, P. Puthavathana, J. Kublin, K. Brown Collins, J. Chisi, R. Pendame, S. J. Thaler, G. Gray, J. McIntyre, W. L. Straus, J. H. Condra, D. V. Mehrotra, H. A. Guess, E. A. Emini, and J. W. Shiver.** 2005. Cross-reactivity of anti-HIV-1 T cell immune responses among the major HIV-1 clades in HIV-1-positive individuals from 4 continents. J Infect Dis **191:**1427-34.

62. **Coren, L. V., J. A. Thomas, E. Chertova, R. C. Sowder, 2nd, T. D. Gagliardi, R. J. Gorelick, and D. E. Ott.** 2007. Mutational analysis of the C-terminal gag cleavage sites in human immunodeficiency virus type 1. J Virol **81:**10047-54.

63. **Corey, L., and M. J. McElrath.** 2010. HIV vaccines: mosaic approach to virus diversity. Nat Med **16:**268-70.

64. **Craigie, R.** 2001. HIV integrase, a brief overview from chemistry to therapeutics. J Biol Chem **276:**23213-6.

65. **Crawford, H., J. G. Prado, A. Leslie, S. Hue, I. Honeyborne, S. Reddy, M. van der Stok, Z. Mncube, C. Brander, C. Rousseau, J. I. Mullins, R. Kaslow, P. Goepfert, S. Allen, E. Hunter, J. Mulenga, P. Kiepiela, B. D. Walker, and P. J. Goulder.** 2007. Compensatory mutation partially restores fitness and delays reversion of escape mutation within the immunodominant HLA-B*5703-restricted Gag epitope in chronic human immunodeficiency virus type 1 infection. J Virol **81:**8346-51.

66. **Currier, J. R., W. E. Dowling, K. M. Wasunna, U. Alam, C. J. Mason, M. L. Robb, J. K. Carr, F. E. McCutchan, D. L. Birx, and J. H. Cox.** 2003. Detection of high frequencies of HIV-1 cross-subtype reactive CD8 T lymphocytes in the peripheral blood of HIV-1-infected Kenyans. AIDS **17:**2149-57.

67. **Dahirel, V., K. Shekhar, F. Pereyra, T. Miura, M. Artyomov, S. Talsania, T. M. Allen, M. Altfeld, M. Carrington, D. J. Irvine, B. D. Walker, and A. K. Chakraborty.** 2011. Coordinate linkage of HIV evolution reveals regions of immunological vulnerability. Proc Natl Acad Sci U S A **108:**11530-5.

68. **Dalmasso, C., W. Carpentier, L. Meyer, C. Rouzioux, C. Goujard, M. L. Chaix, O. Lambotte, V. Avettand-Fenoel, S. Le Clerc, L. D. de Senneville, C. Deveau, F. Boufassa, P. Debre, J. F. Delfraissy, P. Broet, I. Theodorou, and A. G. W. Association.** 2008. Distinct genetic loci control plasma HIV-RNA and cellular HIV-DNA levels in HIV-1 infection: the ANRS Genome Wide Association 01 study. PLoS One **3:**e3907.

69. **Datta, S. A., F. Heinrich, S. Raghunandan, S. Krueger, J. E. Curtis, A. Rein, and H. Nanda.** 2011. HIV-1 Gag extension: conformational changes require simultaneous interaction with membrane and nucleic acid. J Mol Biol **406:**205-14.

70. **Davenport, M. P., L. Loh, J. Petravic, and S. J. Kent.** 2008. Rates of HIV immune escape and reversion: implications for vaccination. Trends Microbiol **16:**561-6.

71. **Day, C. L., P. Kiepiela, A. J. Leslie, M. van der Stok, K. Nair, N. Ismail, I. Honeyborne, H. Crawford, H. M. Coovadia, P. J. Goulder, B. D. Walker, and P. Klenerman.** 2007. Proliferative capacity of epitope-specific CD8 T-cell responses is inversely related to viral load in chronic human immunodeficiency virus type 1 infection. J Virol **81:**434-8.

72. **Dazza, M. C., M. Ekwalanga, M. Nende, K. B. Shamamba, P. Bitshi, D. Paraskevis, and S. Saragosti.** 2005. Characterization of a novel vpu-harboring simian immunodeficiency virus from a Dent's Mona monkey (Cercopithecus mona denti). J Virol **79:**8560-71.

73. **De Rosa, S. C., F. X. Lu, J. Yu, S. P. Perfetto, J. Falloon, S. Moser, T. G. Evans, R. Koup, C. J. Miller, and M. Roederer.** 2004. Vaccination in humans generates broad T cell cytokine responses. J Immunol **173:**5372-80.

74. **Deacon, N. J., A. Tsykin, A. Solomon, K. Smith, M. Ludford-Menting, D. J. Hooker, D. A. McPhee, A. L. Greenway, A. Ellett, C. Chatfield, V. A. Lawson, S. Crowe, A. Maerz, S. Sonza, J. Learmont, J. S. Sullivan, A. Cunningham, D. Dwyer, D. Dowton, and J. Mills.** 1995. Genomic structure of an attenuated quasi species of HIV-1 from a blood transfusion donor and recipients. Science **270:**988-91.

75. **Deng, H., R. Liu, W. Ellmeier, S. Choe, D. Unutmaz, M. Burkhart, P. Di Marzio, S. Marmon, R. E. Sutton, C. M. Hill, C. B. Davis, S. C. Peiper, T. J. Schall, D. R. Littman, and N. R. Landau.** 1996. Identification of a major co-receptor for primary isolates of HIV-1. Nature **381:**661-6.

76. **Deora, A., and L. Ratner.** 2001. Viral protein U (Vpu)-mediated enhancement of human immunodeficiency virus type 1 particle release depends on the rate of cellular proliferation. J Virol **75:**6714-8.

77. **Derby, M., M. Alexander-Miller, R. Tse, and J. Berzofsky.** 2001. High-avidity CTL exploit two complementary mechanisms to provide better protection against viral infection than low-avidity CTL. J Immunol **166:**1690-7.

78. **Doherty, R. S., T. De Oliveira, C. Seebregts, S. Danaviah, M. Gordon, and S. Cassol.** 2005. BioAfrica's HIV-1 proteomics resource: combining protein data with bioinformatics tools. Retrovirology **2:**18.

79. **Dong, T., Y. Zhang, K. Y. Xu, H. Yan, I. James, Y. Peng, M. E. Blais, S. Gaudieri, X. Chen, W. Lun, H. Wu, W. Y. Qu, T. Rostron, N. Li, Y. Mao, S. Mallal, X. Xu, A. McMichael, M. John, and S. L. Rowland-Jones.** 2011. Extensive HLA-driven viral diversity following a narrow-source HIV-1 outbreak in rural China. Blood **118:**98-106.

80. **Dorfman, T., and H. G. Gottlinger.** 1996. The human immunodeficiency virus type 1 capsid p2 domain confers sensitivity to the cyclophilin-binding drug SDZ NIM 811. J Virol **70:**5751-7.

81. **Doria-Rose, N. A., G. H. Learn, A. G. Rodrigo, D. C. Nickle, F. Li, M. Mahalanabis, M. T. Hensel, S. McLaughlin, P. F. Edmonson, D. Montefiori, S.**

W. Barnett, N. L. Haigwood, and J. I. Mullins. 2005. Human immunodeficiency virus type 1 subtype B ancestral envelope protein is functional and elicits neutralizing antibodies in rabbits similar to those elicited by a circulating subtype B envelope. J Virol **79:**11214-24.

82. **Douglas, N. W., G. H. Munro, and R. S. Daniels.** 1997. HIV/SIV glycoproteins: structure-function relationships. J Mol Biol **273:**122-49.

83. **Doytchinova, I. A., P. Guan, and D. R. Flower.** 2004. Identifiying human MHC supertypes using bioinformatic methods. J Immunol **172:**4314-23.

84. **Draenert, R., S. Le Gall, K. J. Pfafferott, A. J. Leslie, P. Chetty, C. Brander, E. C. Holmes, S. C. Chang, M. E. Feeney, M. M. Addo, L. Ruiz, D. Ramduth, P. Jeena, M. Altfeld, S. Thomas, Y. Tang, C. L. Verrill, C. Dixon, J. G. Prado, P. Kiepiela, J. Martinez-Picado, B. D. Walker, and P. J. Goulder.** 2004. Immune selection for altered antigen processing leads to cytotoxic T lymphocyte escape in chronic HIV-1 infection. J Exp Med **199:**905-15.

85. **Draenert, R., C. L. Verrill, Y. Tang, T. M. Allen, A. G. Wurcel, M. Boczanowski, A. Lechner, A. Y. Kim, T. Suscovich, N. V. Brown, M. M. Addo, and B. D. Walker.** 2004. Persistent recognition of autologous virus by high-avidity CD8 T cells in chronic, progressive human immunodeficiency virus type 1 infection. J Virol **78:**630-41.

86. **du Toit, E. D., J. C. Emmanuel, G. West, D. G. Taljaard, and M. Oudshoorn.** 1990. HLA frequencies in black Zimbabweans. Tissue Antigens **36:**122-4.

87. **Dube, M., B. B. Roy, P. Guiot-Guillain, J. Binette, J. Mercier, A. Chiasson, and E. A. Cohen.** 2010. Antagonism of tetherin restriction of HIV-1 release by Vpu involves binding and sequestration of the restriction factor in a perinuclear compartment. PLoS Pathog **6:**e1000856.

88. **Dussupt, V., P. Sette, N. F. Bello, M. P. Javid, K. Nagashima, and F. Bouamr.** 2011. Basic residues in the nucleocapsid domain of Gag are critical for late events of HIV-1 budding. J Virol **85:**2304-15.

89. **Eckstein, D. A., M. P. Sherman, M. L. Penn, P. S. Chin, C. M. De Noronha, W. C. Greene, and M. A. Goldsmith.** 2001. HIV-1 Vpr enhances viral burden by facilitating infection of tissue macrophages but not nondividing CD4+ T cells. J Exp Med **194:**1407-19.

90. **Edwards, B. H., A. Bansal, S. Sabbaj, J. Bakari, M. J. Mulligan, and P. A. Goepfert.** 2002. Magnitude of functional CD8+ T-cell responses to the gag protein of human immunodeficiency virus type 1 correlates inversely with viral load in plasma. J Virol **76:**2298-305.

91. **Elahi, S., W. L. Dinges, N. Lejarcegui, K. J. Laing, A. C. Collier, D. M. Koelle, M. J. McElrath, and H. Horton.** 2011. Protective HIV-specific CD8+ T cells evade Treg cell suppression. Nat Med **17:**989-95.

92. **Ellenberger, D., Bage, S., Butera, S., Folks, T., Kraiseburd, E., Lanier, N., Li, B., Martinez, M., McNicholl, J., Moss, B., Otten, R., Robinson, H., Rodriguez, I., Sariol, C., Smith, J., Wyatt, L., Yi, H.** 2003. Fellows and New Investigators Workshop on HIV Vaccine Development: Immunological and Biological Challenges: Twenty Years of HIV Research: From Discovery to Understanding, p. 134, Keystone Symposia, Silverthone.

93. **Ellenberger, D. L., B. Li, L. D. Lupo, S. M. Owen, J. Nkengasong, M. S. Kadio-Morokro, J. Smith, H. Robinson, M. Ackers, A. Greenberg, T. Folks, and S. Butera.** 2002. Generation of a consensus sequence from prevalent and incident HIV-1 infections in West Africa to guide AIDS vaccine development. Virology **302:**155-63.

94. **Fan, J., G. Ma, K. Nosaka, J. Tanabe, Y. Satou, A. Koito, S. Wain-Hobson, J. P. Vartanian, and M. Matsuoka.** 2010. APOBEC3G generates nonsense mutations in human T-cell leukemia virus type 1 proviral genomes in vivo. J Virol **84:**7278-87.

95. **Fassati, A.** 2006. HIV infection of non-dividing cells: a divisive problem. Retrovirology **3:**74.

96. **Feinberg, M. B., D. Baltimore, and A. D. Frankel.** 1991. The role of Tat in the human immunodeficiency virus life cycle indicates a primary effect on transcriptional elongation. Proc Natl Acad Sci U S A **88:**4045-9.

97. **Fellay, J., D. Ge, K. V. Shianna, S. Colombo, B. Ledergerber, E. T. Cirulli, T. J. Urban, K. Zhang, C. E. Gumbs, J. P. Smith, A. Castagna, A. Cozzi-Lepri, A. De Luca, P. Easterbrook, H. F. Gunthard, S. Mallal, C. Mussini, J. Dalmau, J. Martinez-Picado, J. M. Miro, N. Obel, S. M. Wolinsky, J. J. Martinson, R. Detels, J. B. Margolick, L. P. Jacobson, P. Descombes, S. E. Antonarakis, J. S. Beckmann, S. J. O'Brien, N. L. Letvin, A. J. McMichael, B. F. Haynes, M. Carrington, S. Feng, A. Telenti, D. B. Goldstein, and N. C. f. H. A. V. Immunology.** 2009. Common genetic variation and the control of HIV-1 in humans. PLoS Genet **5:**e1000791.

98. **Fellay, J., K. V. Shianna, D. Ge, S. Colombo, B. Ledergerber, M. Weale, K. Zhang, C. Gumbs, A. Castagna, A. Cossarizza, A. Cozzi-Lepri, A. De Luca, P. Easterbrook, P. Francioli, S. Mallal, J. Martinez-Picado, J. M. Miro, N. Obel, J. P. Smith, J. Wyniger, P. Descombes, S. E. Antonarakis, N. L. Letvin, A. J. McMichael, B. F. Haynes, A. Telenti, and D. B. Goldstein.** 2007. A whole-genome association study of major determinants for host control of HIV-1. Science **317:**944-7.

99. **Feng, Y., C. C. Broder, P. E. Kennedy, and E. A. Berger.** 1996. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. Science **272:**872-7.

100. **Ferrari, G., J. R. Currier, M. E. Harris, S. Finkelstein, A. de Oliveira, D. Barkhan, J. H. Cox, M. Zeira, K. J. Weinhold, N. Reinsmoen, F. McCutchan, D. L. Birx, S. Osmanov, and S. Maayan.** 2004. HLA-A and -B allele expression and ability to develop anti-Gag cross-clade responses in subtype C HIV-1-infected Ethiopians. Hum Immunol **65:**648-59.

101. **Ferrari, G., W. Humphrey, M. J. McElrath, J. L. Excler, A. M. Duliege, M. L. Clements, L. C. Corey, D. P. Bolognesi, and K. J. Weinhold.** 1997. Clade B-based HIV-1 vaccines elicit cross-clade cytotoxic T lymphocyte reactivities in uninfected volunteers. Proc Natl Acad Sci U S A **94:**1396-401.

102. **Fischer, W., H. X. Liao, B. F. Haynes, N. L. Letvin, and B. Korber.** 2008. Coping with viral diversity in HIV vaccine design: a response to Nickle et al. PLoS Comput Biol **4:**e15; author reply e25.

103. **Fischer, W., S. Perkins, J. Theiler, T. Bhattacharya, K. Yusim, R. Funkhouser, C. Kuiken, B. Haynes, N. L. Letvin, B. D. Walker, B. H. Hahn, and B. T. Korber.** 2007. Polyvalent vaccines for optimal coverage of potential T-cell epitopes in global HIV-1 variants. Nat Med **13:**100-6.

104. **Frahm, N., S. Adams, P. Kiepiela, C. H. Linde, H. S. Hewitt, M. Lichterfeld, K. Sango, N. V. Brown, E. Pae, A. G. Wurcel, M. Altfeld, M. E. Feeney, T. M. Allen, T. Roach, M. A. St John, E. S. Daar, E. Rosenberg, B. Korber, F. Marincola, B. D. Walker, P. J. Goulder, and C. Brander.** 2005. HLA-B63 presents HLA-B57/B58-restricted cytotoxic T-lymphocyte epitopes and is associated with low human immunodeficiency virus load. J Virol **79:**10218-25.

105. **Frahm, N., D. E. Kaufmann, K. Yusim, M. Muldoon, C. Kesmir, C. H. Linde, W. Fischer, T. M. Allen, B. Li, B. H. McMahon, K. L. Faircloth, H. S. Hewitt, E. W. Mackey, T. Miura, A. Khatri, S. Wolinsky, A. McMichael, R. K. Funkhouser, B. D. Walker, C. Brander, and B. T. Korber.** 2007. Increased sequence diversity coverage improves detection of HIV-specific T cell responses. J Immunol **179:**6638-50.

106. **Frahm, N., P. Kiepiela, S. Adams, C. H. Linde, H. S. Hewitt, K. Sango, M. E. Feeney, M. M. Addo, M. Lichterfeld, M. P. Lahaie, E. Pae, A. G. Wurcel, T. Roach, M. A. St John, M. Altfeld, F. M. Marincola, C. Moore, S. Mallal, M. Carrington, D. Heckerman, T. M. Allen, J. I. Mullins, B. T. Korber, P. J. Goulder, B. D. Walker, and C. Brander.** 2006. Control of human immunodeficiency virus replication by cytotoxic T lymphocytes targeting subdominant epitopes. Nat Immunol **7:**173-8.

107. **Frahm, N., D. C. Nickle, C. H. Linde, D. E. Cohen, R. Zuniga, A. Lucchetti, T. Roach, B. D. Walker, T. M. Allen, B. T. Korber, J. I. Mullins, and C. Brander.** 2008. Increased detection of HIV-specific T cell responses by combination of central sequences with comparable immunogenicity. AIDS **22:**447-56.

108. **Friedrich, T. C., L. E. Valentine, L. J. Yant, E. G. Rakasz, S. M. Piaskowski, J. R. Furlott, K. L. Weisgrau, B. Burwitz, G. E. May, E. J. Leon, T. Soma, G. Napoe, S. V. Capuano, 3rd, N. A. Wilson, and D. I. Watkins.** 2007. Subdominant CD8+ T-cell responses are involved in durable control of AIDS virus replication. J Virol **81:**3465-76.

109. **Ganser-Pornillos, B. K., M. Yeager, and W. I. Sundquist.** 2008. The structural biology of HIV assembly. Curr Opin Struct Biol **18:**203-17.

110. **Gao, F., E. Bailes, D. L. Robertson, Y. Chen, C. M. Rodenburg, S. F. Michael, L. B. Cummins, L. O. Arthur, M. Peeters, G. M. Shaw, P. M. Sharp, and B. H. Hahn.** 1999. Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. Nature **397:**436-41.

111. **Gao, F., S. G. Morrison, D. L. Robertson, C. L. Thornton, S. Craig, G. Karlsson, J. Sodroski, M. Morgado, B. Galvao-Castro, H. von Briesen, S. Beddows, J. Weber, P. M. Sharp, G. M. Shaw, and B. H. Hahn.** 1996. Molecular cloning and analysis of functional envelope genes from human immunodeficiency virus type 1 sequence subtypes A through G. The WHO and NIAID Networks for HIV Isolation and Characterization. J Virol **70:**1651-67.

112. **Gao, F., D. L. Robertson, C. D. Carruthers, Y. Li, E. Bailes, L. G. Kostrikis, M. O. Salminen, F. Bibollet-Ruche, M. Peeters, D. D. Ho, G. M. Shaw, P. M. Sharp, and B. H. Hahn.** 1998. An isolate of human immunodeficiency virus type 1 originally classified as subtype I represents a complex mosaic comprising three different group M subtypes (A, G, and I). J Virol **72:**10234-41.

113. **Gao, F., N. Vidal, Y. Li, S. A. Trask, Y. Chen, L. G. Kostrikis, D. D. Ho, J. Kim, M. D. Oh, K. Choe, M. Salminen, D. L. Robertson, G. M. Shaw, B. H. Hahn, and M. Peeters.** 2001. Evidence of two distinct subsubtypes within the HIV-1 subtype A radiation. AIDS Res Hum Retroviruses **17:**675-88.

114. **Gao, X., G. W. Nelson, P. Karacki, M. P. Martin, J. Phair, R. Kaslow, J. J. Goedert, S. Buchbinder, K. Hoots, D. Vlahov, S. J. O'Brien, and M. Carrington.** 2001. Effect of a single amino acid change in MHC class I molecules on the rate of progression to AIDS. N Engl J Med **344:**1668-75.

115. **Garber, D. A., G. Silvestri, and M. B. Feinberg.** 2004. Prospects for an AIDS vaccine: three big questions, no easy answers. Lancet Infect Dis **4:**397-413.

116. **Gaschen, B., J. Taylor, K. Yusim, B. Foley, F. Gao, D. Lang, V. Novitsky, B. Haynes, B. H. Hahn, T. Bhattacharya, and B. Korber.** 2002. Diversity considerations in HIV-1 vaccine selection. Science **296:**2354-60.

117. **Gebe, J. A., E. Swanson, and W. W. Kwok.** 2002. HLA class II peptide-binding and autoimmunity. Tissue Antigens **59:**78-87.

118. **Geels, M. J., S. A. Dubey, K. Anderson, E. Baan, M. Bakker, G. Pollakis, W. A. Paxton, J. W. Shiver, and J. Goudsmit.** 2005. Broad cross-clade T-cell responses to gag in individuals infected with human immunodeficiency virus type 1 non-B clades (A to G): importance of HLA anchor residue conservation. J Virol **79:**11247-58.

119. **Geijtenbeek, T. B., D. S. Kwon, R. Torensma, S. J. van Vliet, G. C. van Duijnhoven, J. Middel, I. L. Cornelissen, H. S. Nottet, V. N. KewalRamani, D. R. Littman, C. G. Figdor, and Y. van Kooyk.** 2000. DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. Cell **100:**587-97.

120. **Geldmacher, C., J. R. Currier, M. Gerhardt, A. Haule, L. Maboko, D. Birx, C. Gray, A. Meyerhans, J. Cox, and M. Hoelscher.** 2007. In a mixed subtype

epidemic, the HIV-1 Gag-specific T-cell response is biased towards the infecting subtype. AIDS **21:**135-43.

121. **Geldmacher, C., J. R. Currier, E. Herrmann, A. Haule, E. Kuta, F. McCutchan, L. Njovu, S. Geis, O. Hoffmann, L. Maboko, C. Williamson, D. Birx, A. Meyerhans, J. Cox, and M. Hoelscher.** 2007. CD8 T-cell recognition of multiple epitopes within specific Gag regions is associated with maintenance of a low steady-state viremia in human immunodeficiency virus type 1-seropositive patients. J Virol **81:**2440-8.

122. **Gesprasert, G., N. Wichukchinda, M. Mori, T. Shiino, W. Auwanit, B. Sriwanthana, P. Pathipvanich, P. Sawanpanyalert, T. Miura, P. Auewarakul, A. Thitithanyanont, and K. Ariyoshi.** 2010. HLA-associated immune pressure on Gag protein in CRF01_AE-infected individuals and its association with plasma viral load. PLoS One **5:**e11179.

123. **Gillespie, G. M., R. Kaul, T. Dong, H. B. Yang, T. Rostron, J. J. Bwayo, P. Kiama, T. Peto, F. A. Plummer, A. J. McMichael, and S. L. Rowland-Jones.** 2002. Cross-reactive cytotoxic T lymphocytes against a HIV-1 p24 epitope in slow progressors with B*57. AIDS **16:**961-72.

124. **Gonzalez-Ortega, E., E. Ballana, R. Badia, B. Clotet, and J. A. Este.** 2011. Compensatory mutations rescue the virus replicative capacity of VIRIP-resistant HIV-1. Antiviral Res **92:**479-83.

125. **Goonetilleke, N., M. K. Liu, J. F. Salazar-Gonzalez, G. Ferrari, E. Giorgi, V. V. Ganusov, B. F. Keele, G. H. Learn, E. L. Turnbull, M. G. Salazar, K. J. Weinhold, S. Moore, C. C. C. B, N. Letvin, B. F. Haynes, M. S. Cohen, P. Hraber, T. Bhattacharya, P. Borrow, A. S. Perelson, B. H. Hahn, G. M. Shaw, B. T. Korber, and A. J. McMichael.** 2009. The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection. J Exp Med **206:**1253-72.

126. **Gottlinger, H. G.** 2001. The HIV-1 assembly machine. AIDS **15 Suppl 5:**S13-20.

127. **Goulder, P. J., M. Bunce, P. Krausa, K. McIntyre, S. Crowley, B. Morgan, A. Edwards, P. Giangrande, R. E. Phillips, and A. J. McMichael.** 1996. Novel, cross-restricted, conserved, and immunodominant cytotoxic T lymphocyte epitopes in slow progressors in HIV type 1 infection. AIDS Res Hum Retroviruses **12:**1691-8.

128. **Goulder, P. J., A. Edwards, R. E. Phillips, and A. J. McMichael.** 1997. Identification of a novel HLA-B*2705-restricted cytotoxic T-lymphocyte epitope within a conserved region of HIV-1 Nef. AIDS **11:**536-8.

129. **Goulder, P. J., R. E. Phillips, R. A. Colbert, S. McAdam, G. Ogg, M. A. Nowak, P. Giangrande, G. Luzzi, B. Morgan, A. Edwards, A. J. McMichael, and S. Rowland-Jones.** 1997. Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. Nat Med **3:**212-7.

130. **Goulder, P. J., and D. I. Watkins.** 2004. HIV and SIV CTL escape: implications for vaccine design. Nat Rev Immunol **4:**630-40.

131. **Gray, C. M., M. Mlotshwa, C. Riou, T. Mathebula, D. de Assis Rosa, T. Mashishi, C. Seoighe, N. Ngandu, F. van Loggerenberg, L. Morris, K. Mlisana, C. Williamson, S. A. Karim, and C. A. I. S. Team.** 2009. Human immunodeficiency virus-specific gamma interferon enzyme-linked immunospot assay responses targeting specific regions of the proteome during primary subtype C infection are poor predictors of the course of viremia and set point. J Virol **83:**470-8.

132. **Gray, G., S. Buchbinder, and A. Duerr.** 2010. Overview of STEP and Phambili trial results: two phase IIb test-of-concept studies investigating the efficacy of MRK adenovirus type 5 gag/pol/nef subtype B HIV vaccine. Curr Opin HIV AIDS **5:**357-61.

133. **Gray, G. E., M. Allen, Z. Moodie, G. Churchyard, L. G. Bekker, M. Nchabeleng, K. Mlisana, B. Metch, G. de Bruyn, M. H. Latka, S. Roux, M. Mathebula, N. Naicker, C. Ducar, D. K. Carter, A. Puren, N. Eaton, M. J. McElrath, M. Robertson, L. Corey, J. G. Kublin, and H. P. s. team.** 2011. Safety and efficacy of the HVTN 503/Phambili study of a clade-B-based HIV-1 vaccine in South Africa: a double-blind, randomised, placebo-controlled test-of-concept phase 2b study. Lancet Infect Dis **11:**507-15.

134. **Gudmundsdotter, L., D. Bernasconi, B. Hejdeman, E. Sandstrom, A. Alaeus, K. Lidman, B. Ensoli, B. Wahren, and S. Butto.** 2008. Cross-clade immune responses to Gag p24 in patients infected with different HIV-1 subtypes and correlation with HLA class I and II alleles. Vaccine **26:**5182-7.

135. **Gupta, S. B., C. T. Mast, N. D. Wolfe, V. Novitsky, S. A. Dubey, E. G. Kallas, M. Schechter, B. Mbewe, E. Vardas, P. Pitisuttithum, D. Burke, D. Freed, R. Mogg, P. M. Coplan, J. H. Condra, R. S. Long, K. Anderson, D. R. Casimiro, J. W. Shiver, and W. L. Straus.** 2006. Cross-clade reactivity of HIV-1-specific T-cell responses in HIV-1-infected individuals from Botswana and Cameroon. J Acquir Immune Defic Syndr **42:**135-9.

136. **Gurtler, L. G., P. H. Hauser, J. Eberle, A. von Brunn, S. Knapp, L. Zekeng, J. M. Tsague, and L. Kaptue.** 1994. A new subtype of human immunodeficiency virus type 1 (MVP-5180) from Cameroon. J Virol **68:**1581-5.

137. **Haanen, J. B., M. C. Wolkers, A. M. Kruisbeek, and T. N. Schumacher.** 1999. Selective expansion of cross-reactive CD8(+) memory T cells by viral variants. J Exp Med **190:**1319-28.

138. **Haynes, B. F., P. B. Gilbert, M. J. McElrath, S. Zolla-Pazner, G. D. Tomaras, S. M. Alam, D. T. Evans, D. C. Montefiori, C. Karnasuta, R. Sutthent, H. X. Liao, A. L. DeVico, G. K. Lewis, C. Williams, A. Pinter, Y. Fong, H. Janes, A. DeCamp, Y. Huang, M. Rao, E. Billings, N. Karasavvas, M. L. Robb, V. Ngauy, M. S. de Souza, R. Paris, G. Ferrari, R. T. Bailer, K. A. Soderberg, C. Andrews, P. W. Berman, N. Frahm, S. C. De Rosa, M. D. Alpert, N. L. Yates, X. Shen, R. A. Koup, P. Pitisuttithum, J. Kaewkungwal, S. Nitayaphan, S. Rerks-Ngarm, N.**

L. Michael, and J. H. Kim. 2012. Immune-correlates analysis of an HIV-1 vaccine efficacy trial. N Engl J Med **366:**1275-86.

139. **He, N., C. K. Chan, B. Sobhian, S. Chou, Y. Xue, M. Liu, T. Alber, M. Benkirane, and Q. Zhou.** 2011. Human Polymerase-Associated Factor complex (PAFc) connects the Super Elongation Complex (SEC) to RNA polymerase II on chromatin. Proc Natl Acad Sci U S A **108:**E636-45.

140. **Heath, M. J., S. S. Derebail, R. J. Gorelick, and J. J. DeStefano.** 2003. Differing roles of the N- and C-terminal zinc fingers in human immunodeficiency virus nucleocapsid protein-enhanced nucleic acid annealing. J Biol Chem **278:**30755-63.

141. **Hemelaar, J., E. Gouws, P. D. Ghys, and S. Osmanov.** 2006. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. AIDS **20:**W13-23.

142. **Hemelaar, J., E. Gouws, P. D. Ghys, S. Osmanov, W.-U. N. f. H. Isolation, and Characterisation.** 2011. Global trends in molecular epidemiology of HIV-1 during 2000-2007. AIDS **25:**679-89.

143. **Henderson, L. E., M. A. Bowers, R. C. Sowder, 2nd, S. A. Serabyn, D. G. Johnson, J. W. Bess, Jr., L. O. Arthur, D. K. Bryant, and C. Fenselau.** 1992. Gag proteins of the highly replicative MN strain of human immunodeficiency virus type 1: posttranslational modifications, proteolytic processings, and complete amino acid sequences. J Virol **66:**1856-65.

144. **Hersperger, A. R., J. N. Martin, L. Y. Shin, P. M. Sheth, C. M. Kovacs, G. L. Cosma, G. Makedonas, F. Pereyra, B. D. Walker, R. Kaul, S. G. Deeks, and M. R. Betts.** 2011. Increased HIV-specific CD8+ T-cell cytotoxic potential in HIV elite controllers is associated with T-bet expression. Blood **117:**3799-808.

145. **Hersperger, A. R., S. A. Migueles, M. R. Betts, and M. Connors.** 2011. Qualitative features of the HIV-specific CD8+ T-cell response associated with immunologic control. Curr Opin HIV AIDS **6:**169-73.

146. **Hersperger, A. R., F. Pereyra, M. Nason, K. Demers, P. Sheth, L. Y. Shin, C. M. Kovacs, B. Rodriguez, S. F. Sieg, L. Teixeira-Johnson, D. Gudonis, P. A. Goepfert, M. M. Lederman, I. Frank, G. Makedonas, R. Kaul, B. D. Walker, and M. R. Betts.** 2010. Perforin expression directly ex vivo by HIV-specific CD8 T-cells is a correlate of HIV elite control. PLoS Pathog **6:**e1000917.

147. **Ho, D. D., A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard, and M. Markowitz.** 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature **373:**123-6.

148. **Hockett, R. D., J. M. Kilby, C. A. Derdeyn, M. S. Saag, M. Sillers, K. Squires, S. Chiz, M. A. Nowak, G. M. Shaw, and R. P. Bucy.** 1999. Constant mean viral copy number per infected cell in tissues regardless of high, low, or undetectable plasma HIV RNA. J Exp Med **189:**1545-54.

149. **Honeyborne, I., A. Rathod, R. Buchli, D. Ramduth, E. Moodley, P. Rathnavalu, S. Chetty, C. Day, C. Brander, W. Hildebrand, B. D. Walker, P. Kiepiela, and P. J. Goulder.** 2006. Motif inference reveals optimal CTL epitopes presented by HLA class I alleles highly prevalent in southern Africa. J Immunol **176:**4699-705.

150. **Hope, T. J.** 1999. The ins and outs of HIV Rev. Arch Biochem Biophys **365:**186-91.

151. **Horton, H., I. Frank, R. Baydo, E. Jalbert, J. Penn, S. Wilson, J. P. McNevin, M. D. McSweyn, D. Lee, Y. Huang, S. C. De Rosa, and M. J. McElrath.** 2006. Preservation of T cell proliferation restricted by protective HLA alleles is critical for immune control of HIV-1 infection. J Immunol **177:**7406-15.

152. **Hu, W. S., and H. M. Temin.** 1990. Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. Proc Natl Acad Sci U S A **87:**1556-60.

153. **Hu, W. S., and H. M. Temin.** 1990. Retroviral recombination and reverse transcription. Science **250:**1227-33.

154. **Huang, J., J. J. Goedert, E. J. Sundberg, T. D. Cung, P. S. Burke, M. P. Martin, L. Preiss, J. Lifson, M. Lichterfeld, M. Carrington, and X. G. Yu.** 2009. HLA-B*35-Px-mediated acceleration of HIV-1 infection by increased inhibitory immunoregulatory impulses. J Exp Med **206:**2959-66.

155. **Huang, K. H., M. R. Loutfy, C. M. Tsoukas, and N. F. Bernard.** 2008. Immune correlates of CD4 decline in HIV-infected patients experiencing virologic failure before undergoing treatment interruption. BMC Infect Dis **8:**59.

156. **Huang, M., and M. A. Martin.** 1997. Incorporation of Pr160(gag-pol) into virus particles requires the presence of both the major homology region and adjacent C-terminal capsid sequences within the Gag-Pol polyprotein. J Virol **71:**4472-8.

157. **Hulot, S. L., E. M. Cale, B. T. Korber, and N. L. Letvin.** 2011. Vaccine-induced CD8+ T lymphocytes of rhesus monkeys recognize variant forms of an HIV epitope but do not mediate optimal functional activity. J Immunol **186:**5663-74.

158. **Hwang, S. S., T. J. Boyle, H. K. Lyerly, and B. R. Cullen.** 1991. Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1. Science **253:**71-4.

159. **Igarashi, T., C. R. Brown, Y. Endo, A. Buckler-White, R. Plishka, N. Bischofberger, V. Hirsch, and M. A. Martin.** 2001. Macrophage are the principal reservoir and sustain high virus loads in rhesus macaques after the depletion of CD4+ T cells by a highly pathogenic simian immunodeficiency virus/HIV type 1 chimera (SHIV): Implications for HIV-1 infections of humans. Proc Natl Acad Sci U S A **98:**658-63.

160. **International, H. I. V. C. S., F. Pereyra, X. Jia, P. J. McLaren, A. Telenti, P. I. de Bakker, B. D. Walker, S. Ripke, C. J. Brumme, S. L. Pulit, M. Carrington, C. M. Kadie, J. M. Carlson, D. Heckerman, R. R. Graham, R. M. Plenge, S. G. Deeks,**

L. Gianniny, G. Crawford, J. Sullivan, E. Gonzalez, L. Davies, A. Camargo, J. M. Moore, N. Beattie, S. Gupta, A. Crenshaw, N. P. Burtt, C. Guiducci, N. Gupta, X. Gao, Y. Qi, Y. Yuki, A. Piechocka-Trocha, E. Cutrell, R. Rosenberg, K. L. Moss, P. Lemay, J. O'Leary, T. Schaefer, P. Verma, I. Toth, B. Block, B. Baker, A. Rothchild, J. Lian, J. Proudfoot, D. M. Alvino, S. Vine, M. M. Addo, T. M. Allen, M. Altfeld, M. R. Henn, S. Le Gall, H. Streeck, D. W. Haas, D. R. Kuritzkes, G. K. Robbins, R. W. Shafer, R. M. Gulick, C. M. Shikuma, R. Haubrich, S. Riddler, P. E. Sax, E. S. Daar, H. J. Ribaudo, B. Agan, S. Agarwal, R. L. Ahern, B. L. Allen, S. Altidor, E. L. Altschuler, S. Ambardar, K. Anastos, B. Anderson, V. Anderson, U. Andrady, D. Antoniskis, D. Bangsberg, D. Barbaro, W. Barrie, J. Bartczak, S. Barton, P. Basden, N. Basgoz, S. Bazner, N. C. Bellos, A. M. Benson, J. Berger, N. F. Bernard, A. M. Bernard, C. Birch, S. J. Bodner, R. K. Bolan, E. T. Boudreaux, M. Bradley, J. F. Braun, J. E. Brndjar, S. J. Brown, K. Brown, et al. 2010. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. Science **330:**1551-7.

161. **Iweriebor, B. C., L. G. Mavhandu, T. Masebe, D. Rekosh, M. L. Hammarskjold, J. M. Mphahlele, and P. O. Bessong.** 2011. Molecular epidemiology of HIV in two highly endemic areas of northeastern South Africa. Arch Virol.

162. **Jaffray, A., E. Shephard, J. van Harmelen, C. Williamson, A. L. Williamson, and E. P. Rybicki.** 2004. Human immunodeficiency virus type 1 subtype C Gag virus-like particle boost substantially improves the immune response to a subtype C gag DNA vaccine in mice. J Gen Virol **85:**409-13.

163. **Jansen, C. A., S. Kostense, K. Vandenberghe, N. M. Nanlohy, I. M. De Cuyper, E. Piriou, E. H. Manting, F. Miedema, and D. van Baarle.** 2005. High responsiveness of HLA-B57-restricted Gag-specific CD8+ T cells in vitro may contribute to the protective effect of HLA-B57 in HIV-infection. Eur J Immunol **35:**150-8.

164. **Janssen, E. M., E. E. Lemmens, T. Wolfe, U. Christen, M. G. von Herrath, and S. P. Schoenberger.** 2003. CD4+ T cells are required for secondary expansion and memory in CD8+ T lymphocytes. Nature **421:**852-6.

165. **Janssens, W., L. Heyndrickx, K. Fransen, M. Temmerman, A. Leonaers, T. Ivens, J. Motte, P. Piot, and G. Van der Groen.** 1994. Genetic variability of HIV type 1 in Kenya. AIDS Res Hum Retroviruses **10:**1577-9.

166. **Jaoko, W., F. N. Nakwagala, O. Anzala, G. O. Manyonyi, J. Birungi, A. Nanvubya, F. Bashir, K. Bhatt, H. Ogutu, S. Wakasiaka, L. Matu, W. Waruingi, J. Odada, M. Oyaro, J. Indangasi, J. Ndinya-Achola, C. Konde, E. Mugisha, P. Fast, C. Schmidt, J. Gilmour, T. Tarragona, C. Smith, B. Barin, L. Dally, B. Johnson, A. Muluubya, L. Nielsen, P. Hayes, M. Boaz, P. Hughes, T. Hanke, A. McMichael, J. Bwayo, and P. Kaleebu.** 2008. Safety and immunogenicity of recombinant low-dosage HIV-1 A vaccine candidates vectored by plasmid pTHr DNA or modified vaccinia virus Ankara (MVA) in humans in East Africa. Vaccine **26:**2788-95.

167. **Jin, X., D. E. Bauer, S. E. Tuttleton, S. Lewin, A. Gettie, J. Blanchard, C. E. Irwin, J. T. Safrit, J. Mittler, L. Weinberger, L. G. Kostrikis, L. Zhang, A. S. Perelson, and D. D. Ho.** 1999. Dramatic rise in plasma viremia after CD8(+) T cell depletion in simian immunodeficiency virus-infected macaques. J Exp Med **189:**991-8.

168. **John, M., D. Heckerman, I. James, L. P. Park, J. M. Carlson, A. Chopra, S. Gaudieri, D. Nolan, D. W. Haas, S. A. Riddler, R. Haubrich, and S. Mallal.** 2010. Adaptive interactions between HLA and HIV-1: highly divergent selection imposed by HLA class I molecules with common supertype motifs. J Immunol **184:**4368-77.

169. **Johnson, W. E., J. Morgan, J. Reitter, B. A. Puffer, S. Czajak, R. W. Doms, and R. C. Desrosiers.** 2002. A replication-competent, neutralization-sensitive variant of simian immunodeficiency virus lacking 100 amino acids of envelope. J Virol **76:**2075-86.

170. **Johnston, M. I., and A. S. Fauci.** 2007. An HIV vaccine--evolving concepts. N Engl J Med **356:**2073-81.

171. **Jones, K. A.** 1997. Taking a new TAK on tat transactivation. Genes Dev **11:**2593-9.

172. **Jowett, J. B., V. Planelles, B. Poon, N. P. Shah, M. L. Chen, and I. S. Chen.** 1995. The human immunodeficiency virus type 1 vpr gene arrests infected T cells in the G2 + M phase of the cell cycle. J Virol **69:**6304-13.

173. **Kangueane, P., M. K. Sakharkar, G. Rajaseger, S. Bolisetty, B. Sivasekari, B. Zhao, M. Ravichandran, P. Shapshak, and S. Subbiah.** 2005. A framework to sub-type HLA supertypes. Front Biosci **10:**879-86.

174. **Kannanganat, S., C. Ibegbu, L. Chennareddi, H. L. Robinson, and R. R. Amara.** 2007. Multiple-cytokine-producing antiviral CD4 T cells are functionally superior to single-cytokine-producing cells. J Virol **81:**8468-76.

175. **Kannanganat, S., B. G. Kapogiannis, C. Ibegbu, L. Chennareddi, P. Goepfert, H. L. Robinson, J. Lennox, and R. R. Amara.** 2007. Human immunodeficiency virus type 1 controllers but not noncontrollers maintain CD4 T cells coexpressing three cytokines. J Virol **81:**12071-6.

176. **Kao, S. Y., A. F. Calman, P. A. Luciw, and B. M. Peterlin.** 1987. Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. Nature **330:**489-93.

177. **Kaul, R., T. Dong, F. A. Plummer, J. Kimani, T. Rostron, P. Kiama, E. Njagi, E. Irungu, B. Farah, J. Oyugi, R. Chakraborty, K. S. MacDonald, J. J. Bwayo, A. McMichael, and S. L. Rowland-Jones.** 2001. CD8(+) lymphocytes respond to different HIV epitopes in seronegative and infected subjects. J Clin Invest **107:**1303-10.

178. **Kawashima, Y., K. Pfafferott, J. Frater, P. Matthews, R. Payne, M. Addo, H. Gatanaga, M. Fujiwara, A. Hachiya, H. Koizumi, N. Kuse, S. Oka, A. Duda, A.**

Prendergast, H. Crawford, A. Leslie, Z. Brumme, C. Brumme, T. Allen, C. Brander, R. Kaslow, J. Tang, E. Hunter, S. Allen, J. Mulenga, S. Branch, T. Roach, M. John, S. Mallal, A. Ogwu, R. Shapiro, J. G. Prado, S. Fidler, J. Weber, O. G. Pybus, P. Klenerman, T. Ndung'u, R. Phillips, D. Heckerman, P. R. Harrigan, B. D. Walker, M. Takiguchi, and P. Goulder. 2009. Adaptation of HIV-1 to human leukocyte antigen class I. Nature **458:**641-5.

179. **Keane, N. M., S. G. Roberts, C. A. Almeida, T. Krishnan, A. Chopra, E. Demaine, R. Laird, M. Tschochner, J. M. Carlson, S. Mallal, D. Heckerman, I. James, and M. John.** 2011. High-avidity, high-IFNgamma-producing CD8 T-cell responses following immune selection during HIV-1 infection. Immunol Cell Biol.

180. **Kearney, M., F. Maldarelli, W. Shao, J. B. Margolick, E. S. Daar, J. W. Mellors, V. Rao, J. M. Coffin, and S. Palmer.** 2009. Human immunodeficiency virus type 1 population genetics and adaptation in newly infected individuals. J Virol **83:**2715-27.

181. **Keating, S. M., R. C. Bollinger, T. C. Quinn, J. B. Jackson, and L. M. Carruth.** 2002. Cross-clade T lymphocyte-mediated immunity to HIV type 1: implications for vaccine design and immunodetection assays. AIDS Res Hum Retroviruses **18:**1067-79.

182. **Kent, S. J., C. S. Fernandez, C. J. Dale, and M. P. Davenport.** 2005. Reversion of immune escape HIV variants upon transmission: insights into effective viral immunity. Trends Microbiol **13:**243-6.

183. **Kiepiela, P., A. J. Leslie, I. Honeyborne, D. Ramduth, C. Thobakgale, S. Chetty, P. Rathnavalu, C. Moore, K. J. Pfafferott, L. Hilton, P. Zimbwa, S. Moore, T. Allen, C. Brander, M. M. Addo, M. Altfeld, I. James, S. Mallal, M. Bunce, L. D. Barber, J. Szinger, C. Day, P. Klenerman, J. Mullins, B. Korber, H. M. Coovadia, B. D. Walker, and P. J. Goulder.** 2004. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. Nature **432:**769-75.

184. **Kiepiela, P., K. Ngumbela, C. Thobakgale, D. Ramduth, I. Honeyborne, E. Moodley, S. Reddy, C. de Pierres, Z. Mncube, N. Mkhwanazi, K. Bishop, M. van der Stok, K. Nair, N. Khan, H. Crawford, R. Payne, A. Leslie, J. Prado, A. Prendergast, J. Frater, N. McCarthy, C. Brander, G. H. Learn, D. Nickle, C. Rousseau, H. Coovadia, J. I. Mullins, D. Heckerman, B. D. Walker, and P. Goulder.** 2007. CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. Nat Med **13:**46-53.

185. **Kijak, G. H., E. Sanders-Buell, N. D. Wolfe, E. Mpoudi-Ngole, B. Kim, B. Brown, M. L. Robb, D. L. Birx, D. S. Burke, J. K. Carr, and F. E. McCutchan.** 2004. Development and application of a high-throughput HIV type 1 genotyping assay to identify CRF02_AG in West/West Central Africa. AIDS Res Hum Retroviruses **20:**521-30.

186. **Kim, S., K. Ikeuchi, R. Byrn, J. Groopman, and D. Baltimore.** 1989. Lack of a negative influence on viral growth by the nef gene of human immunodeficiency virus type 1. Proc Natl Acad Sci U S A **86:**9544-8.

187. **Kirchhoff, F.** 2010. Immune evasion and counteraction of restriction factors by HIV-1 and other primate lentiviruses. Cell Host Microbe **8:**55-67.

188. **Kirchhoff, F., M. Schindler, A. Specht, N. Arhel, and J. Munch.** 2008. Role of Nef in primate lentiviral immunopathogenesis. Cell Mol Life Sci **65:**2621-36.

189. **Kong, W. P., L. Wu, T. C. Wallstrom, W. Fischer, Z. Y. Yang, S. Y. Ko, N. L. Letvin, B. F. Haynes, B. H. Hahn, B. Korber, and G. J. Nabel.** 2009. Expanded breadth of the T-cell response to mosaic human immunodeficiency virus type 1 envelope DNA vaccination. J Virol **83:**2201-15.

190. **Korber, B., B. Gaschen, K. Yusim, R. Thakallapally, C. Kesmir, and V. Detours.** 2001. Evolutionary and immunological implications of contemporary HIV-1 variation. Br Med Bull **58:**19-42.

191. **Kothe, D. L., Y. Li, J. M. Decker, F. Bibollet-Ruche, K. P. Zammit, M. G. Salazar, Y. Chen, Z. Weng, E. A. Weaver, F. Gao, B. F. Haynes, G. M. Shaw, B. T. Korber, and B. H. Hahn.** 2006. Ancestral and consensus envelope immunogens for HIV-1 subtype C. Virology **352:**438-49.

192. **Koup, R. A.** 1994. Virus escape from CTL recognition. J Exp Med **180:**779-82.

193. **Koup, R. A., J. T. Safrit, Y. Cao, C. A. Andrews, G. McLeod, W. Borkowsky, C. Farthing, and D. D. Ho.** 1994. Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. J Virol **68:**4650-5.

194. **Kresge, K. J.** 2009. South African AIDS Vaccine Initiative launches phase I trial. IAVI Rep **13:**19.

195. **Kuiken, C., Foley, B., Hahn, B., Korber, B., Marx, PA., McCutchan, F., Mellors, JW., Mullins, JI., Sodroski, J., Wolinsky, S.** 2000. HIV Sequence Compendium 2000, vol. 01. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM, LA-UR.

196. **Kuiken, C., Foley, B., Leitner, T., Apetrei, C., Hahn, B., Mizrachi, I., Mullins, J., Rambaut, A., Wolinsky, S., Korber, B.** 2010. HIV Sequence Compendium 2010, vol. 10-03684. Published by Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, NM, LA-UR.

197. **Landau, N. R., M. Warton, and D. R. Littman.** 1988. The envelope glycoprotein of the human immunodeficiency virus binds to the immunoglobulin-like domain of CD4. Nature **334:**159-62.

198. **Lazaro, A. M., Y. Xiao, J. Henry, J. Ng, C. K. Hurley, and P. E. Posch.** 2011. Ninety-six novel HLA class I and II alleles identified in volunteers for the National Marrow Donor Program Registry in 2009. Tissue Antigens **78:**195-202.

199. **Lazaro, A. M., Y. Xiao, C. Masaberg, W. Y. Hwang, A. E. Yeoh, Y. Weiyan, J. Ng, C. K. Hurley, and P. E. Posch.** 2011. Novel HLA class I and II alleles identified during routine registry typing in 2010. Tissue Antigens **78:**263-6.

200. **Lazaro, A. M., Y. Xiao, C. Masaberg, B. Tu, J. Ng, C. K. Hurley, and P. E. Posch.** 2011. Seventy-eight novel HLA class I and II alleles identified during routine registry typing in 2008 and 2009. Tissue Antigens **77:**54-61.

201. **Lazaro, E., C. Kadie, P. Stamegna, S. C. Zhang, P. Gourdain, N. Y. Lai, M. Zhang, S. A. Martinez, D. Heckerman, and S. Le Gall.** 2011. Variable HIV peptide stability in human cytosol is critical to epitope presentation and immune escape. J Clin Invest **121:**2480-92.

202. **Lazaro, E., L. T. Tram, P. Bellecave, G. L. Guidicelli, G. Anies, H. H. Thu, M. P. Debelleix, M. Vray, P. Recordon-Pinson, J. L. Taupin, T. T. Lien, and H. Fleury.** 2011. Molecular characterization of HIV-1 CRF01_AE in Mekong Delta, Vietnam, and impact of T-cell epitope mutations on HLA recognition (ANRS 12159). PLoS One **6:**e26244.

203. **Le Vu, S., Y. Le Strat, F. Barin, J. Pillonel, F. Cazein, V. Bousquet, S. Brunet, D. Thierry, C. Semaille, L. Meyer, and J. C. Desenclos.** 2010. Population-based HIV-1 incidence in France, 2003-08: a modelling analysis. Lancet Infect Dis **10:**682-7.

204. **Leslie, A. J., K. J. Pfafferott, P. Chetty, R. Draenert, M. M. Addo, M. Feeney, Y. Tang, E. C. Holmes, T. Allen, J. G. Prado, M. Altfeld, C. Brander, C. Dixon, D. Ramduth, P. Jeena, S. A. Thomas, A. St John, T. A. Roach, B. Kupfer, G. Luzzi, A. Edwards, G. Taylor, H. Lyall, G. Tudor-Williams, V. Novelli, J. Martinez-Picado, P. Kiepiela, B. D. Walker, and P. J. Goulder.** 2004. HIV evolution: CTL escape mutation and reversion after transmission. Nat Med **10:**282-9.

205. **Letourneau, S., E. J. Im, T. Mashishi, C. Brereton, A. Bridgeman, H. Yang, L. Dorrell, T. Dong, B. Korber, A. J. McMichael, and T. Hanke.** 2007. Design and pre-clinical evaluation of a universal HIV-1 vaccine. PLoS One **2:**e984.

206. **Li, F., A. C. Finnefrock, S. A. Dubey, B. T. Korber, J. Szinger, S. Cole, M. J. McElrath, J. W. Shiver, D. R. Casimiro, L. Corey, and S. G. Self.** 2011. Mapping HIV-1 vaccine induced T-cell responses: bias towards less-conserved regions and potential impact on vaccine efficacy in the Step study. PLoS One **6:**e20479.

207. **Li, Y., S. O'Dell, L. M. Walker, X. Wu, J. Guenaga, Y. Feng, S. D. Schmidt, K. McKee, M. K. Louder, J. E. Ledgerwood, B. S. Graham, B. F. Haynes, D. R. Burton, R. T. Wyatt, and J. R. Mascola.** 2011. Mechanism of neutralization by the broadly neutralizing HIV-1 monoclonal antibody VRC01. J Virol **85:**8954-67.

208. **Liang, C., J. Hu, R. S. Russell, A. Roldan, L. Kleiman, and M. A. Wainberg.** 2002. Characterization of a putative alpha-helix across the capsid-SP1 boundary that is critical for the multimerization of human immunodeficiency virus type 1 gag. J Virol **76:**11729-37.

209. **Lichterfeld, M., D. E. Kaufmann, X. G. Yu, S. K. Mui, M. M. Addo, M. N. Johnston, D. Cohen, G. K. Robbins, E. Pae, G. Alter, A. Wurcel, D. Stone, E. S. Rosenberg, B. D. Walker, and M. Altfeld.** 2004. Loss of HIV-1-specific CD8+ T cell proliferation after acute HIV-1 infection and restoration by vaccine-induced HIV-1-specific CD4+ T cells. J Exp Med **200:**701-12.

210. **Lifson, J. D., J. L. Rossio, M. Piatak, Jr., T. Parks, L. Li, R. Kiser, V. Coalter, B. Fisher, B. M. Flynn, S. Czajak, V. M. Hirsch, K. A. Reimann, J. E. Schmitz, J. Ghrayeb, N. Bischofberger, M. A. Nowak, R. C. Desrosiers, and D. Wodarz.** 2001. Role of CD8(+) lymphocytes in control of simian immunodeficiency virus infection and resistance to rechallenge after transient early antiretroviral treatment. J Virol **75:**10187-99.

211. **Limou, S., S. Le Clerc, C. Coulonges, W. Carpentier, C. Dina, O. Delaneau, T. Labib, L. Taing, R. Sladek, C. Deveau, R. Ratsimandresey, M. Montes, J. L. Spadoni, J. D. Lelievre, Y. Levy, A. Therwath, F. Schachter, F. Matsuda, I. Gut, P. Froguel, J. F. Delfraissy, S. Hercberg, J. F. Zagury, and A. G. Group.** 2009. Genomewide association study of an AIDS-nonprogression cohort emphasizes the role played by HLA genes (ANRS Genomewide Association Study 02). J Infect Dis **199:**419-26.

212. **Lindwasser, O. W., R. Chaudhuri, and J. S. Bonifacino.** 2007. Mechanisms of CD4 downregulation by the Nef and Vpu proteins of primate immunodeficiency viruses. Curr Mol Med **7:**171-84.

213. **Liu, Y., J. McNevin, H. Zhao, D. M. Tebit, R. M. Troyer, M. McSweyn, A. K. Ghosh, D. Shriner, E. J. Arts, M. J. McElrath, and J. I. Mullins.** 2007. Evolution of human immunodeficiency virus type 1 cytotoxic T-lymphocyte epitopes: fitness-balanced escape. J Virol **81:**12179-88.

214. **Loffredo, J. T., T. C. Friedrich, E. J. Leon, J. J. Stephany, D. S. Rodrigues, S. P. Spencer, A. T. Bean, D. R. Beal, B. J. Burwitz, R. A. Rudersdorf, L. T. Wallace, S. M. Piaskowski, G. E. May, J. Sidney, E. Gostick, N. A. Wilson, D. A. Price, E. G. Kallas, H. Piontkivska, A. L. Hughes, A. Sette, and D. I. Watkins.** 2007. CD8+ T cells from SIV elite controller macaques recognize Mamu-B*08-bound epitopes and select for widespread viral variation. PLoS One **2:**e1152.

215. **Lopez, C. S., J. D. Eccles, A. Still, R. E. Sloan, R. L. Barklis, S. M. Tsagli, and E. Barklis.** 2011. Determinants of the HIV-1 core assembly pathway. Virology **417:**137-46.

216. **Louwagie, J., F. E. McCutchan, M. Peeters, T. P. Brennan, E. Sanders-Buell, G. A. Eddy, G. van der Groen, K. Fransen, G. M. Gershy-Damet, R. Deleys, and et al.** 1993. Phylogenetic analysis of gag genes from 70 international HIV-1 isolates provides evidence for multiple genotypes. AIDS **7:**769-80.

217. **Magadan, J. G., and J. S. Bonifacino.** 2012. Transmembrane domain determinants of CD4 Downregulation by HIV-1 Vpu. J Virol **86:**757-72.

218. **Makedonas, G., N. Hutnick, D. Haney, A. C. Amick, J. Gardner, G. Cosma, A. R. Hersperger, D. Dolfi, E. J. Wherry, G. Ferrari, and M. R. Betts.** 2010. Perforin and IL-2 upregulation define qualitative differences among highly functional virus-specific human CD8 T cells. PLoS Pathog **6:**e1000798.

219. **Malhotra, U., F. Li, J. Nolin, M. Allison, H. Zhao, J. I. Mullins, S. Self, and M. J. McElrath.** 2007. Enhanced detection of human immunodeficiency virus type 1 (HIV-

1) Nef-specific T cells recognizing multiple variants in early HIV-1 infection. J Virol **81:**5225-37.

220. **Malhotra, U., J. Nolin, H. Horton, F. Li, L. Corey, J. I. Mullins, and M. J. McElrath.** 2009. Functional properties and epitope characteristics of T-cells recognizing natural HIV-1 variants. Vaccine **27:**6678-87.

221. **Malhotra, U., J. Nolin, J. I. Mullins, and M. J. McElrath.** 2007. Comprehensive epitope analysis of cross-clade Gag-specific T-cell responses in individuals with early HIV-1 infection in the US epidemic. Vaccine **25:**381-90.

222. **Malim, M. H., J. Hauber, S. Y. Le, J. V. Maizel, and B. R. Cullen.** 1989. The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. Nature **338:**254-7.

223. **Mangeat, B., G. Gers-Huber, M. Lehmann, M. Zufferey, J. Luban, and V. Piguet.** 2009. HIV-1 Vpu neutralizes the antiviral factor Tetherin/BST-2 by binding it and directing its beta-TrCP2-dependent degradation. PLoS Pathog **5:**e1000574.

224. **Martinez-Picado, J., J. G. Prado, E. E. Fry, K. Pfafferott, A. Leslie, S. Chetty, C. Thobakgale, I. Honeyborne, H. Crawford, P. Matthews, T. Pillay, C. Rousseau, J. I. Mullins, C. Brander, B. D. Walker, D. I. Stuart, P. Kiepiela, and P. Goulder.** 2006. Fitness cost of escape mutations in p24 Gag in association with control of human immunodeficiency virus type 1. J Virol **80:**3617-23.

225. **Masemola, A., T. Mashishi, G. Khoury, P. Mohube, P. Mokgotho, E. Vardas, M. Colvin, L. Zijenah, D. Katzenstein, R. Musonda, S. Allen, N. Kumwenda, T. Taha, G. Gray, J. McIntyre, S. A. Karim, H. W. Sheppard, C. M. Gray, and H. S. Team.** 2004. Hierarchical targeting of subtype C human immunodeficiency virus type 1 proteins by CD8+ T cells: correlation with viral load. J Virol **78:**3233-43.

226. **McAdam, S., P. Kaleebu, P. Krausa, P. Goulder, N. French, B. Collin, T. Blanchard, J. Whitworth, A. McMichael, and F. Gotch.** 1998. Cross-clade recognition of p55 by cytotoxic T lymphocytes in HIV-1 infection. AIDS **12:**571-9.

227. **McAdam, S., P. Klenerman, L. Tussey, S. Rowland-Jones, D. Lalloo, R. Phillips, A. Edwards, P. Giangrande, A. L. Brown, F. Gotch, and et al.** 1995. Immunogenic HIV variant peptides that bind to HLA-B8 can fail to stimulate cytotoxic T lymphocyte responses. J Immunol **155:**2729-36.

228. **McCutchan, F. E.** 2006. Global epidemiology of HIV. J Med Virol **78 Suppl 1:**S7-S12.

229. **McCutchan, F. E.** 2000. Understanding the genetic diversity of HIV-1. AIDS **14 Suppl 3:**S31-44.

230. **McCutchan, F. E., J. K. Carr, M. Bajani, E. Sanders-Buell, T. O. Harry, T. C. Stoeckli, K. E. Robbins, W. Gashau, A. Nasidi, W. Janssens, and M. L. Kalish.** 1999. Subtype G and multiple forms of A/G intersubtype recombinant human immunodeficiency virus type 1 in Nigeria. Virology **254:**226-34.

231. **McCutchan, F. E., P. A. Hegerich, T. P. Brennan, P. Phanuphak, P. Singharaj, A. Jugsudee, P. W. Berman, A. M. Gray, A. K. Fowler, and D. S. Burke.** 1992. Genetic variants of HIV-1 in Thailand. AIDS Res Hum Retroviruses **8:**1887-95.

232. **McCutchan, F. E., B. L. Ungar, P. Hegerich, C. R. Roberts, A. K. Fowler, S. K. Hira, P. L. Perine, and D. S. Burke.** 1992. Genetic analysis of HIV-1 isolates from Zambia and an expanded phylogenetic tree for HIV-1. J Acquir Immune Defic Syndr **5:**441-9.

233. **McDonald, A.** 2000. HIV infection, AIDS, hepatitis C, and sexually transmissible infections in Australia: National surveillance results to December 1998. N S W Public Health Bull **11:**58-60.

234. **McKay, P. F., D. H. Barouch, J. E. Schmitz, R. S. Veazey, D. A. Gorgone, M. A. Lifton, K. C. Williams, and N. L. Letvin.** 2003. Global dysfunction of CD4 T-lymphocyte cytokine expression in simian-human immunodeficiency virus/SIV-infected monkeys is prevented by vaccination. J Virol **77:**4695-702.

235. **McKinney, D. M., R. Skvoretz, B. D. Livingston, C. C. Wilson, M. Anders, R. W. Chesnut, A. Sette, M. Essex, V. Novitsky, and M. J. Newman.** 2004. Recognition of variant HIV-1 epitopes from diverse viral subtypes by vaccine-induced CTL. J Immunol **173:**1941-50.

236. **McKinnon, L. R., T. B. Ball, J. Kimani, C. Wachihi, L. Matu, M. Luo, J. Embree, K. R. Fowke, and F. A. Plummer.** 2005. Cross-clade CD8(+) T-cell responses with a preference for the predominant circulating clade. J Acquir Immune Defic Syndr **40:**245-9.

237. **McKinnon, L. R., T. B. Ball, C. Wachihi, P. J. McLaren, J. L. Waruk, X. Mao, S. Ramdahin, A. O. Anzala, J. Kamene, M. Luo, K. R. Fowke, and F. A. Plummer.** 2007. Epitope cross-reactivity frequently differs between central and effector memory HIV-specific CD8+ T cells. J Immunol **178:**3750-6.

238. **McKinnon, L. R., R. Kaul, J. Kimani, N. J. Nagelkerke, C. Wachihi, K. R. Fowke, T. B. Ball, and F. A. Plummer.** 2011. HIV-specific CD8(+) T-cell proliferation is prospectively associated with delayed disease progression. Immunol Cell Biol.

239. **Mervis, R. J., N. Ahmad, E. P. Lillehoj, M. G. Raum, F. H. Salazar, H. W. Chan, and S. Venkatesan.** 1988. The gag gene products of human immunodeficiency virus type 1: alignment within the gag open reading frame, identification of posttranslational modifications, and evidence for alternative gag precursors. J Virol **62:**3993-4002.

240. **Michael, N. L.** 1999. Host genetic influences on HIV-1 pathogenesis. Curr Opin Immunol **11:**466-74.

241. **Michel, N., I. Allespach, S. Venzke, O. T. Fackler, and O. T. Keppler.** 2005. The Nef protein of human immunodeficiency virus establishes superinfection immunity by a dual strategy to downregulate cell-surface CCR5 and CD4. Curr Biol **15:**714-23.

242. **Michel, N., K. Ganter, S. Venzke, J. Bitzegeio, O. T. Fackler, and O. T. Keppler.** 2006. The Nef protein of human immunodeficiency virus is a broad-spectrum modulator of chemokine receptor cell surface levels that acts independently of classical motifs for receptor endocytosis and Galphai signaling. Mol Biol Cell **17:**3578-90.

243. **Migueles, S. A., A. C. Laborico, H. Imamichi, W. L. Shupert, C. Royce, M. McLaughlin, L. Ehler, J. Metcalf, S. Liu, C. W. Hallahan, and M. Connors.** 2003. The differential ability of HLA B*5701+ long-term nonprogressors and progressors to restrict human immunodeficiency virus replication is not caused by loss of recognition of autologous viral gag sequences. J Virol **77:**6889-98.

244. **Migueles, S. A., A. C. Laborico, W. L. Shupert, M. S. Sabbaghian, R. Rabin, C. W. Hallahan, D. Van Baarle, S. Kostense, F. Miedema, M. McLaughlin, L. Ehler, J. Metcalf, S. Liu, and M. Connors.** 2002. HIV-specific CD8+ T cell proliferation is coupled to perforin expression and is maintained in nonprogressors. Nat Immunol **3:**1061-8.

245. **Migueles, S. A., M. S. Sabbaghian, W. L. Shupert, M. P. Bettinotti, F. M. Marincola, L. Martino, C. W. Hallahan, S. M. Selig, D. Schwartz, J. Sullivan, and M. Connors.** 2000. HLA B*5701 is highly associated with restriction of virus replication in a subgroup of HIV-infected long term nonprogressors. Proc Natl Acad Sci U S A **97:**2709-14.

246. **Migueles, S. A., J. C. Tilton, and M. Connors.** 2004. Advances in understanding immunologic control of HIV infection. Curr HIV/AIDS Rep **1:**12-7.

247. **Migueles, S. A., J. C. Tilton, and M. Connors.** 2006. Qualitative host factors associated with immunological control of HIV infection by CD8 T cells. Curr Opin HIV AIDS **1:**28-33.

248. **Milicic, A., C. T. Edwards, S. Hue, J. Fox, H. Brown, T. Pillay, J. W. Drijfhout, J. N. Weber, E. C. Holmes, S. J. Fidler, H. T. Zhang, and R. E. Phillips.** 2005. Sexual transmission of single human immunodeficiency virus type 1 virions encoding highly polymorphic multisite cytotoxic T-lymphocyte escape variants. J Virol **79:**13953-62.

249. **Miller, R. H., and N. Sarver.** 1997. HIV accessory proteins as therapeutic targets. Nat Med **3:**389-94.

250. **Miura, T., M. A. Brockman, A. Schneidewind, M. Lobritz, F. Pereyra, A. Rathod, B. L. Block, Z. L. Brumme, C. J. Brumme, B. Baker, A. C. Rothchild, B. Li, A. Trocha, E. Cutrell, N. Frahm, C. Brander, I. Toth, E. J. Arts, T. M. Allen, and B. D. Walker.** 2009. HLA-B57/B*5801 human immunodeficiency virus type 1 elite controllers select for rare gag variants associated with reduced viral replication capacity and strong cytotoxic T-lymphocyte [corrected] recognition. J Virol **83:**2743-55.

251. **Moore, C. B., M. John, I. R. James, F. T. Christiansen, C. S. Witt, and S. A. Mallal.** 2002. Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. Science **296:**1439-43.

252. **Mori, M., B. Sriwanthana, N. Wichukchinda, C. Boonthimat, N. Tsuchiya, T. Miura, P. Pathipvanich, K. Ariyoshi, and P. Sawanpanyalert.** 2011. Unique CRF01_AE Gag CTL epitopes associated with lower HIV-viral load and delayed disease progression in a cohort of HIV-infected Thais. PLoS One **6:**e22680.

253. **Mothe, B., A. Llano, J. Ibarrondo, M. Daniels, C. Miranda, J. Zamarreno, V. Bach, R. Zuniga, S. Perez-Alvarez, C. T. Berger, M. C. Puertas, J. Martinez-Picado, M. Rolland, M. Farfan, J. J. Szinger, W. H. Hildebrand, O. O. Yang, V. Sanchez-Merino, C. J. Brumme, Z. L. Brumme, D. Heckerman, T. M. Allen, J. I. Mullins, G. Gomez, P. J. Goulder, B. D. Walker, J. M. Gatell, B. Clotet, B. T. Korber, J. Sanchez, and C. Brander.** 2011. Definition of the viral targets of protective HIV-1-specific T cell responses. J Transl Med **9:**208.

254. **Mothe, B. R., J. Weinfurter, C. Wang, W. Rehrauer, N. Wilson, T. M. Allen, D. B. Allison, and D. I. Watkins.** 2003. Expression of the major histocompatibility complex class I molecule Mamu-A*01 is associated with control of simian immunodeficiency virus SIVmac239 replication. J Virol **77:**2736-40.

255. **Motomura, K., S. Kusagawa, K. Kato, K. Nohtomi, H. H. Lwin, K. M. Tun, M. Thwe, K. Y. Oo, S. Lwin, O. Kyaw, M. Zaw, Y. Nagai, and Y. Takebe.** 2000. Emergence of new forms of human immunodeficiency virus type 1 intersubtype recombinants in central Myanmar. AIDS Res Hum Retroviruses **16:**1831-43.

256. **Muhl, T., M. Krawczak, P. Ten Haaft, G. Hunsmann, and U. Sauermann.** 2002. MHC class I alleles influence set-point viral load and survival time in simian immunodeficiency virus-infected rhesus monkeys. J Immunol **169:**3438-46.

257. **Mullins, J. I., D. C. Nickle, L. Heath, A. G. Rodrigo, and G. H. Learn.** 2004. Immunogen sequence: the fourth tier of AIDS vaccine design. Expert Rev Vaccines **3:**S151-9.

258. **Musyoki, A. M., K. M. Mothapo, J. N. Rakgole, A. Lukhwareni, P. Bessong, G. S. Selabe, C. Williamson, J. Mphahlele, and H. Bredell.** 2011. Genetic charaterisation of HIV before widespread testing of HIV vaccine candidates at a clinical trial site in Pretoria, South Africa. AIDS Res Hum Retroviruses.

259. **Myers, G.** 1994. Molecular investigation of HIV transmission. Ann Intern Med **121:**889-90.

260. **Naderi, H. R., M. Tagliamonte, M. L. Tornesello, M. Ciccozzi, G. Rezza, R. Farid, F. M. Buonaguro, and L. Buonaguro.** 2006. Molecular and phylogenetic analysis of HIV-1 variants circulating among injecting drug users in Mashhad-Iran. Infect Agent Cancer **1:**4.

261. **Navis, M., I. Schellens, D. van Baarle, J. Borghans, P. van Swieten, F. Miedema, N. Kootstra, and H. Schuitemaker.** 2007. Viral replication capacity as a correlate of

HLA B57/B5801-associated nonprogressive HIV-1 infection. J Immunol **179:**3133-43.

262. **Ndhlovu, Z. M., A. Piechocka-Trocha, S. Vine, A. McMullen, K. C. Koofhethile, P. J. Goulder, T. Ndung'u, D. H. Barouch, and B. D. Walker.** 2011. Mosaic HIV-1 Gag antigens can be processed and presented to human HIV-specific CD8+ T cells. J Immunol **186:**6914-24.

263. **Nickle, D. C., M. A. Jensen, G. S. Gottlieb, D. Shriner, G. H. Learn, A. G. Rodrigo, and J. I. Mullins.** 2003. Consensus and ancestral state HIV vaccines. Science **299:**1515-8; author reply 1515-8.

264. **Nielsen, M., C. Lundegaard, P. Worning, S. L. Lauemoller, K. Lamberth, S. Buus, S. Brunak, and O. Lund.** 2003. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Sci **12:**1007-17.

265. **Novitsky, V., U. R. Smith, P. Gilbert, M. F. McLane, P. Chigwedere, C. Williamson, T. Ndung'u, I. Klein, S. Y. Chang, T. Peter, I. Thior, B. T. Foley, S. Gaolekwe, N. Rybak, S. Gaseitsiwe, F. Vannberg, R. Marlink, T. H. Lee, and M. Essex.** 2002. Human immunodeficiency virus type 1 subtype C molecular phylogeny: consensus sequence for an AIDS vaccine design? J Virol **76:**5435-51.

266. **Novitsky, V., R. Wang, L. Margolin, J. Baca, L. Kebaabetswe, R. Rossenkhan, C. Bonney, M. Herzig, D. Nkwe, S. Moyo, R. Musonda, E. Woldegabriel, E. van Widenfelt, J. Makhema, S. Lagakos, and M. Essex.** 2009. Timing constraints of in vivo gag mutations during primary HIV-1 subtype C infection. PLoS One **4:**e7727.

267. **Oliva, R., L. Falcigno, G. D'Auria, M. Dettin, C. Scarinci, A. Pasquato, C. Di Bello, and L. Paolillo.** 2003. Structural investigation of the HIV-1 envelope glycoprotein gp160 cleavage site, 2: relevance of an N-terminal helix. Chembiochem **4:**727-33.

268. **Op de Coul, E. L., M. Prins, M. Cornelissen, A. van der Schoot, F. Boufassa, R. P. Brettle, L. Hernandez-Aguado, V. Schiffer, J. McMenamin, G. Rezza, R. Robertson, R. Zangerle, J. Goudsmit, R. A. Coutinho, V. V. Lukashov, European, and S. Italian Seroconverter.** 2001. Using phylogenetic analysis to trace HIV-1 migration among western European injecting drug users seroconverting from 1984 to 1997. AIDS **15:**257-66.

269. **Opi, S., S. Kao, R. Goila-Gaur, M. A. Khan, E. Miyagi, H. Takeuchi, and K. Strebel.** 2007. Human immunodeficiency virus type 1 Vif inhibits packaging and antiviral activity of a degradation-resistant APOBEC3G variant. J Virol **81:**8236-46.

270. **Osmanov, S., C. Pattou, N. Walker, B. Schwardlander, J. Esparza, W.-U. N. f. H. Isolation, and Characterization.** 2002. Estimated global distribution and regional spread of HIV-1 genetic subtypes in the year 2000. J Acquir Immune Defic Syndr **29:**184-90.

271. **Ou, C. Y., Y. Takebe, C. C. Luo, M. Kalish, W. Auwanit, C. Bandea, N. de la Torre, J. L. Moore, G. Schochetman, S. Yamazaki, and et al.** 1992. Wide

distribution of two subtypes of HIV-1 in Thailand. AIDS Res Hum Retroviruses **8:**1471-2.

272. **Paxton, W., R. I. Connor, and N. R. Landau.** 1993. Incorporation of Vpr into human immunodeficiency virus type 1 virions: requirement for the p6 region of gag and mutational analysis. J Virol **67:**7229-37.

273. **Peeters, M.** 2001. The genetic variability of HIV-1 and its implications. Transfus Clin Biol **8:**222-5.

274. **Pejchal, R., K. J. Doores, L. M. Walker, R. Khayat, P. S. Huang, S. K. Wang, R. L. Stanfield, J. P. Julien, A. Ramos, M. Crispin, R. Depetris, U. Katpally, A. Marozsan, A. Cupo, S. Maloveste, Y. Liu, R. McBride, Y. Ito, R. W. Sanders, C. Ogohara, J. C. Paulson, T. Feizi, C. N. Scanlan, C. H. Wong, J. P. Moore, W. C. Olson, A. B. Ward, P. Poignard, W. R. Schief, D. R. Burton, and I. A. Wilson.** 2011. A potent and broad neutralizing antibody recognizes and penetrates the HIV glycan shield. Science **334:**1097-103.

275. **Pejchal, R., L. M. Walker, R. L. Stanfield, S. K. Phogat, W. C. Koff, P. Poignard, D. R. Burton, and I. A. Wilson.** 2010. Structure and function of broadly reactive antibody PG16 reveal an H3 subdomain that mediates potent neutralization of HIV-1. Proc Natl Acad Sci U S A **107:**11483-8.

276. **Pelak, K., D. B. Goldstein, N. M. Walley, J. Fellay, D. Ge, K. V. Shianna, C. Gumbs, X. Gao, J. M. Maia, K. D. Cronin, S. K. Hussain, M. Carrington, N. L. Michael, A. C. Weintrob, H. I. V. W. G. Infectious Disease Clinical Research Program, A. National Institute of, and H. I. V. A. V. I. Infectious Diseases Center for.** 2010. Host determinants of HIV-1 control in African Americans. J Infect Dis **201:**1141-9.

277. **Peretz, Y., G. Alter, M. P. Boisvert, G. Hatzakis, C. M. Tsoukas, and N. F. Bernard.** 2005. Human immunodeficiency virus (HIV)-specific gamma interferon secretion directed against all expressed HIV genes: relationship to rate of CD4 decline. J Virol **79:**4908-17.

278. **Pereyra, F., M. M. Addo, D. E. Kaufmann, Y. Liu, T. Miura, A. Rathod, B. Baker, A. Trocha, R. Rosenberg, E. Mackey, P. Ueda, Z. Lu, D. Cohen, T. Wrin, C. J. Petropoulos, E. S. Rosenberg, and B. D. Walker.** 2008. Genetic and immunologic heterogeneity among persons who control HIV infection in the absence of therapy. J Infect Dis **197:**563-71.

279. **Perrin, L., L. Kaiser, and S. Yerly.** 2003. Travel and the spread of HIV-1 genetic variants. Lancet Infect Dis **3:**22-7.

280. **Peters, B. S., W. Jaoko, E. Vardas, G. Panayotakopoulos, P. Fast, C. Schmidt, J. Gilmour, M. Bogoshi, G. Omosa-Manyonyi, L. Dally, L. Klavinskis, B. Farah, T. Tarragona, P. A. Bart, A. Robinson, C. Pieterse, W. Stevens, R. Thomas, B. Barin, A. J. McMichael, J. A. McIntyre, G. Pantaleo, T. Hanke, and J. Bwayo.** 2007. Studies of a prophylactic HIV-1 vaccine candidate based on modified vaccinia

virus Ankara (MVA) with and without DNA priming: effects of dosage and route on safety and immunogenicity. Vaccine **25:**2120-7.

281. **Pitisuttithum, P., P. Gilbert, M. Gurwith, W. Heyward, M. Martin, F. van Griensven, D. Hu, J. W. Tappero, K. Choopanya, and G. Bangkok Vaccine Evaluation.** 2006. Randomized, double-blind, placebo-controlled efficacy trial of a bivalent recombinant glycoprotein 120 HIV-1 vaccine among injection drug users in Bangkok, Thailand. J Infect Dis **194:**1661-71.

282. **Piyasirisilp, S., F. E. McCutchan, J. K. Carr, E. Sanders-Buell, W. Liu, J. Chen, R. Wagner, H. Wolf, Y. Shao, S. Lai, C. Beyrer, and X. F. Yu.** 2000. A recent outbreak of human immunodeficiency virus type 1 infection in southern China was initiated by two highly homogeneous, geographically separated strains, circulating recombinant form AE and a novel BC recombinant. J Virol **74:**11286-95.

283. **Planelles, V., F. Bachelerie, J. B. Jowett, A. Haislip, Y. Xie, P. Banooni, T. Masuda, and I. S. Chen.** 1995. Fate of the human immunodeficiency virus type 1 provirus in infected cells: a role for vpr. J Virol **69:**5883-9.

284. **Plantier, J. C., M. Leoz, J. E. Dickerson, F. De Oliveira, F. Cordonnier, V. Lemee, F. Damond, D. L. Robertson, and F. Simon.** 2009. A new human immunodeficiency virus derived from gorillas. Nat Med **15:**871-2.

285. **Pond, S. L., S. D. Frost, Z. Grossman, M. B. Gravenor, D. D. Richman, and A. J. Brown.** 2006. Adaptation to different human populations by HIV-1 revealed by codon-based analyses. PLoS Comput Biol **2:**e62.

286. **Popov, S., M. Rexach, L. Ratner, G. Blobel, and M. Bukrinsky.** 1998. Viral protein R regulates docking of the HIV-1 preintegration complex to the nuclear pore complex. J Biol Chem **273:**13347-52.

287. **Price, D. A., P. J. Goulder, P. Klenerman, A. K. Sewell, P. J. Easterbrook, M. Troop, C. R. Bangham, and R. E. Phillips.** 1997. Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection. Proc Natl Acad Sci U S A **94:**1890-5.

288. **Qi, M., R. Yang, and C. Aiken.** 2008. Cyclophilin A-dependent restriction of human immunodeficiency virus type 1 capsid mutants for infection of nondividing cells. J Virol **82:**12001-8.

289. **Qing, M., T. Li, Y. Han, Z. Qiu, and Y. Jiao.** 2006. Accelerating effect of human leukocyte antigen-Bw6 homozygosity on disease progression in Chinese HIV-1-infected patients. J Acquir Immune Defic Syndr **41:**137-9.

290. **Ragupathy, V., J. Zhao, O. Wood, S. Tang, S. Lee, P. Nyambi, and I. Hewlett.** 2011. Identification of new, emerging HIV-1 unique recombinant forms and drug resistant viruses circulating in Cameroon. Virol J **8:**185.

291. **Rayfield, M. A., P. Sullivan, C. I. Bandea, L. Britvan, R. A. Otten, C. P. Pau, D. Pieniazek, S. Subbarao, P. Simon, C. A. Schable, A. C. Wright, J. Ward, and G.**

Schochetman. 1996. HIV-1 group O virus identified for the first time in the United States. Emerg Infect Dis **2:**209-12.

292. **Rerks-Ngarm, S., A. E. Brown, C. Khamboonruang, P. Thongcharoen, and P. Kunasol.** 2006. HIV/AIDS preventive vaccine 'prime-boost' phase III trial: foundations and initial lessons learned from Thailand. AIDS **20:**1471-9.

293. **Rerks-Ngarm, S., P. Pitisuttithum, S. Nitayaphan, J. Kaewkungwal, J. Chiu, R. Paris, N. Premsri, C. Namwat, M. de Souza, E. Adams, M. Benenson, S. Gurunathan, J. Tartaglia, J. G. McNeil, D. P. Francis, D. Stablein, D. L. Birx, S. Chunsuttiwat, C. Khamboonruang, P. Thongcharoen, M. L. Robb, N. L. Michael, P. Kunasol, J. H. Kim, and M.-T. Investigators.** 2009. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. N Engl J Med **361:**2209-20.

294. **Rhodes, D. I., L. Ashton, A. Solomon, A. Carr, D. Cooper, J. Kaldor, and N. Deacon.** 2000. Characterization of three nef-defective human immunodeficiency virus type 1 strains associated with long-term nonprogression. Australian Long-Term Nonprogressor Study Group. J Virol **74:**10581-8.

295. **Richmond, M., L. R. McKinnon, S. A. Kiazyk, C. Wachihi, M. Kimani, J. Kimani, F. A. Plummer, and T. B. Ball.** 2011. Epitope mapping of HIV-specific CD8+ T cell responses by multiple immunological readouts reveals distinct specificities defined by function. J Virol **85:**1275-86.

296. **Roberts, J. D., K. Bebenek, and T. A. Kunkel.** 1988. The accuracy of reverse transcriptase from HIV-1. Science **242:**1171-3.

297. **Robertson, D. L., J. P. Anderson, J. A. Bradac, J. K. Carr, B. Foley, R. K. Funkhouser, F. Gao, B. H. Hahn, M. L. Kalish, C. Kuiken, G. H. Learn, T. Leitner, F. McCutchan, S. Osmanov, M. Peeters, D. Pieniazek, M. Salminen, P. M. Sharp, S. Wolinsky, and B. Korber.** 2000. HIV-1 nomenclature proposal. Science **288:**55-6.

298. **Robinson, J., and S. G. Marsh.** 2003. HLA informatics. Accessing HLA sequences from sequence databases. Methods Mol Biol **210:**3-21.

299. **Robinson, J., M. J. Waller, S. C. Fail, H. McWilliam, R. Lopez, P. Parham, and S. G. Marsh.** 2009. The IMGT/HLA database. Nucleic Acids Res **37:**D1013-7.

300. **Robinson, J., M. J. Waller, P. Parham, N. de Groot, R. Bontrop, L. J. Kennedy, P. Stoehr, and S. G. Marsh.** 2003. IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. Nucleic Acids Res **31:**311-4.

301. **Roederer, M., J. M. Brenchley, M. R. Betts, and S. C. De Rosa.** 2004. Flow cytometric analysis of vaccine responses: how many colors are enough? Clin Immunol **110:**199-205.

302. **Rolland, M., D. C. Nickle, and J. I. Mullins.** 2007. HIV-1 group M conserved elements vaccine. PLoS Pathog **3:**e157.

303. **Rolland, M., S. Tovanabutra, A. C. deCamp, N. Frahm, P. B. Gilbert, E. Sanders-Buell, L. Heath, C. A. Magaret, M. Bose, A. Bradfield, A. O'Sullivan, J. Crossler, T. Jones, M. Nau, K. Wong, H. Zhao, D. N. Raugi, S. Sorensen, J. N. Stoddard, B. S. Maust, W. Deng, J. Hural, S. Dubey, N. L. Michael, J. Shiver, L. Corey, F. Li, S. G. Self, J. Kim, S. Buchbinder, D. R. Casimiro, M. N. Robertson, A. Duerr, M. J. McElrath, F. E. McCutchan, and J. I. Mullins.** 2011. Genetic impact of vaccination on breakthrough HIV-1 sequences from the STEP trial. Nat Med **17:**366-71.

304. **Rosario, M., A. Bridgeman, E. D. Quakkelaar, M. F. Quigley, B. J. Hill, M. L. Knudsen, V. Ammendola, K. Ljungberg, N. Borthwick, E. J. Im, A. J. McMichael, J. W. Drijfhout, H. Y. Greenaway, V. Venturi, D. C. Douek, S. Colloca, P. Liljestrom, A. Nicosia, D. A. Price, C. J. Melief, and T. Hanke.** 2010. Long peptides induce polyfunctional T cells against conserved regions of HIV-1 with superior breadth to single-gene vaccines in macaques. Eur J Immunol **40:**1973-84.

305. **Roy, S., N. T. Parkin, C. Rosen, J. Itovitch, and N. Sonenberg.** 1990. Structural requirements for trans activation of human immunodeficiency virus type 1 long terminal repeat-directed gene expression by tat: importance of base pairing, loop sequence, and bulges in the tat-responsive sequence. J Virol **64:**1402-6.

306. **Saez-Cirion, A., C. Lacabaratz, O. Lambotte, P. Versmisse, A. Urrutia, F. Boufassa, F. Barre-Sinoussi, J. F. Delfraissy, M. Sinet, G. Pancino, A. Venet, and E. P. H. I. V. C. S. G. Agence Nationale de Recherches sur le Sida.** 2007. HIV controllers exhibit potent CD8 T cell capacity to suppress HIV infection ex vivo and peculiar cytotoxic T lymphocyte activation phenotype. Proc Natl Acad Sci U S A **104:**6776-81.

307. **Salvi, R., A. R. Garbuglia, A. Di Caro, S. Pulciani, F. Montella, and A. Benedetto.** 1998. Grossly defective nef gene sequences in a human immunodeficiency virus type 1-seropositive long-term nonprogressor. J Virol **72:**3646-57.

308. **Santra, S., H. X. Liao, R. Zhang, M. Muldoon, S. Watson, W. Fischer, J. Theiler, J. Szinger, H. Balachandran, A. Buzby, D. Quinn, R. J. Parks, C. Y. Tsao, A. Carville, K. G. Mansfield, G. N. Pavlakis, B. K. Felber, B. F. Haynes, B. T. Korber, and N. L. Letvin.** 2010. Mosaic vaccines elicit CD8+ T lymphocyte responses that confer enhanced immune coverage of diverse HIV strains in monkeys. Nat Med **16:**324-8.

309. **Scarlata, S., and C. Carter.** 2003. Role of HIV-1 Gag domains in viral assembly. Biochim Biophys Acta **1614:**62-72.

310. **Schmitz, J. E., M. J. Kuroda, S. Santra, V. G. Sasseville, M. A. Simon, M. A. Lifton, P. Racz, K. Tenner-Racz, M. Dalesandro, B. J. Scallon, J. Ghrayeb, M. A. Forman, D. C. Montefiori, E. P. Rieber, N. L. Letvin, and K. A. Reimann.** 1999.

Control of viremia in simian immunodeficiency virus infection by CD8+ lymphocytes. Science **283:**857-60.

311. **Schneidewind, A., M. A. Brockman, R. Yang, R. I. Adam, B. Li, S. Le Gall, C. R. Rinaldo, S. L. Craggs, R. L. Allgaier, K. A. Power, T. Kuntzen, C. S. Tung, M. X. LaBute, S. M. Mueller, T. Harrer, A. J. McMichael, P. J. Goulder, C. Aiken, C. Brander, A. D. Kelleher, and T. M. Allen.** 2007. Escape from the dominant HLA-B27-restricted cytotoxic T-lymphocyte response in Gag is associated with a dramatic reduction in human immunodeficiency virus type 1 replication. J Virol **81:**12382-93.

312. **Schroeder, T. L., H. Burger, B. Weiser, V. Bengualid, J. Kimani, A. O. Anzala, M. M. Parker, D. Lamson, and S. M. Philpott.** 2005. Characterization of intersubtype recombinant HIV type 1 genomes using a nonradioactive heteroduplex tracking assay. AIDS Res Hum Retroviruses **21:**314-8.

313. **Schwartz, O., V. Marechal, S. Le Gall, F. Lemonnier, and J. M. Heard.** 1996. Endocytosis of major histocompatibility complex class I molecules is induced by the HIV-1 Nef protein. Nat Med **2:**338-42.

314. **Seaman, M. S., H. Janes, N. Hawkins, L. E. Grandpre, C. Devoy, A. Giri, R. T. Coffey, L. Harris, B. Wood, M. G. Daniels, T. Bhattacharya, A. Lapedes, V. R. Polonis, F. E. McCutchan, P. B. Gilbert, S. G. Self, B. T. Korber, D. C. Montefiori, and J. R. Mascola.** 2010. Tiered categorization of a diverse panel of HIV-1 Env pseudoviruses for assessment of neutralizing antibodies. J Virol **84:**1439-52.

315. **Serwanga, J., S. Mugaba, E. Pimego, B. Nanteza, F. Lyagoba, S. Nakubulwa, L. Heath, R. N. Nsubuga, N. Ndembi, F. Gotch, and P. Kaleebu.** 2011. Profile of T Cell Recognition of HIV Type 1 Consensus Group M Gag and Nef Peptides in a Clade A1- and D-Infected Ugandan Population. AIDS Res Hum Retroviruses.

316. **Sette, A., and J. Sidney.** 1998. HLA supertypes and supermotifs: a functional perspective on HLA polymorphism. Curr Opin Immunol **10:**478-82.

317. **Sette, A., and J. Sidney.** 1999. Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. Immunogenetics **50:**201-12.

318. **Sheehy, A. M., N. C. Gaddis, J. D. Choi, and M. H. Malim.** 2002. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature **418:**646-50.

319. **Sherman, M. P., C. M. de Noronha, L. A. Eckstein, J. Hataye, P. Mundt, S. A. Williams, J. A. Neidleman, M. A. Goldsmith, and W. C. Greene.** 2003. Nuclear export of Vpr is required for efficient replication of human immunodeficiency virus type 1 in tissue macrophages. J Virol **77:**7582-9.

320. **Shiina, T., K. Hosomichi, H. Inoko, and J. K. Kulski.** 2009. The HLA genomic loci map: expression, interaction, diversity and disease. J Hum Genet **54:**15-39.

321. **Short, G. F., 3rd, A. L. Laikhter, M. Lodder, Y. Shayo, T. Arslan, and S. M. Hecht.** 2000. Probing the S1/S1' substrate binding pocket geometry of HIV-1 protease with modified aspartic acid analogues. Biochemistry **39:**8768-81.

322. **Simon, F., I. Loussert-Ajaka, F. Damond, S. Saragosti, F. Barin, and F. Brun-Vezinet.** 1996. HIV type 1 diversity in northern Paris, France. AIDS Res Hum Retroviruses **12:**1427-33.

323. **Srinivasakumar, N., M. L. Hammarskjold, and D. Rekosh.** 1995. Characterization of deletion mutations in the capsid region of human immunodeficiency virus type 1 that affect particle formation and Gag-Pol precursor incorporation. J Virol **69:**6106-14.

324. **Stephens, H. A.** 2005. HIV-1 diversity versus HLA class I polymorphism. Trends Immunol **26:**41-7.

325. **Stevens, W., A. Kamali, E. Karita, O. Anzala, E. J. Sanders, W. Jaoko, P. Kaleebu, J. Mulenga, L. Dally, P. Fast, J. Gilmour, B. Farah, J. Birungi, P. Hughes, O. Manigart, G. Stevens, S. Yates, H. Thomson, A. von Lieven, M. Krebs, M. A. Price, L. Stoll-Johnson, and N. Ketter.** 2008. Baseline morbidity in 2,990 adult African volunteers recruited to characterize laboratory reference intervals for future HIV vaccine clinical trials. PLoS One **3:**e2043.

326. **Stranford, S. A., J. C. Ong, B. Martinez-Marino, M. Busch, F. M. Hecht, J. Kahn, and J. A. Levy.** 2001. Reduction in CD8+ cell noncytotoxic anti-HIV activity in individuals receiving highly active antiretroviral therapy during primary infection. Proc Natl Acad Sci U S A **98:**597-602.

327. **Swanstrom, R., and J. W. Wills.** 1997. Synthesis, Assembly, and Processing of Viral Proteins.

328. **Tebit, D. M., and E. J. Arts.** 2011. Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease. Lancet Infect Dis **11:**45-56.

329. **Temin, H. M., and D. P. Bolognesi.** 1993. AIDS. Where has HIV been hiding? Nature **362:**292-3.

330. **Thali, M., A. Bukovsky, E. Kondo, B. Rosenwirth, C. T. Walsh, J. Sodroski, and H. G. Gottlinger.** 1994. Functional association of cyclophilin A with HIV-1 virions. Nature **372:**363-5.

331. **Thurmond, J., H. Yoon, C. Kuiken, K. Yusim, S. Perkins, J. Theiler, T. Bhattacharya, B. Korber, and W. Fischer.** 2008. Web-based design and evaluation of T-cell vaccine candidates. Bioinformatics **24:**1639-40.

332. **Tian, W., D. A. Boggs, G. Uko, A. Essiet, M. Inyama, B. Banjoko, T. Adewole, W. Z. Ding, M. Mohseni, R. Fritz, D. F. Chen, L. J. Palmer, and P. A. Fraser.** 2003. MICA, HLA-B haplotypic variation in five population groups of sub-Saharan African ancestry. Genes Immun **4:**500-5.

333. **Tolstrup, M., A. L. Laursen, J. Gerstoft, F. S. Pedersen, L. Ostergaard, and M. Duch.** 2006. Cysteine 138 mutation in HIV-1 Nef from patients with delayed disease progression. Sex Health **3:**281-6.

334. **Triques, K., A. Bourgeois, N. Vidal, E. Mpoudi-Ngole, C. Mulanga-Kabeya, N. Nzilambi, N. Torimiro, E. Saman, E. Delaporte, and M. Peeters.** 2000. Near-full-length genome sequencing of divergent African HIV type 1 subtype F viruses leads to the identification of a new HIV type 1 subtype designated K. AIDS Res Hum Retroviruses **16:**139-51.

335. **Troyer, R. M., J. McNevin, Y. Liu, S. C. Zhang, R. W. Krizan, A. Abraha, D. M. Tebit, H. Zhao, S. Avila, M. A. Lobritz, M. J. McElrath, S. Le Gall, J. I. Mullins, and E. J. Arts.** 2009. Variable fitness impact of HIV-1 escape mutations to cytotoxic T lymphocyte (CTL) response. PLoS Pathog **5:**e1000365.

336. **Ueno, T., H. Tomiyama, M. Fujiwara, S. Oka, and M. Takiguchi.** 2004. Functionally impaired HIV-specific CD8 T cells show high affinity TCR-ligand interactions. J Immunol **173:**5451-7.

337. **Valentine, L. E., S. M. Piaskowski, E. G. Rakasz, N. L. Henry, N. A. Wilson, and D. I. Watkins.** 2008. Recognition of escape variants in ELISPOT does not always predict CD8+ T-cell recognition of simian immunodeficiency virus-infected cells expressing the same variant sequences. J Virol **82:**575-81.

338. **Vallari, A., V. Holzmayer, B. Harris, J. Yamaguchi, C. Ngansop, F. Makamche, D. Mbanya, L. Kaptue, N. Ndembi, L. Gurtler, S. Devare, and C. A. Brennan.** 2011. Confirmation of putative HIV-1 group P in Cameroon. J Virol **85:**1403-7.

339. **van Harmelen, J., R. Wood, M. Lambrick, E. P. Rybicki, A. L. Williamson, and C. Williamson.** 1997. An association between HIV-1 subtypes and mode of transmission in Cape Town, South Africa. AIDS **11:**81-7.

340. **Vardas, E., P. Kaleebu, L. G. Bekker, A. Hoosen, E. Chomba, P. R. Johnson, P. Anklesaria, J. Birungi, B. Barin, M. Boaz, J. Cox, J. Lehrman, G. Stevens, J. Gilmour, T. Tarragona, P. Hayes, S. Lowenbein, E. Kizito, P. Fast, A. E. Heald, and C. Schmidt.** 2010. A phase 2 study to evaluate the safety and immunogenicity of a recombinant HIV type 1 vaccine based on adeno-associated virus. AIDS Res Hum Retroviruses **26:**933-42.

341. **Vasan, S., S. J. Schlesinger, Z. Chen, A. Hurley, A. Lombardo, S. Than, P. Adesanya, C. Bunce, M. Boaz, R. Boyle, E. Sayeed, L. Clark, D. Dugin, M. Boente-Carrera, C. Schmidt, Q. Fang, LeiBa, Y. Huang, G. J. Zaharatos, D. F. Gardiner, M. Caskey, L. Seamons, M. Ho, L. Dally, C. Smith, J. Cox, D. Gill, J. Gilmour, M. C. Keefer, P. Fast, and D. D. Ho.** 2010. Phase 1 safety and immunogenicity evaluation of ADMVA, a multigenic, modified vaccinia Ankara-HIV-1 B'/C candidate vaccine. PLoS One **5:**e8816.

342. **Vasan, S., S. J. Schlesinger, Y. Huang, A. Hurley, A. Lombardo, Z. Chen, S. Than, P. Adesanya, C. Bunce, M. Boaz, R. Boyle, E. Sayeed, L. Clark, D. Dugin, C. Schmidt, Y. Song, L. Seamons, L. Dally, M. Ho, C. Smith, M. Markowitz, J.**

Cox, D. K. Gill, J. Gilmour, M. C. Keefer, P. Fast, and D. D. Ho. 2010. Phase 1 safety and immunogenicity evaluation of ADVAX, a multigenic, DNA-based clade C/B' HIV-1 candidate vaccine. PLoS One **5:**e8617.

343. **Verheyen, J., E. Litau, T. Sing, M. Daumer, M. Balduin, M. Oette, G. Fatkenheuer, J. K. Rockstroh, U. Schuldenzucker, D. Hoffmann, H. Pfister, and R. Kaiser.** 2006. Compensatory mutations at the HIV cleavage sites p7/p1 and p1/p6-gag in therapy-naive and therapy-experienced patients. Antivir Ther **11:**879-87.

344. **Vodicka, M. A., D. M. Koepp, P. A. Silver, and M. Emerman.** 1998. HIV-1 Vpr interacts with the nuclear transport pathway to promote macrophage infection. Genes Dev **12:**175-85.

345. **Vogt, R. L., S. Richmond-Crum, and A. Diwan.** 1997. Hepatitis C virus infection in a human immunodeficiency virus-positive cohort in Hawaii. J Infect Dis **176:**542-3.

346. **Vogt, V. M.** 1996. Proteolytic processing and particle maturation. Curr Top Microbiol Immunol **214:**95-131.

347. **von Schwedler, U. K., M. Stuchell, B. Muller, D. M. Ward, H. Y. Chung, E. Morita, H. E. Wang, T. Davis, G. P. He, D. M. Cimbora, A. Scott, H. G. Krausslich, J. Kaplan, S. G. Morham, and W. I. Sundquist.** 2003. The protein network of HIV budding. Cell **114:**701-13.

348. **Votteler, J., L. Neumann, S. Hahn, F. Hahn, P. Rauch, K. Schmidt, N. Studtrucker, S. M. Solbak, T. Fossen, P. Henklein, D. E. Ott, G. Holland, N. Bannert, and U. Schubert.** 2011. Highly conserved serine residue 40 in HIV-1 p6 regulates capsid processing and virus core assembly. Retrovirology **8:**11.

349. **Walk, S. F., M. Alexander, B. Maier, M. L. Hammarskjold, D. M. Rekosh, and K. S. Ravichandran.** 2001. Design and use of an inducibly activated human immunodeficiency virus type 1 Nef to study immune modulation. J Virol **75:**834-43.

350. **Walker, B. D., and B. T. Korber.** 2001. Immune control of HIV: the obstacles of HLA and viral diversity. Nat Immunol **2:**473-5.

351. **Walker, L. M., M. D. Simek, F. Priddy, J. S. Gach, D. Wagner, M. B. Zwick, S. K. Phogat, P. Poignard, and D. R. Burton.** 2010. A limited number of antibody specificities mediate broad and potent serum neutralization in selected HIV-1 infected individuals. PLoS Pathog **6:**e1001028.

352. **Wang, J. Y., H. Ling, W. Yang, and R. Craigie.** 2001. Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. EMBO J **20:**7333-43.

353. **Weaver, E. A., Z. T. Camacho, and F. Gao.** 2010. Similar T-cell immune responses induced by group M consensus env immunogens with wild-type or minimum consensus variable regions. AIDS Res Hum Retroviruses **26:**577-84.

354. **Weaver, E. A., Z. Lu, Z. T. Camacho, F. Moukdar, H. X. Liao, B. J. Ma, M. Muldoon, J. Theiler, G. J. Nabel, N. L. Letvin, B. T. Korber, B. H. Hahn, B. F.**

Haynes, and F. Gao. 2006. Cross-subtype T-cell immune responses induced by a human immunodeficiency virus type 1 group m consensus env immunogen. J Virol **80:**6745-56.

355. **Welker, R., M. Harris, B. Cardel, and H. G. Krausslich.** 1998. Virion incorporation of human immunodeficiency virus type 1 Nef is mediated by a bipartite membrane-targeting signal: analysis of its role in enhancement of viral infectivity. J Virol **72:**8833-40.

356. **Wiegers, K., G. Rutter, H. Kottler, U. Tessmer, H. Hohenberg, and H. G. Krausslich.** 1998. Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. J Virol **72:**2846-54.

357. **Williams, F., A. Meenagh, C. Darke, A. Acosta, A. S. Daar, C. Gorodezky, M. Hammond, E. Nascimento, and D. Middleton.** 2001. Analysis of the distribution of HLA-B alleles in populations from five continents. Hum Immunol **62:**645-50.

358. **Williamson, C., L. Morris, M. F. Maughan, L. H. Ping, S. A. Dryga, R. Thomas, E. A. Reap, T. Cilliers, J. van Harmelen, A. Pascual, G. Ramjee, G. Gray, R. Johnston, S. A. Karim, and R. Swanstrom.** 2003. Characterization and selection of HIV-1 subtype C isolates for use in vaccine development. AIDS Res Hum Retroviruses **19:**133-44.

359. **Wills, J. W., and R. C. Craven.** 1991. Form, function, and use of retroviral gag proteins. AIDS **5:**639-54.

360. **Wolinsky, S. M., K. J. Kunstman, J. T. Safrit, R. A. Koup, A. U. Neumann, and B. T. Korber.** 1996. Response: HIV-1 Evolution and Disease Progression. Science **274:**1010-1.

361. **Woo, J., D. L. Robertson, and S. C. Lovell.** 2010. Constraints on HIV-1 diversity from protein structure. J Virol **84:**12995-3003.

362. **Xiang, J., H. Huang, and Y. Liu.** 2005. A new dynamic model of CD8+ T effector cell responses via CD4+ T helper-antigen-presenting cells. J Immunol **174:**7497-505.

363. **Yang, O. O.** 2009. Candidate vaccine sequences to represent intra- and inter-clade HIV-1 variation. PLoS One **4:**e7388.

364. **Yang, O. O., P. T. Sarkis, A. Trocha, S. A. Kalams, R. P. Johnson, and B. D. Walker.** 2003. Impacts of avidity and specificity on the antiviral efficiency of HIV-1-specific CTL. J Immunol **171:**3718-24.

365. **Yang, Y., J. Huang, I. Toth, M. Lichterfeld, and X. G. Yu.** 2010. Mutational escape in HIV-1 CTL epitopes leads to increased binding to inhibitory myelomonocytic MHC class I receptors. PLoS One **5:**e15084.

366. **Yant, L. J., T. C. Friedrich, R. C. Johnson, G. E. May, N. J. Maness, A. M. Enz, J. D. Lifson, D. H. O'Connor, M. Carrington, and D. I. Watkins.** 2006. The high-frequency major histocompatibility complex class I allele Mamu-B*17 is associated

with control of simian immunodeficiency virus SIVmac239 replication. J Virol **80:**5074-7.

367. **Yoo, S., D. G. Myszka, C. Yeh, M. McMurray, C. P. Hill, and W. I. Sundquist.** 1997. Molecular recognition in the HIV-1 capsid/cyclophilin A complex. J Mol Biol **269:**780-95.

368. **Younes, S. A., B. Yassine-Diab, A. R. Dumont, M. R. Boulassel, Z. Grossman, J. P. Routy, and R. P. Sekaly.** 2003. HIV-1 viremia prevents the establishment of interleukin 2-producing HIV-specific memory CD4+ T cells endowed with proliferative capacity. J Exp Med **198:**1909-22.

369. **Yu, G., and R. L. Felsted.** 1992. Effect of myristoylation on p27 nef subcellular distribution and suppression of HIV-LTR transcription. Virology **187:**46-55.

370. **Yu, Q., R. Konig, S. Pillai, K. Chiles, M. Kearney, S. Palmer, D. Richman, J. M. Coffin, and N. R. Landau.** 2004. Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. Nat Struct Mol Biol **11:**435-42.

371. **Yu, X. F., Z. Matsuda, Q. C. Yu, T. H. Lee, and M. Essex.** 1995. Role of the C terminus Gag protein in human immunodeficiency virus type 1 virion assembly and maturation. J Gen Virol **76 (Pt 12):**3171-9.

372. **Yu, X. G., M. Lichterfeld, B. Perkins, E. Kalife, S. Mui, J. Chen, M. Cheng, W. Kang, G. Alter, C. Brander, B. D. Walker, and M. Altfeld.** 2005. High degree of inter-clade cross-reactivity of HIV-1-specific T cell responses at the single peptide level. AIDS **19:**1449-56.

373. **Zhao, S., S. Zhai, Y. Zhuang, S. Wang, D. Huang, W. Kang, X. Li, X. G. Yu, B. D. Walker, M. A. Altfeld, and Y. Sun.** 2007. Inter-clade cross-reactivity of HIV-1-specific T cell responses in human immunodeficiency virus type 1 infection in China. Curr HIV Res **5:**251-9.

374. **Zuniga, R., A. Lucchetti, P. Galvan, S. Sanchez, C. Sanchez, A. Hernandez, H. Sanchez, N. Frahm, C. H. Linde, H. S. Hewitt, W. Hildebrand, M. Altfeld, T. M. Allen, B. D. Walker, B. T. Korber, T. Leitner, J. Sanchez, and C. Brander.** 2006. Relative dominance of Gag p24-specific cytotoxic T lymphocytes is associated with human immunodeficiency virus control. J Virol **80:**3122-5.

375. **Zydowsky, L. D., F. A. Etzkorn, H. Y. Chang, S. B. Ferguson, L. A. Stolz, S. I. Ho, and C. T. Walsh.** 1992. Active site mutants of human cyclophilin A separate peptidyl-prolyl isomerase activity from cyclosporin A binding and calcineurin inhibition. Protein Sci **1:**1092-9.

WEBSITES CITED

UNAIDS epidemic update
www.unaids.org
Accessed 10 January 2012

Los Alamos national Laboratory Databases
http://hiv-web.lanl.gov
www.hiv.lanl.gov
www.hiv.lanl.gov/content/immunology
Accessed 04 February 2012

HLA database
http://www.ebi.ac.uk/imgt/hla/stats.html
Accessed 10 October 2011

NetMHC
www.abs.dtu.dk/services/NetMHC
Accessed 15 November 2011

IAVI
http://www.iavi.org/archives/2010
Accesses 07 February 2012

AIDS Vaccine 2011 Conference, Bangkok, Thianland; 12-15 September 2011
http://www.hivvaccineenterprise.org/conference/2011/detailed-program
Accessed 08 February 2012

Mapping database
www.naturalearthdata.com
Accessed 26 November 2011

APPENDICES

Appendix A

Table A1. Amino acid codes

Amino Acid	3-Letter Code	IUB Code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cystine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenyalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Table A2. Amino acids and the genetic code

3-letter code	Unambiguous codons	IUB codons
Ala	GCT, GCC, GCA, GCG	GCX
Asp, Asn	GAT, GAC, AAT, AAC	RAY
Cys	TGT, TGC	TGY
Asp	GAT, GAC	GAY
Glu	GAA, GAG	GAR
Phe	TTT, TTC	TTY
Gly	GGT, GGC, GGA, GGG	GGX
His	CAT, CAC	CAY
Ile	ATT, ATC, ATA	ATH
Lys	AAA, AAG	AAR
Leu	TTG, TTA,	TTR
	CTT, CTC, CTA, CTG	CTX
Met	ATG	ATG
Asn	AAT, AAC	AAY
Pro	CCT, CCC, CCA, CCG	CCX
Gln	CAA, CAG	CAR
Arg	CGT, CGC, CGA, CGG,	CGX
	AGA, AGG	AGR
Ser	TCT, TCC, TCA, TCG, AGT,	TCX
	AGC	AGY
Thr	ACT, ACC, ACA, ACG	ACX
Val	GTT, GTC, GTA, GTG	GTX
Trp	TGG	TGG
XXX		XXX
Tyr	TAT, TAC	TAY
Glu, Gln	GAA, GAG, CAA, CAG	SAR
End (Stop)	TAA, TAG	TAR
	TGA	TRA

The table shows the 20 amino acids that are found in proteins. The single and three-letter code used to represent these amino acids are also listed along with the 64 possible three-letter combinations of the DNA coding units namely T (Thymine), C (Cytosine), A (Adenine) and G (Guanine), used either to encode one of these amino acids or as a stop codon that signals the end of a protein. This is the information incorporated into sequence analysis programs used to translate viral DNA sequences into protein sequences for analyses together with ELISpot peptide reagents. X-denotes that the last position of certain triple codons can be ambiguous.

Appendix B. PCR and cycling conditions

These conditions apply for study individuals whose data is reported in Chapter 3 for *gag* and *nef* sequencing. PCR conditions for study participants in Chapter 2 have already been reported in MSc Thesis (2007). However, the conditions are similar except that for Chapter 2 participants, only full-length *gag* amplification was performed while in the following conditions, both *gag* and *nef* genes were amplified. In addition, different primers for *gag* gene were used that amplified the whole of the *gag* gene continuously while previously, the *gag* gene was amplified in three fragments A, B and C (MSc Thesis, 2007). The sequences of the primers for Gag and Nef amplification are shown in Table B.

Table B1. Gag and Nef primer sequences

Primer	Sequence	HXB2 position
Gag D Reverse	5′-AAT TCC TCC TAT CAT TTT TGG-3′	2382 - 2402
Gag D Forward	5′-TCT CTA GCA GTG GCG CCC G-3′	626 - 644
Gag A Forward	5′-CTC TCG ACG CAG GAC TCG GCT T-3′	683 - 704
Gag C Reverse	5′-TCT TCT AAT ACT GTA TCA TCT GC-3	2334 - 2356
Nef O Reverse	5′-AGG CAA GCT TTA TTG AGG-3′	9608 - 9625
Nef SQ15FC	5′-GAG AGC GGT GGA ACT TCT-3	8561 - 8578
Nef Forward	5′-CCT AGA AGA ATA AGA CAG GGC TT-3	8754 - 8776
Nef Reverse	5′-CCT GGA ACG CCC CAG TGG-3′	9443 - 9461

Appendix B1. HIV-1 *gag* PCR and cycling conditions

Table B1.1. HIV-1 *gag* cDNA synthesis

Reagent	Volume (µl)
Primer Gag D Reverse	**1**
RNA	**7**
Total	8

Table B1.2. HIV-1 *gag* RT PCR

Reagent	Volume (µl)
H_2O	**0.7**
Buffer (X5)	**4.0**
$MgCl_2$	**4.8**
dNTPs (10mM)	**1.0**
RNasin (40U/µl)	**0.5**
RT Enzyme (15U/µl)	**1.0**
Total	12

Table B1.3. HIV-1 *gag* first round PCR

Reagent	**Volume (µl)**
H_2O	**37.25**
Buffer (10X) + $MgCl_2$ (15mM)	**5.0**
dNTPs (10mM each)	**1.0**
Primer Gag D Reverse (10 pmoles/µl)	**1.0**
Primer Gag D Forward (10 pmoles/µl)	**1.0**
Expand High Fidelity Enzyme (5U/µl)	**0.75**
cDNA	**4.0**
Total	50.0

Table B1.4. HIV-1 *gag* first round PCR cycling conditions

	10 cycles	15 cycles		
95^0C (2 min)	**95^0C (15 sec) 52^0C (30 sec) 72^0C (1 min)**	**95^0C (15 sec) 55^0C (30 sec) 72^0C (1.5 min)**	**72^0C (7 min)**	**4^0C (∞)**

Table B1.5. HIV-1 *gag* second round PCR

Reagent	**Volume (µl)**
H_2O	**37.25**
Buffer (10X) + $MgCl_2$ (15mM)	**5.0**
dNTPs (10mM each)	**1.0**
Primer Gag A Forward (10 pmoles/µl)	**1.0**
Primer Gag C Reverse (10 pmoles/µl)	**1.0**
Expand High Fidelity Enzyme (5U/µl)	**0.75**
cDNA	**4.0**
Total	50.0

Table B1.6. HIV-1 *gag* second round PCR cycling conditions

	10 cycles	15 cycles		
95^0C (2 min)	**95^0C (15 sec) 52^0C (30 sec) 72^0C (1 min)**	**95^0C (15 sec) 55^0C (30 sec) 72^0C (1. min)**	**72^0C (7 min)**	**4^0C (∞)**

Appendix B2. HIV-1 *nef* PCR and cycling conditions

Table B2.1. HIV-1 *nef* cDNA synthesis

Reagent	Volume (µl)
Primer Nef OR	**1**
RNA	**7**
Total	8

Table B2.2. HIV-1 *nef* RT PCR

Reagent	Volume (µl)
H_2O	**0.7**
Buffer (X5)	**4.0**
$MgCl_2$	**4.8**
dNTPs (10mM)	**1.0**
RNasin (40U/µl)	**0.5**
RT Enzyme (15U/µl)	**1.0**
Total	12

Table B2.3. HIV-1 *nef* first round PCR

Reagent	Volume (µl)
H_2O	**37.25**
Buffer (10X) + $MgCl_2$ (15mM)	**5.0**
dNTPs (10mM each)	**1.0**
Primer Nef SQ15FC or E200 (10 pmoles/µl)	**1.0**
Primer Nef OR (10 pmoles/µl)	**1.0**
Expand High Fidelity Enzyme (5U/µl)	**0.75**
cDNA	**4.0**
Total	50.0

Table B2.4. HIV-1 *nef* first round PCR cycling conditions

	10 cycles	15 cycles		
95^0C (2 min)	**95^0C (15 sec)** **53^0C (30 sec)** **72^0C (45 sec)**	**95^0C (15 sec)** **55^0C (30 sec)** **72^0C (1 min)**	**72^0C (7 min)**	**4^0C (∞)**

Table B2.5. HIV-1 *nef* second round PCR

Reagent	Volume (µl)
H_2O	**37.25**
Buffer (10X) + $MgCl_2$ (15mM)	**5.0**
dNTPs (10mM each)	**1.0**
Primer Nef Forward (10 pmoles/µl)	**1.0**
Primer Nef Reverse (10 pmoles/µl)	**1.0**
Expand High Fidelity Enzyme (5U/µl)	**0.75**
cDNA	**4.0**
Total	50.0

Table B2.6. HIV-1 *nef* second round PCR cycling conditions

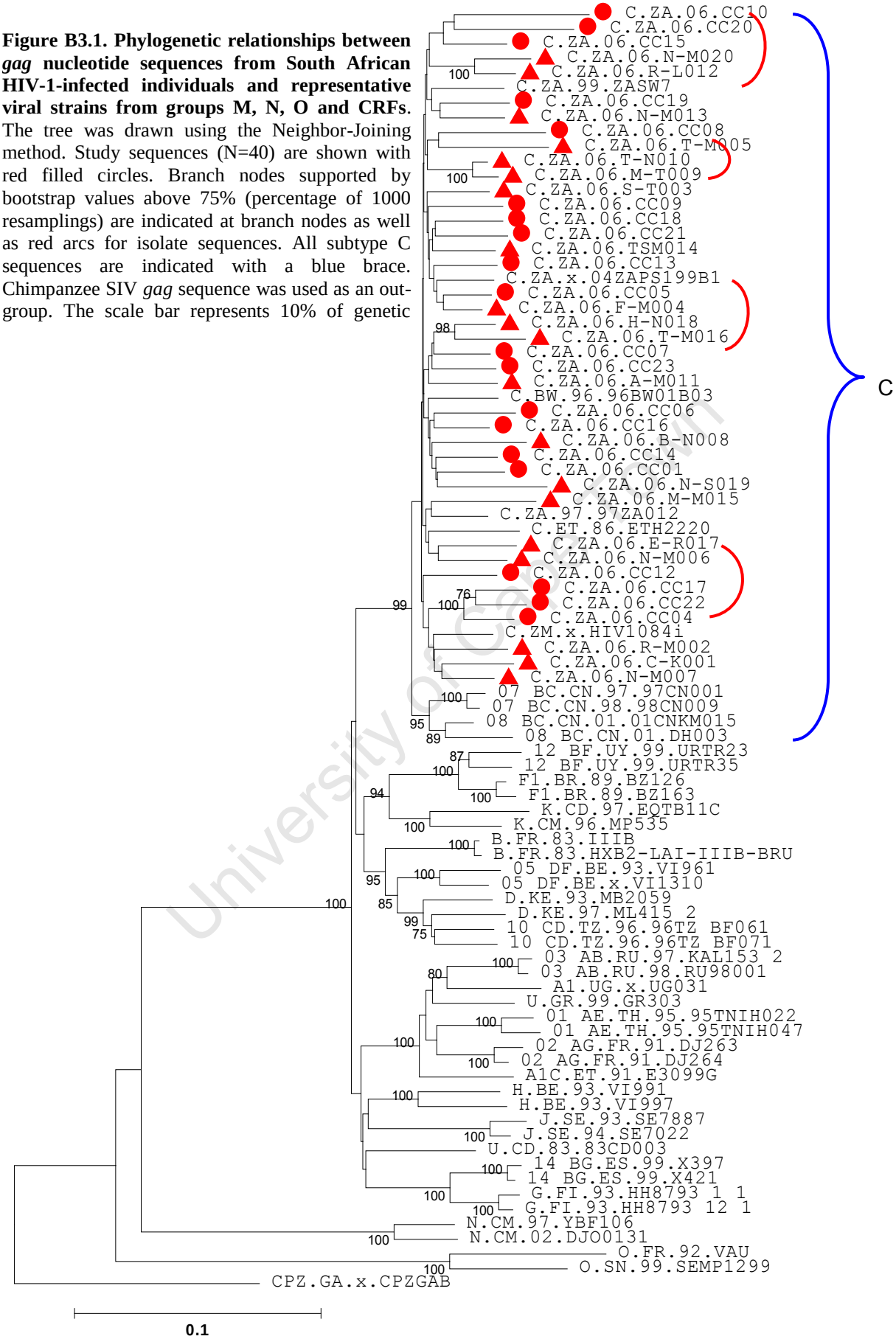
	35 cycles		
95^0C (2 min)	**95^0C (15 sec)** **55^0C (30 sec)** **72^0C (1 min)**	**72^0C (7 min)**	**4^0C (∞)**

Appendix B3. Sequence analyses

HIV-1 clade classification (subtyping) of infecting viral sequences

Infecting viral HIV-1 full-length *gag* sequences were aligned together with reference sequences from the Los Alamos HIBV sequence database using BioEdit sequence alignment editor and phylogenetic relatedness inferred using the Neighbor-Joining method, Kimura 2-parameter model with 1000 botstraps test of phylogeny as described in MSc Thesis, Lycias Zembe, UCT (2007).

Figure B3.1. Phylogenetic relationships between *gag* nucleotide sequences from South African HIV-1-infected individuals and representative viral strains from groups M, N, O and CRFs. The tree was drawn using the Neighbor-Joining method. Study sequences (N=40) are shown with red filled circles. Branch nodes supported by bootstrap values above 75% (percentage of 1000 resamplings) are indicated at branch nodes as well as red arcs for isolate sequences. All subtype C sequences are indicated with a blue brace. Chimpanzee SIV *gag* sequence was used as an out-group. The scale bar represents 10% of genetic



Genetic relatedness of infecting viral sequences and ELISpot peptide reagents

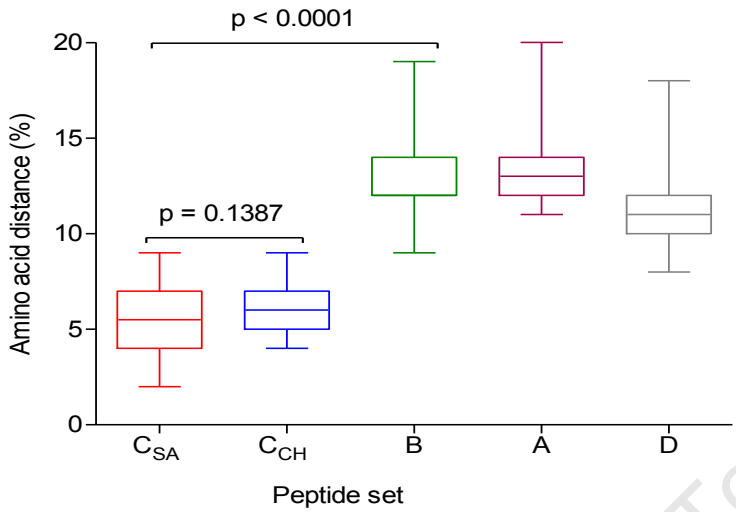


Figure B3.2. Amino Acid distances between study subject sequences and ELISpot peptide reagents. Shown in the figure are the minimum, the 25th, 50th (median), 75th percent quartiles and maximum values for each peptide set. The line in the box represents the median or 50th percentile. The p value is from the non-parametric Mann-Whitney test.

B4: Recombination analysis

Each of the study participants' full length *gag* sequence was analyzed for recombination using REGA version 6.4.1, a software program which compares each nucleotide position to other HIV-1 pure subtype and CRF reference sequences to determine whether there is sufficient phylogenetic signal to classify the sequence as a pure subtype of CRF, recombinant viruses or unclassified viral subtypes (de Oliveira *et al*., 2005).

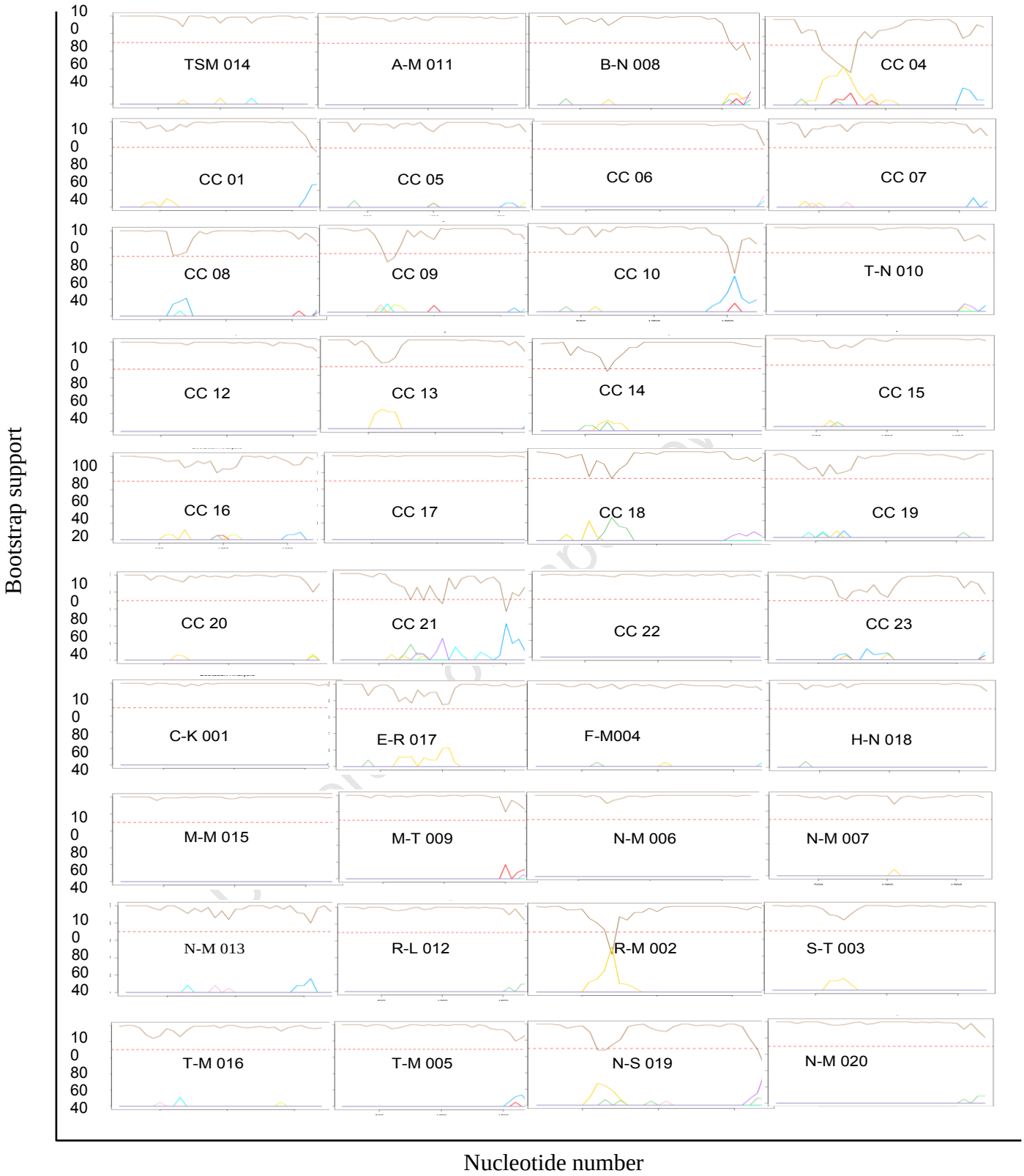


Figure B4.1. Recombination analysis of isolate sequences using REGA. The red dotted lines represent 70% cut-off value of 100 replicates. The different isolate sequences are indicated on each graph.

Appendix C. Mosaic analysis for each peptide reagent

The Mosaic Vaccine Tool

This is a set of on-line tools that can be used to design candidate vaccine immunogen sequences as well as assessing their ability as antigens using coverage of short fragments, k-mers (9-12 amino acids) as potential epitopes (Fischer *et al*., 2007). The tool is a suite made up of three separate tools namely, the Mosaic Vaccine Designer Tool, The Vaccine Epitope Coverage Tool (Epicover) and the Positional Epitope Coverage Tool (Posicover). The vaccine designer tool uses a computer algorithm mimicking recombination events in natural viral infection in a series of iterations chosen by the user to design mosaic sequences with the best coverage of the test set sequences (Fischer *et al*., 2007; Thurmond *et al*., 2008). The two tools, Epicover and Posicover are assessment tools that compute how well a suggested set of peptides covers potential epitopes in a test set of natural sequences (Thurmond *et al*., 2008). The suggested set of peptides can be from the mosaic vaccine designer tool or any other vaccine candidate peptides for assessing T-cell responses in assays.

Epicover calculates the average coverage of the test-set population by antigen-set of k-mers while Posicover provides detailed positional coverage information relative to a test-set alignment. Studies of T-cell recognition of viral peptides have shown that some peptides are recognized even if they have amino acid mismatches with the optimally defined epitope. Therefore, the tool also computes coverage based on exact matches of epitopes as well as those mismatching by 1 or 2 amino acids. Both tools provide graphical outputs and allow detailed user control for example epitope length to use (which is 9-mers by default) and the number of antigen sets to use for computing. In Figures C1 to Figure C5, graphical representations of Posicover results for individual peptide reagent (C_{Du422}, C_{CH}, A, B and D) used in ELISpot assay are shown.

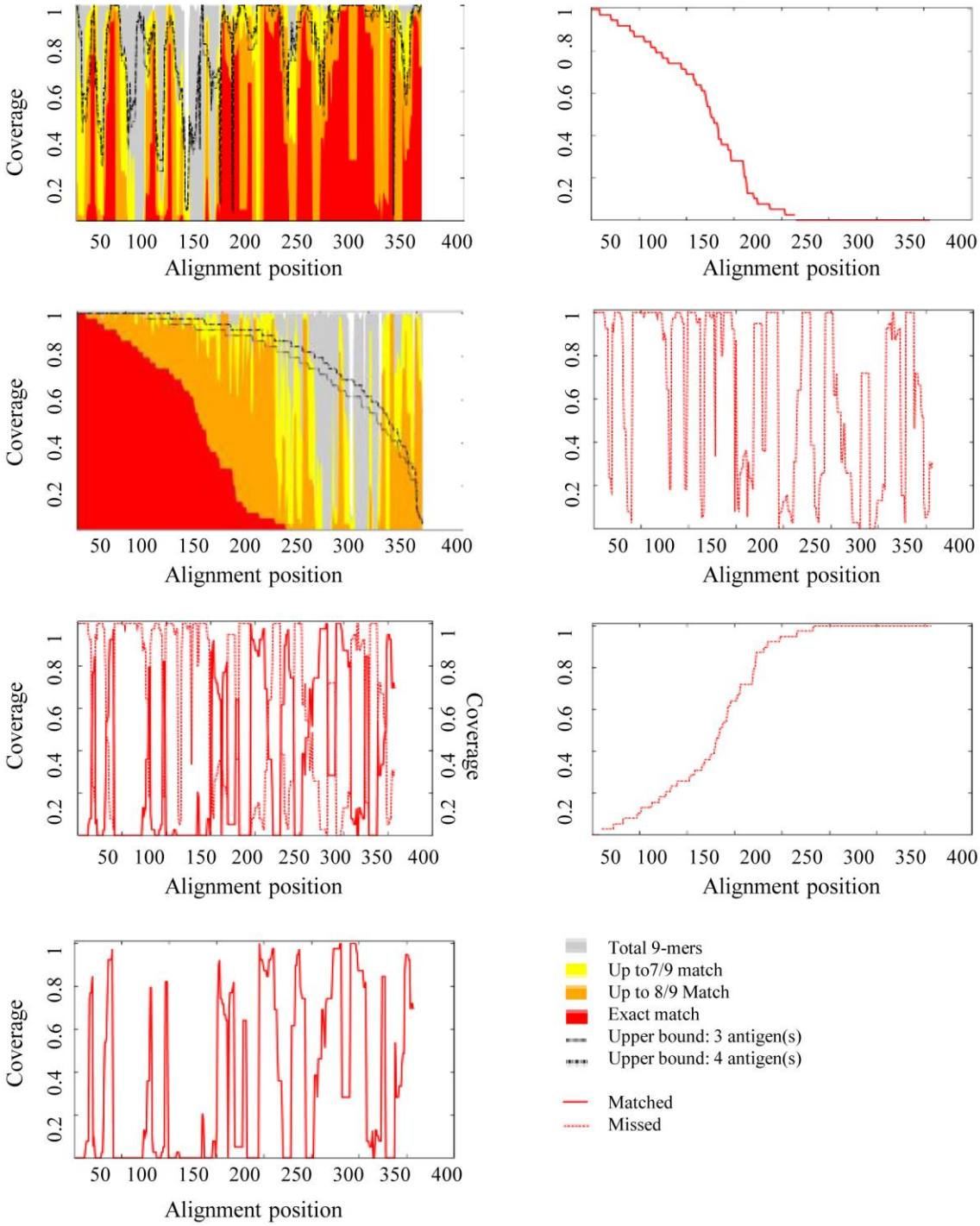


Figure C1. Positional epitope coverage (Posicover) of infecting viral sequences by HIV-1 clade A peptide reagent sequence. (A) Coverage by position in the alignment of the Gag protein (p17p24p2). **(B)** Coverage ranked. **(C)** Matched 9-mer epitopes and missed 9-mer epitopes coverage by natural position in the alignment. **(D)** Matched 9-mers epitopes ranked by coverage. (E) Missed 9-mers epitopes coverage by natural position along the Gag alignment. **(F)** Missed 9-mers by natural position in along the Gag alignment. **(G)** Missed 9-mers ranked by coverage.

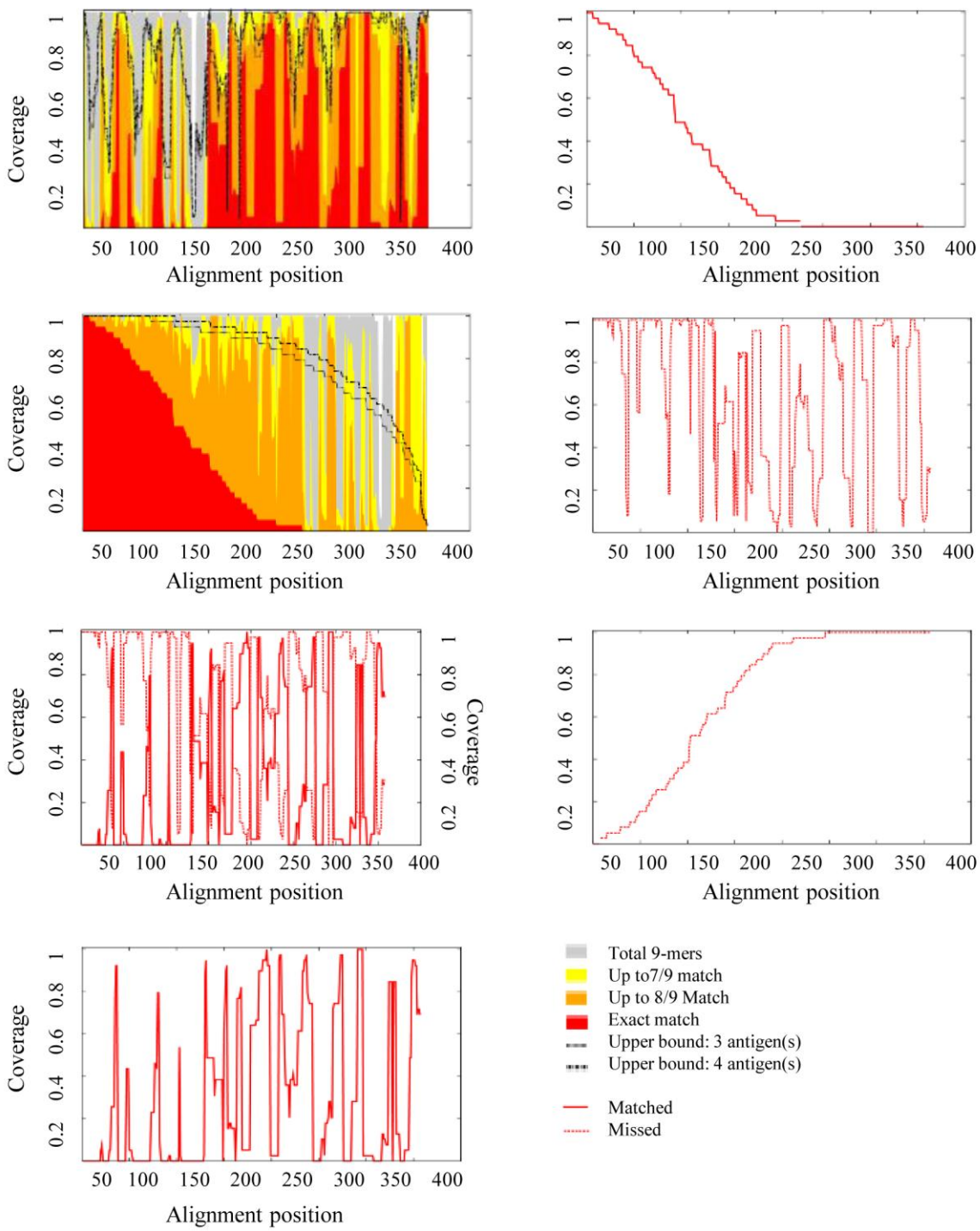


Figure C2. Positional epitope coverage (Posicover) of infecting viral sequences by HIV-1 clade B peptide reagent sequence. (A) Coverage by position in the alignment of the Gag protein (p17p24p2). **(B)** Coverage ranked. **(C)** Matched 9-mer epitopes and missed 9-mer epitopes coverage by natural position in the alignment. **(D)** Matched 9-mers epitopes ranked by coverage. **(E)** Missed 9-mers epitopes coverage by natural position along the Gag alignment**. (F)** Missed 9-mers by natural position in along the Gag alignment. **(G)** Missed 9-mers ranked by coverage.

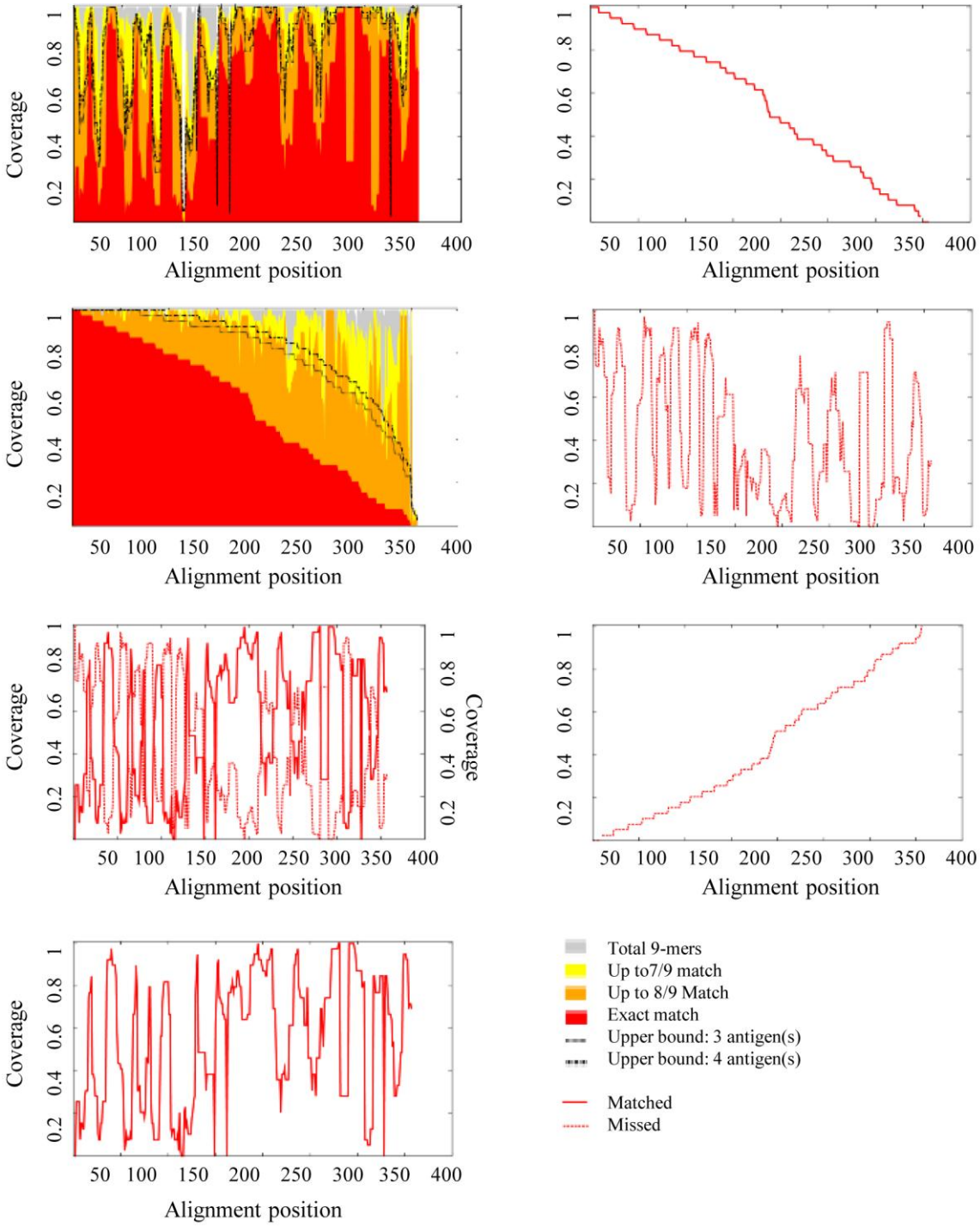


Figure C3. Positional epitope coverage (Posicover) of infecting viral sequences by HIV-1 clade C (C_{Du422}, South African) peptide reagent sequence. (A) Coverage by position in the alignment of the Gag protein (p17p24p2). **(B)** Coverage ranked. **(C)** Matched 9-mer epitopes and missed 9-mer epitopes coverage by natural position in the alignment. **(D)** Matched 9-mers epitopes ranked by coverage. **(E)** Missed 9-mers epitopes coverage by natural position along the Gag alignment. **(F)** Missed 9-mers by natural position in along the Gag alignment. **(G)** Missed 9-mers ranked by coverage.

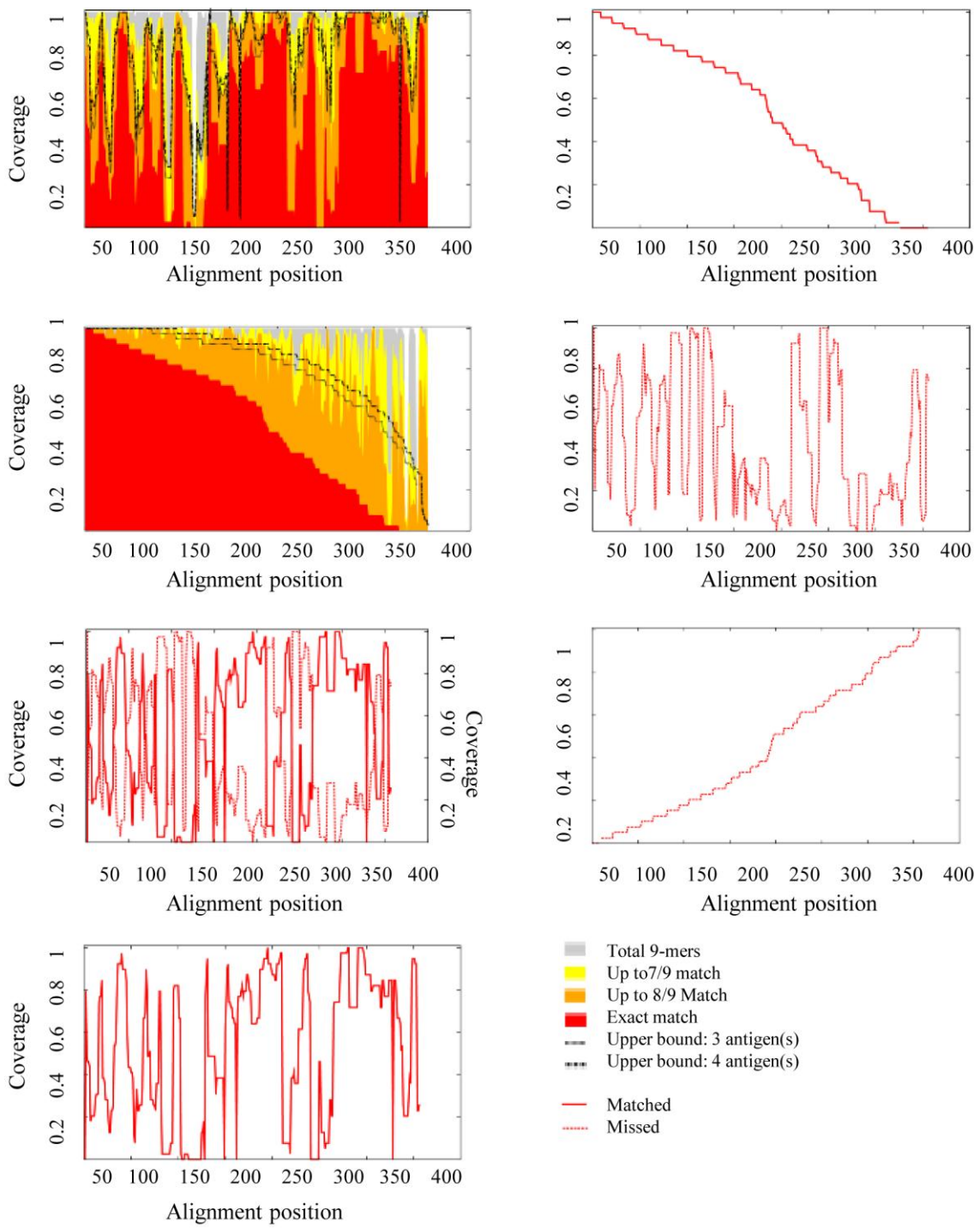


Figure C4. Positional epitope coverage (Posicover) of infecting viral sequences by HIV-1 clade C (C_{CH}, Chinese) peptide reagent sequence. (A) Coverage by position in the alignment of the Gag protein (p17p24p2). **(B)** Coverage ranked. **(C)** Matched 9-mer epitopes and missed 9-mer epitopes coverage by natural position in the alignment. **(D)** Matched 9-mers epitopes ranked by coverage. **(E)** Missed 9-mers epitopes coverage by natural position along the Gag alignment. (F) Missed 9-mers by natural position in along the Gag alignment. **(G)** Missed 9-mers ranked by coverage.

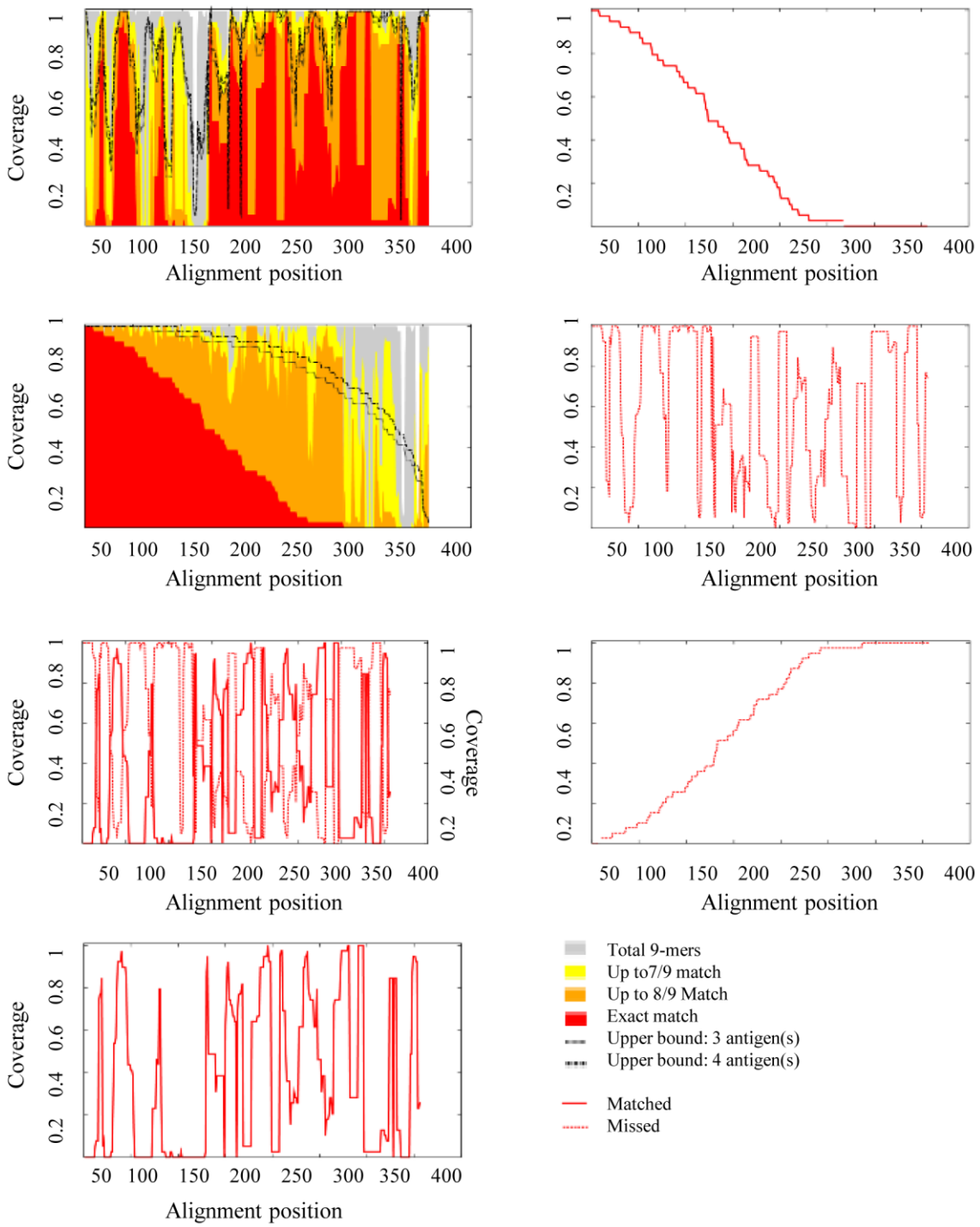


Figure C5. Positional epitope coverage (Posicover) of infecting viral sequences by HIV-1 clade D peptide reagent sequence. (A) Coverage by position in the alignment of the Gag protein (p17p24p2). **(B)** Coverage ranked. **(C)** Matched 9-mer epitopes and missed 9-mer epitopes coverage by natural position in the alignment. **(D)** Matched 9-mers epitopes ranked by coverage. **(E)** Missed 9-mers epitopes coverage by natural position along the Gag alignment. (F) Missed 9-mers by natural position in along the Gag alignment. **(G)** Missed 9-mers ranked by coverage.

Appendix D

Appendix D1. ELISpot layout, quality assurance and test acceptance criteria

Peptide sets

The peptide sets used in the cross-clade study belong to Gag subtype A, Gag subtype D, Gag subtype B, Gag C Du422 (South African subtype C) and Gag subtype C China. A pool and matrix approach was used in which five pools were made up for each of the five peptide variants (Table D1.1) and twenty-four matrices were designed to include all the single Gag peptides, which make up the five different peptide variants (Table D1.2). The consensus Gag B peptides were available as 1mg lyophilized peptides while as the other four peptide-variants were supplied at 500µg/peptide. Single peptides from the five clades were reconstituted to 10 µl aliquots of 10mg/ml stocks and then further reconstituted to 30µg/ml and stored at -80^{0}C. The peptides were used at a final concentration of 1.5µg/ml in the ELISpot assay. Peptides were confirmed by another IFN-γ ELISpot assay using single peptides to identify specific epitope stretches.

CEF peptide pool and PHA were used as positive controls on all tested PBMCs. The CEF peptide pool constituted a panel of 32 8-11-mer CMV, EBV and Flu virus peptide epitopes recognized by CD8$^+$ T cells. The pool was reconstituted at 20µg/ml in 90% PBS/10% DMSO and stored at -80^{0}C.

Assay protocol

The ELISpot assays were conducted for screening of peptide responses using a panel of peptide described in C1 and C2. PBMC from NICD blood donor (QC sample-NICD 063) was used as a positive control sample for each plate. The QC sample had been tested against PHA and CEF and was a known responder and therefore used as a positive control. Participants' PBMCs were thawed as described in Chapter 2 and tested in duplicate against each Gag peptide pool, once against each Gag matrix and twice against CEF and PHA (plate layout in Table D1.3). Eight negative control wells, four positive control wells and four peptide control wells were used.

The negative control wells consisted of six unstimulated PBMC and two unstimulated wells for the QC sample per plate. Each plate also had six wells containing R10 (media only). Positive control wells consisted of two PHA stimulated PBMC and two PHA stimulated wells for the QC sample per plate. The control wells consisted of two CEF stimulated PBMC and two CEF stimulated QC sample on each plate. Peptide confirmations were performed for those peptides that gave a positive response after the screening ELISpot assays.

Test acceptance criteria

A test was regarded as positive when the response was >100 SFU/10^6 PBMC and at least 3 times the mean background response. The positive response in the pool was supposed to match a response in a matrix pool that shares one of the peptides in the pool. Fail criteria for the ELISpot assay included any one or more of the following:

- ✓ Greater than 100 spots in the negative control wells
- ✓ Grater than 5 spots per well for the wells containing media only
- ✓ Less than 400 spots per well for the PHA wells.

Record keeping

An ELISpot worksheet (Table D1.3) was completed with each assay performed. The plates were read by the CTL Immunospot Analyzer and data saved on CD plates. All completed worksheets and ELISpot raw data and calculated data were archived.

Table D1.1. Gag pools

	Pool 1	Pool 2	Pool 3	Pool 4	Pool 5
Gag C1 Du422	1-24	25-48	49-72	73-96	97-120
Gag C2 China	1-24	25-48	49-72	73-96	97-119
Gag B	1-24	25-48	49-72	73-96	97-123
Gag A	1-18	19-36	37-54	55-72	73-90

Table D1.2. Gag matrices

M1	M2	M3	M4	M5	M6	M7	M8
Gag C1-1	Gag C1-2	Gag C1-3	Gag C1-4	Gag C1-5	Gag C1-6	Gag C1-7	Gag C1-8
Gag C1-25	Gag C1-26	Gag C1-27	Gag C1-28	Gag C1-29	Gag C1-30	Gag C1-31	Gag C1-32
Gag C1-49	Gag C1-50	Gag C1-51	Gag C1-52	Gag C1-53	Gag C1-54	Gag C1-55	Gag C1-56
Gag C1-73	Gag C1-74	Gag C1-75	Gag C1-76	Gag C1-77	Gag C1-78	Gag C1-79	Gag C1-80
Gag C1-97	Gag C1-98	Gag C1-99	Gag C1-100	Gag C1-101	Gag C1-102	Gag C1-103	Gag C1-104
Gag C2-1	Gag C2-2	Gag C2-3	Gag C2-4	Gag C2/C1-5	Gag C2/C1-6	Gag C2-7	Gag C2-8
Gag C2-25	Gag C2-26	Gag C2-27	Gag C2-28	Gag C2-29	Gag C2-30	Gag C2-31	Gag C2/C1-32
Gag C2/C1-49	Gag C2-50	Gag C2-51	Gag C2-52	Gag C2-53	Gag C2/C1-54	Gag C2/C1-55	Gag C2/C1-56
Gag C2/C1-73	Gag C2-74	Gag C2-75	Gag C2-76	Gag C2-77	Gag C2/C1-78	Gag C2/C1-79	Gag C2/C1-80
Gag C2/C1-97	Gag C2/C1-98	Gag C2/C1-99	Gag C2/C1-100/B-102	Gag C2/C1-101/B-103	Gag C2/C1-102/B-104	Gag C2/C1-103/B105	Gag C2/C1-104/B106
Gag B-2	Gag B-3	Gag B-4	Gag B-5	Gag B-6	Gag B-7	Gag B-8	Gag B-9
Gag B-26	Gag B-27	Gag B-28	Gag B-29	Gag B-30	Gag B-31	Gag B-32	Gag B-33
Gag B-50	Gag B-51	Gag B-52	Gag B-53	Gag B-54	Gag B-55	Gag B-56	Gag B-57
Gag B-74	Gag B-75	Gag B-76	Gag B-77	Gag B-78	Gag B-79	Gag B-80	Gag B-81
Gag B-99	Gag B-100	Gag B-101	Gag B-102/C1/C2-100	Gag B-103/C1/C2-101	Gag B-104/C1/C2-102	Gag B-105/C1/C2-103	Gag B-106/C1/C2-104
Gag B123	Gag B-95						
Gag A-17/C1/C2-49	Gag A-18/C1-50	Gag A-19/C1-51	Gag A-20	Gag A-21	Gag A-22	Gag A-23	Gag A-24/C1/C2-56
Gag A-41/C1/C2-73	Gag A-42/C1-74	Gag A-43/C1-75	Gag A-44	Gag A-45	Gag A-46	Gag A-47	Gag A-48
Gag A-65	Gag A-66	Gag A-67	Gag A-68	Gag A-69	Gag A-70	Gag A-71	Gag A-72
Gag A-89	Gag A-90						
Gag D-17	Gag D-18	Gag D/A-19/C1-51	Gag D-20	Gag D-21	Gag D-22	Gag D-23	Gag D/A-24/C1/C2-56
Gag D-41	Gag D-42	Gag D-43	Gag D-44	Gag D-45	Gag D/A-46	Gag D-47	Gag D-48
Gag D/A-65	Gag D/A-66	Gag D/A-67					
Gag D-83	Gag D-84	Gag D-85	Gag D-86	Gag D-87	Gag D-88	Gag D-89	Gag D-90
18	20	17	19	18	15	17	14

Table D1.2. ***continued***

	M10	M11	M12	M13	M14	M15	M16
Gag C1-9	Gag C1-10	Gag C1-11	Gag C1-12	Gag C1-13	Gag C1-14	Gag C1-15	Gag C1-16
Gag C1-33	Gag C1-34	Gag C1-35	Gag C1-36	Gag C1-37	Gag C1-38	Gag C1-39	Gag C1-40
Gag C1-57	Gag C1-58	Gag C1-59	Gag C1-60	Gag C1-61	Gag C1-62	Gag C1-63	Gag C1-64
Gag C1-81	Gag C1-82	Gag C1-83	Gag C1-84	Gag C1-85	Gag C1-86	Gag C1-87	Gag C1-88
Gag C1-105	Gag C1-106	Gag C1-107	Gag C1-108	Gag C1-109	Gag C1-110	Gag C1-111	Gag C1-112
Gag C2-9	Gag C2/C1-10	Gag C2/C1-11	Gag C2/C1-12	Gag C2-13	Gag C2-14	Gag C2-15	Gag C2-16
Gag C2/C1-33/B-34	Gag C2/C1-34	Gag C2/C1-35	Gag C2/C1-36	Gag C2/C1-37	Gag C2/C1-38	Gag C2/C1-39	Gag C2/C1-40
Gag C2/C1-57	Gag C2-58	Gag C2-59	Gag C2-60	Gag C2-61	Gag C2-62	Gag C2-63	Gag C2-64
Gag C2/C1-81	Gag C2-82	Gag C2-83	Gag C2-84	Gag C2-85	Gag C2-86	Gag C2-87	Gag C2-88
Gag C2/C1-105/B107	Gag C2/C1-106/B108	Gag C2/C1-107/B109	Gag C2-108	Gag C2-109	Gag C2-110	Gag C2-111	Gag C2/C1-112
Gag B-10	Gag B-11	Gag B-12	Gag B-13	Gag B-14	Gag B-15	Gag B-16	Gag B-17
Gag B-34/C1/C2-33	Gag B-35	Gag B-36	Gag B-37	Gag B-38	Gag B-39	Gag B-40	Gag B-41
Gag B-58	Gag B-59	Gag B-60	Gag B-61	Gag B-62	Gag B-63	Gag B-64	Gag B-65
Gag B-82	Gag B-83	Gag B-84	Gag B-85	Gag B-86	Gag B-87	Gag B-88	Gag B-89
Gag B-107/C1/C2-105	Gag B-108/C1/C2-106	Gag B-109/C1/C2-107	Gag B-110/C2-108	Gag B-111/C2-109	Gag B-112/C2-110	Gag B-113	Gag B-114
Gag A-1	Gag A-2	Gag A-3	Gag A-4	Gag A-5/C1/C2-37	Gag A-6/C1/C2-38	Gag A-7/C1/C2-39	Gag A-8
Gag A-25/C1/C2-57	Gag A-26	Gag A-27	Gag A-28	Gag A-29	Gag A-30/C1-62	Gag A-31/C1-63	Gag A-32/C1-64
Gag A-49	Gag A-50	Gag A-51/C1-83	Gag A-52/C1-84	Gag A-53/C1-85	Gag A-54/C1-86	Gag A-55	Gag A-56
Gag A-73	Gag A-74	Gag A-75	Gag A-76	Gag A-77	Gag A-78	Gag A-79	Gag A-80
Gag D-1/C1/C2-33	Gag D-2/C1/C2-34	Gag D-3/C1/C2-35	Gag D-4/C1/C2-36	Gag D/A-5/C1/C2-37	Gag D/A-6/C1/C2-38	Gag D/A-7/C1/C2-39	Gag D/A-8
Gag D/A-25/C1/C2-57	Gag D-26/C1-58	Gag D-27/C1-59	Gag D-28/C1-60	Gag D-29	Gag D-30	Gag D-31	Gag D-32/C2-64
Gag D-49	Gag D-50	Gag D-51	Gag D-52	Gag D/A-53/C1-85	Gag D-54/C2-86	Gag D-55/C2-87	Gag D-56/C2-88
	Gag D-68/C1-10	Gag D-69/C1-11	Gag D-70/C1-12	Gag D-71	Gag D-72	Gag D-73	Gag D-74
Ψ 13	16	15	16	17	16	18	17

Table D1.2. ***continued***

M17	M18	M19	M20	M21	M22	M23	M24
Gag C1-17	Gag C1-18	Gag C1-19	Gag C1-20	Gag C1-21	Gag C1-22	Gag C1-23	Gag C1-24
Gag C1-41	Gag C1-42	Gag C1-43	Gag C1-44	Gag C1-45	Gag C1-46	Gag C1-47	Gag C1-48
Gag C1-65	Gag C1-66	Gag C1-67	Gag C1-68	Gag C1-69	Gag C1-70	Gag C1-71	Gag C1-72
Gag C1-89	Gag C1-90	Gag C1-91	Gag C1-92	Gag C1-93	Gag C1-94	Gag C1-95	Gag C1-96
Gag C1-113	Gag C1-114	Gag C1-115	Gag C1-116	Gag C1-117	Gag C1-118	Gag C1-119	Gag C1-120
Gag C2-17	Gag C2-18	Gag C2-19	Gag C2-20	Gag C2-21	Gag C2-22	Gag C2-23	Gag C2-24
Gag C2/C1-41	Gag C2/C1-42	Gag C2/C1-43	Gag C2/C1-44	Gag C2/C1-45	Gag C2/C1-46	Gag C2/C1-47	Gag C2/C1-48
Gag C2/C1-65	Gag C2/C1-66	Gag C2/C1-67	Gag C2-68	Gag C2-69	Gag C2-70	Gag C2-71	Gag C2/C1-72
Gag C2-89	Gag C2-90	Gag C2-91	Gag C2-92	Gag C2-93	Gag C2-94	Gag C2-95	Gag C2/C1-96
Gag C2/C1-113	Gag C2/C1-114	Gag C2-115	Gag C2-116	Gag C2-117	Gag C2-118	Gag C2-119	Gag B-1
Gag B-18	Gag B-19	Gag B-20	Gag B-21	Gag B-22	Gag B-23	Gag B-24	Gag B-25
Gag B-42	Gag B-43	Gag B-44	Gag B-45	Gag B-46	Gag B-47	Gag B-48	Gag B-49
Gag B-66	Gag B-67	Gag B-68	Gag B-69	Gag B-70	Gag B-71	Gag B-72	Gag B-73
Gag B-90	Gag B-91	Gag B-92	Gag B-93	Gag B-94	Gag B-96	Gag B-97	Gag B-98
Gag B-115	Gag B-116	Gag B-117	Gag B-118	Gag B-119	Gag B-120	Gag B-121	Gag B-122
Gag A-9	Gag A-10	Gag A-11	Gag A-12	Gag A-13	Gag A-14	Gag A-15	Gag A-16/C1/C2-48
Gag A-33/C1/C2-65	Gag A-34/C1/C2-66	Gag A-35/C1/C2-67	Gag A-36/C1-68	Gag A-37/C1-69	Gag A-38/C1-70	Gag A-39/C1-71	Gag A-40/C1/C2-72
Gag A-57	Gag A-58	Gag A-59	Gag A-60	Gag A-61	Gag A-62	Gag A-63	Gag A-64
Gag A-81	Gag A-82	Gag A-83	Gag A-84	Gag A-85	Gag A-86	Gag A-87	Gag A-88
Gag D/A-9	Gag D/A-10	Gag D-11/C1/C2-43	Gag D-12/C1/C2-44	Gag D-13/C1/C2-45	Gag D-14/C1/C2-46	Gag D-15	Gag D-16
Gag D/A-33/C1/C2-65	Gag D/A-34/C1/C2-66	Gag D/A-35/C1/C2-67	Gag D/A-36/C1-68	Gag D/A-37/C1-69	Gag D/A-38/C1-70	Gag D/A-39/C1-71	Gag D-40
Gag D-57	Gag D-58	Gag D/A-59	Gag D/A-60	Gag D/A-61	Gag D/A-62	Gag D/A-63	Gag D/A-64
Gag D-75	Gag D-76	Gag D-77	Gag D-78	Gag D-79	Gag D-80	Gag D-81	Gag D-82
17	17	17	18	18	18	19	17

Common peptides between C_1 (Du422) and C_2 (Chinese)
No peptide
Number in the last raw of each table indicate the number of different peptides

Table D1.3. ELISpot screening plate layout

	1	2	3	4	5	6	7	8	9	10	11	12
A	Gag C1 Pool 1	Gag C1 Pool 1	Gag C2 Pool 1	Gag C2Pool 1	Gag B Pool 1	Gag B Pool 1	Gag A Pool 1	Gag A Pool 1	Gag D Pool 1	Gag D Pool 1	Cells + Media	Cells + Media
B	Gag C1 Pool 2	Gag C1 Pool 2	Gag C2 Pool 2	Gag C2 Pool 2	Gag B Pool 2	Gag B Pool 2	Gag A Pool 2	Gag A Pool 2	Gag D Pool 2	Gag D Pool 2	Cells + Media	Cells + Media
C	Gag C1 Pool 3	Gag C1 Pool 3	Gag C2 Pool 3	Gag C2 Pool 3	Gag B Pool 3	Gag B Pool 3	Gag A Pool 3	Gag A Pool 3	Gag D Pool 3	Gag D Pool 3	Cells + Media	Cells + Media
D	Gag C1Pool 4	Gag C1 Pool 4	Gag C2 Pool 4	Gag C2 Pool 4	Gag B Pool 4	Gag B Pool 4	Gag A Pool 4	Gag A Pool 4	Gag D Pool 4	Gag D Pool 4	Media Only	Media Only
E	Gag C1 Pool 5	Gag C1 Pool 5	Gag C2 Pool 5	Gag C2 Pool 5	Gag B Pool 5	Gag B Pool 5	Gag A Pool 5	Gag A Pool 5	Gag D Pool 5	Gag D Pool 5	Media Only	Media Only
F	Matrix 1	Matrix 2	Matrix 3	Matrix 4	Matrix 5	Matrix 6	Matrix 7	Matrix 8	Matrix 9	Matrix 10	Matrix 11	Matrix 12
G	Matrix 13	Matrix 14	Matrix 15	Matrix 16	Matrix 17	Matrix 18	Matrix 19	Matrix 20	Matrix 21	Matrix 22	Matrix 23	Matrix 24
H	CEF	CEF	PHA	PHA	Media Only	Media Only	QC Cells + Media	QC Cells + Media	QC + CEF	QC + CEF	QC + PHA	QC + PHA

Pt ID _____

Cells/well _____

QC ID _____

	Reagent number	Initials
Coating ab		
Secondary ab		
Streptavidin		
NovaRed		
Blocking media		
FCS		
Coating date		

	Start Time	End Time	Initials
Blocking	:	:	
O/N incubation	:	:	
2nd ab	:	:	
Streptavidin	:	:	
NovaRed	:	:	
Plate number			
Plate read by		Date	

Comments _____

This table shows the ELISpot layout utilized for those individuals screened for HIV-1-specific T-cell responses to five HIV gag peptide reagents simultaneously, that is matched data was available for these peptide reagents in the individuals screened using this layout. Also, the data for responses to clade CDu422 peptide reagent was combined with that from individuals screened for consensus group M and subtype C_{Du422} peptide reagents simultaneously (Table D2.3) and unmatched data analysis performed.

Appendix D2. ELISpot assay layout and acceptance criteria

The pooled peptides were arranged in 12 Gag pools, five Nef pools and 12 Matrix pools. Peptide pools were tested in duplicate (except for individuals screened for both consensus group M and clade C_{Du422} Gag and Nef HIV-specific T-cell responses that were performed in single wells) and matrices in single wells (see ELISpot plate layout worksheet, Table D2.4).

Stocks of peptide pools (40µg/ml, 25µl aliquots) were made from 1mg of lyophilized peptides. For matrix pools stocks, 135µl of R10 (RPMI with 10% FCS) was added to each aliquot, and for Gag and Nef peptide pools, 225µl of R10 was added per tube.

The CEF peptide pool was used as a positive control as in previous assays (Appendix D1) for each sample plated as well as for testing of the QC (a known CEF responder) sample on each plate. These peptide pools consist of a panel of 32 8-11 mers CMV, EBV and Flu virus CD8+ T-cell epitopes. The 60µl aliquots of these peptides were provided at 20µg/ml and were used at a final concentration of 1µg/ml. Test acceptance criteria were the same as for assays in Appendix D1.

Table D2.1. Peptide for different Pools

	1	2	3	4	5	6	7	8	9	10	11	12
Gag	1−12	13−24	25−33	34−45	46−57	58−69	70−81	82−93	94−95	96−107	108−119	120−129
Nef	1−12	13−24	25−36	37−48	49−53							

Table D2.2. Peptide Pool and Matrix configuration

	Matrix 1	Matrix 2	Matrix 3	Matrix 4	Matrix 5	Matrix 6	Matrix 7	Matrix 8	Matrix 9	Matrix 10	matrix 11	Matrix 12
Gag Pool 1	Gag 1	Gag 2	Gag 3	Gag 4	Gag 5	Gag 6	Gag 7	Gag 8	Gag 9	Gag 10	Gag 11	Gag 12
Gag Pool 2	Gag 13	Gag 14	Gag 15	Gag 16	Gag 17	Gag 18	Gag 19	Gag 20	Gag 21	Gag 22	Gag 23	Gag 24
Gag Pool 3	Gag 25	Gag 26	Gag 27	Gag 28	Gag 29	Gag 30	Gag 31	Gag 32	Gag 33			
Gag Pool 4	Gag 34	Gag 35	Gag 36	Gag 37	Gag 38	Gag 39	Gag 40	Gag 41	Gag 42	Gag 43	Gag 44	Gag 45
Gag Pool 5	Gag 46	Gag 47	Gag 48	Gag 49	Gag 50	Gag 51	Gag 52	Gag 53	Gag 54	Gag 55	Gag 56	Gag 57
Gag Pool 6	Gag 58	Gag 59	Gag 60	Gag 61	Gag 62	Gag 63	Gag 64	Gag 65	Gag 66	Gag 67	Gag 68	Gag 69
Gag Pool 7	Gag 70	Gag 71	Gag 72	Gag 73	Gag 74	Gag 75	Gag 76	Gag 77	Gag 78	Gag 79	Gag 80	Gag 81
Gag Pool 8	Gag 82	Gag 83	Gag 84	Gag 85	Gag 86	Gag 87	Gag 88	Gag 89	Gag 90	Gag 91	Gag 92	Gag 93
Gag Pool 9	Gag 94	Gag 95										
Gag Pool 10	Gag 96	Gag 97	Gag 98	Gag 99	Gag 100	Gag 101	Gag 102	Gag 103	Gag 104	Gag 105	Gag 106	Gag 107
Gag Pool 11	Gag 108	Gag 109	Gag 110	Gag 111	Gag 112	Gag 113	Gag 114	Gag 115	Gag 116	Gag 117	Gag 118	Gag 119
Gag Pool 12	Gag 120	Gag 121	Gag 122	Gag 123	Gag 124	Gag 125	Gag 126	Gag 127	Gag 128	Gag 129		

Nef Pool 1	Nef 1	Nef 2	Nef 3	Nef 4	Nef 5	Nef 6	Nef 7	Nef 8	Nef 9	Nef 10	Nef 11	Nef 12
Nef Pool 2	Nef 13	Nef 14	Nef 15	Nef 16	Nef 17	Nef 18	Nef 19	Nef 20	Nef 21	Nef 22	Nef 23	Nef 24
Nef Pool 3	Nef 25	Nef 26	Nef 27	Nef 28	Nef 29	Nef 30	Nef 31	Nef 32	Nef 33	Nef 34	Nef 35	Nef 36
Nef Pool 4	Nef 37	Nef 38	Nef 39	Nef 40	Nef 41	Nef 42	Nef 43	Nef 44	Nef 45	Nef 46	Nef 47	Nef 48
Nef Pool 5	Nef 49	Nef 50	Nef 51	Nef 52	Nef 53							

Table D2.3. ELISpot plate layout for unmatched data

	1	2	3	4	5	6	7	8	9	10	11	12
A	Gag Pool 1	Gag Pool 5	Gag Pool 9	Nef Pool 1	Nef Pool 5	Matrix 7	Cells + Media	QC Cells + Media				
B	Gag Pool 1	Gag Pool 5	Gag Pool 9	Nef Pool 1	Nef Pool 5	Matrix 8	Cells + Media	QC Cells + Media				
C	Gag Pool 2	Gag Pool 6	Gag Pool 10	Nef Pool 2	Matrix 1	Matrix 9	Cells + Media	QC Cells + Media				
D	Gag Pool 2	Gag Pool 6	Gag Pool 10	Nef Pool 2	Matrix 2	Matrix 10	Cells + Media	QC Cells + Media				
E	Gag Pool 3	Gag Pool 7	Gag Pool 11	Nef Pool 3	Matrix 3	Matrix 11	CEF	QC CEF				
F	Gag Pool 3	Gag Pool 7	Gag Pool 11	Nef Pool 3	Matrix 4	Matrix 12	CEF	QC CEF				
G	Gag Pool 4	Gag Pool 8	Gag Pool 12	Nef Pool 4	Matrix 5	Media Only	PHA	QC PHA				
H	Gag Pool 4	Gag Pool 8	Gag Pool 12	Nef Pool 4	Matrix 6	Media Only	PHA	QC PHA				

		Reagent numb	Initials
Pt ID	Coating ab		
Cells/well	Secondary ab		
	Streptavidin		
QC ID	NovaRed		
	Blocking media		
	FCS		
	Coating date		

	Start Time	End Time	Initials
Blocking	:	:	
O/N incubation	:	:	
2nd ab	:	:	
Streptavidin	:	:	
NovaRed	:	:	
Plate number			
Plate read by		Date	

Comments

This table shows the ELISpot layout utilized for those individuals screened for consensus group M Gag and HIV-1-specific T-cell responses. Therefore information from this screening with subsequent confirmatory assays was used for data analysis involving HIV-1 Gag and Nef consensus group M T-cell responses only.

Table D2.4. ELISpot plate layout for matched data

	1	2	3	4	5	6	7	8	9	10	11	12
A	Gag Pool 1	Gag Pool 9	Nef Pool 5	Matrix 8	Gag Pool 1	Gag Pool 9	Matrix 5	Cells + Media				
B	Gag Pool 2	Gag Pool 10	Matrix 1	Matrix 9	Gag Pool 2	Gag Pool 10	Matrix 6	Cells + Media				
C	Gag Pool 3	Gag Pool 11	Matrix 2	Matrix 10	Gag Pool 3	Gag Pool 11	Matrix 7	Cells + Media				
D	Gag Pool 4	Gag Pool 12	Matrix 3	Matrix 11	Gag Pool 4	Gag Pool 12	Matrix 8	Cells + Media				
E	Gag Pool 5	Nef Pool 1	Matrix 4	Matrix 12	Gag Pool 5	Matrix 1	Matrix 9					
F	Gag Pool 6	Nef Pool 2	Matrix 5		Gag Pool 6	Matrix 2	Matrix 10					
G	Gag Pool 7	Nef Pool 3	Matrix 6	Media only	Gag Pool 7	Matrix 3	CEF	PHA				
H	Gag Pool 8	Nef Pool 4	Matrix 7	Media only	Gag Pool 8	Matrix 4	CEF	PHA				

Pt ID _____

Cells/well _____

QC ID _____

	Reagent numb	Initials
Coating ab		
Secondary ab		
Streptavidin		
NovaRed		
Blocking media		
FCS		
Coating date		

	Start Time	End Time	Initials
Blocking	:	:	
O/N incubation	:	:	
2nd ab	:	:	
Streptavidin	:	:	
NovaRed	:	:	
Plate number			
Plate read by		Date	

Comments _____

This table shows the ELISpot layout utilized for those individuals screened for consensus group M and clade C HIV-1-specific T-cell responses simultaneously, that is matched data was available for individuals screened using this layout.

Appendix D3. Clinical data of study participants with matched and unmatched data

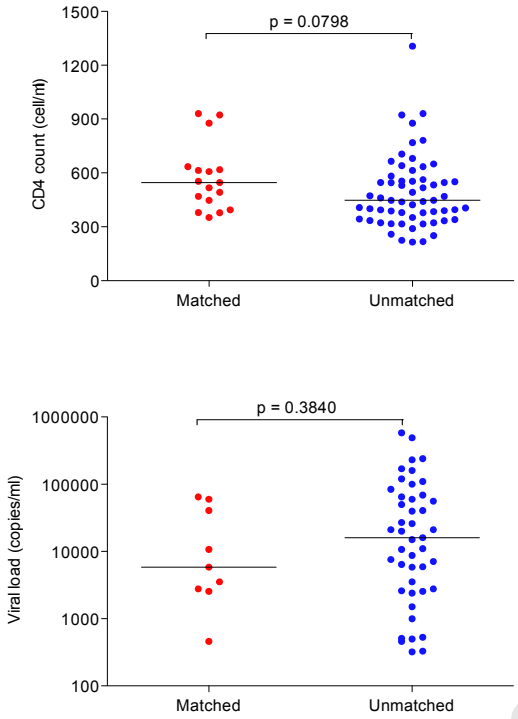


Figure D3.1. Clical characteristics of study participants for matched and unmatched data. (A) CD4 count (cells/µl). **(B)** Viral load (copies/ml). Mann-Whitney test for unmatched differences between median was used at 95% significance level (p>0.05 was considered not significant).

Appendix D4. Peptide maps for peptides reactive in the study that were classified as beneficial peptides in Mothe *et al* (2011) study.

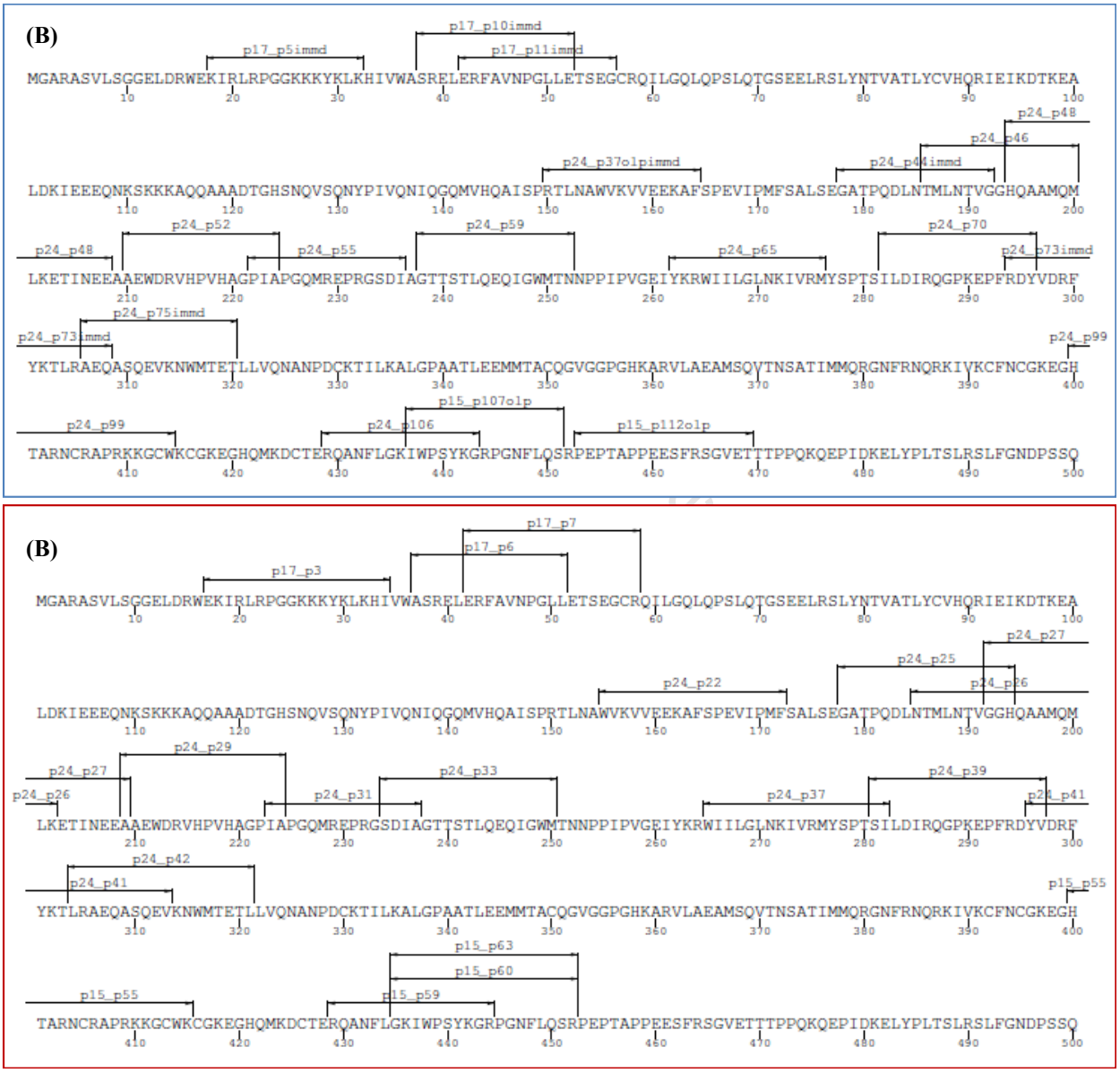


Figure D4.1. Reactive peptide maps. (A) Peptide maps showing peptides identified as reactive in this study and had corresponding peptides/regions whose recognition was classified as beneficial in the Mothe *et al* (2011) study. **(B)** Peptides that were classified as beneficial peptides (protective ratio, PR>1) in Mothe *et al* (2011) study. Peptides with immd after the peptide number for example p24_p37immd denotes that the peptides was classified as immunodominant in this study and olp denotes overlapping with another peptide but with potential for more than one epitope for.

Appendix E. Functional studies

Appendix E1. Cytokine panel

Table E1.1. Cytokine panel and antibodies.

Marker	Type and function	Antibody	Laser
Viability marker			
Vivid	Identification of live/dead cells	CD14/CD19 Pac Blue	
CD4	T-cell marker	FITC	
CD8	T-cell marker	PerCP Cy5.5	
CD3	T-cell marker	APC Cy7	
IFN-γ	Cytokine	Alexa Flour 700	
IL-2	Cytokine	APC	
TNF-α	Cytokine	PE Cy7	
MIP-1β	Chemokine	PE	

Staining protocol

1. Add 50µl PBS and spin at 2100 rpm at 4^0C for 3 minutes and plate blotted on paper towel for 1 second and add 150µl of PBS and spin at same speed for 3 minutes
2. Add 50µl of Vivid, re-suspend and incubate in the DARK for 20 minutes
3. Add 100µl of FACS wash and spin at 2100 rpm for 3 minutes
4. Blot on paper towel for 1 second and add 150µl of FACS wash, re-suspend and spin at same speed for 3 minutes
5. Add 50µl of Surface stain mix and re-suspend and incubate for 20 minutes in the DARK
6. Add 100µl of FACS wash and spin at 2100 rpm for 3 minutes at 4^0C
7. Blot on paper towel for 1 second and add 150µl of FACS wash, re-suspend and spin at same speed
8. Add 100µl of Cytofix/Cytoperm, re-suspend and incubate in the DARK for 20 minutes
9. Add 50µl of Perm Wash and spin at 2100 rpm for 3 minutes at 4^0C
10. Blot on paper towel and add 150µl of Perm Wash, re-suspend and spin at same speed
11. Add 50µl of Intracellular staining mix and incubate in the dark for 20 minutes
12. Add 100µl of Perm wash and spin at 2100 rpm for 3 minutes at 4^0C
13. Blot on paper towel and add 150µl of Perm wash, re-suspend and spin at same speed
14. Blot on paper towel and add 150µl of Cell fix
15. Transfer fixed cell to FACS tubes in their respective stimulations well labeled for acquisition.

Flow cytometry results layouts for each individual participant

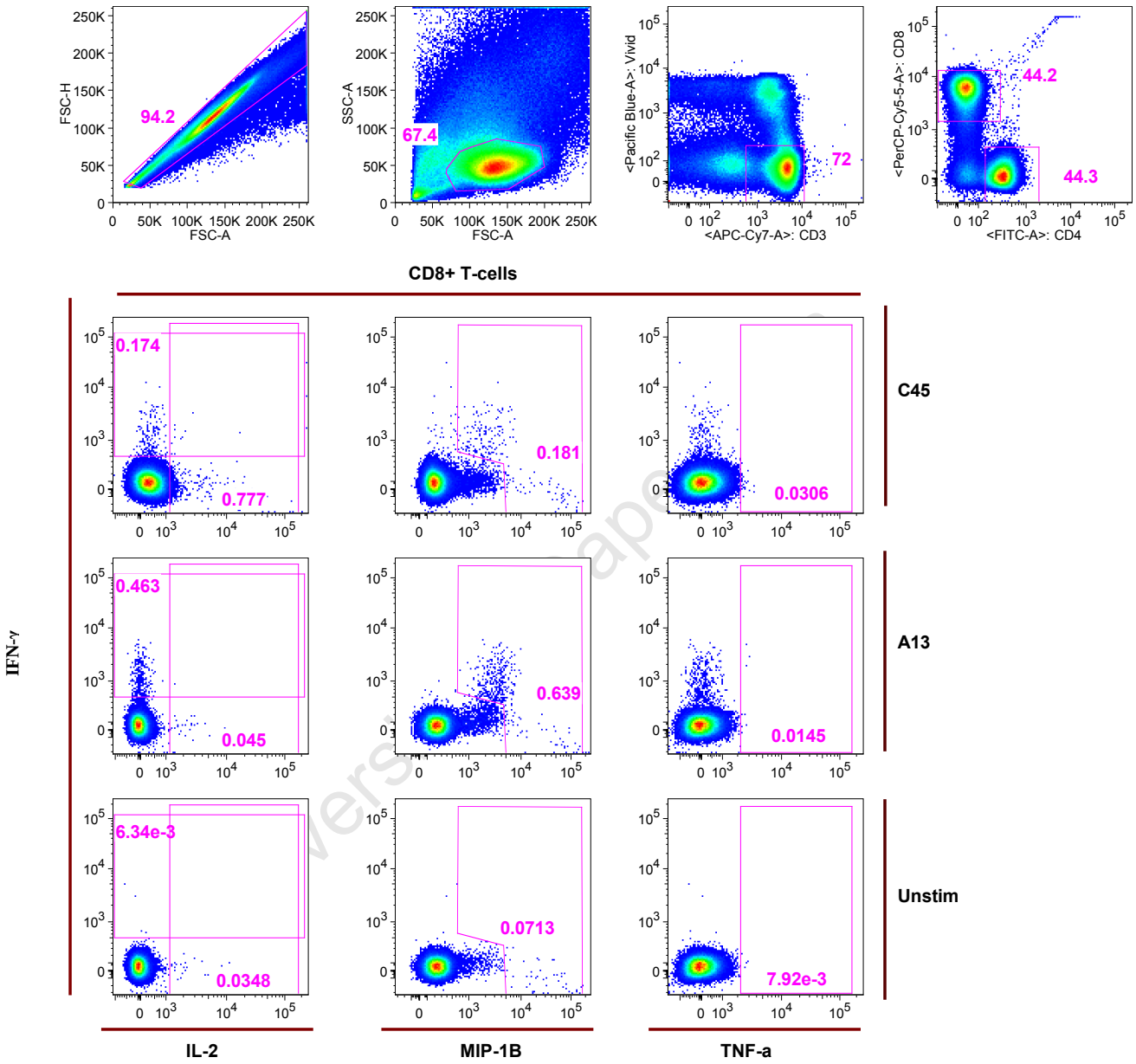


Figure E1.1. Cytokine profile of CD8+ T-cells from study individual ZA023 after stimulation with peptide variants C45 and A13.

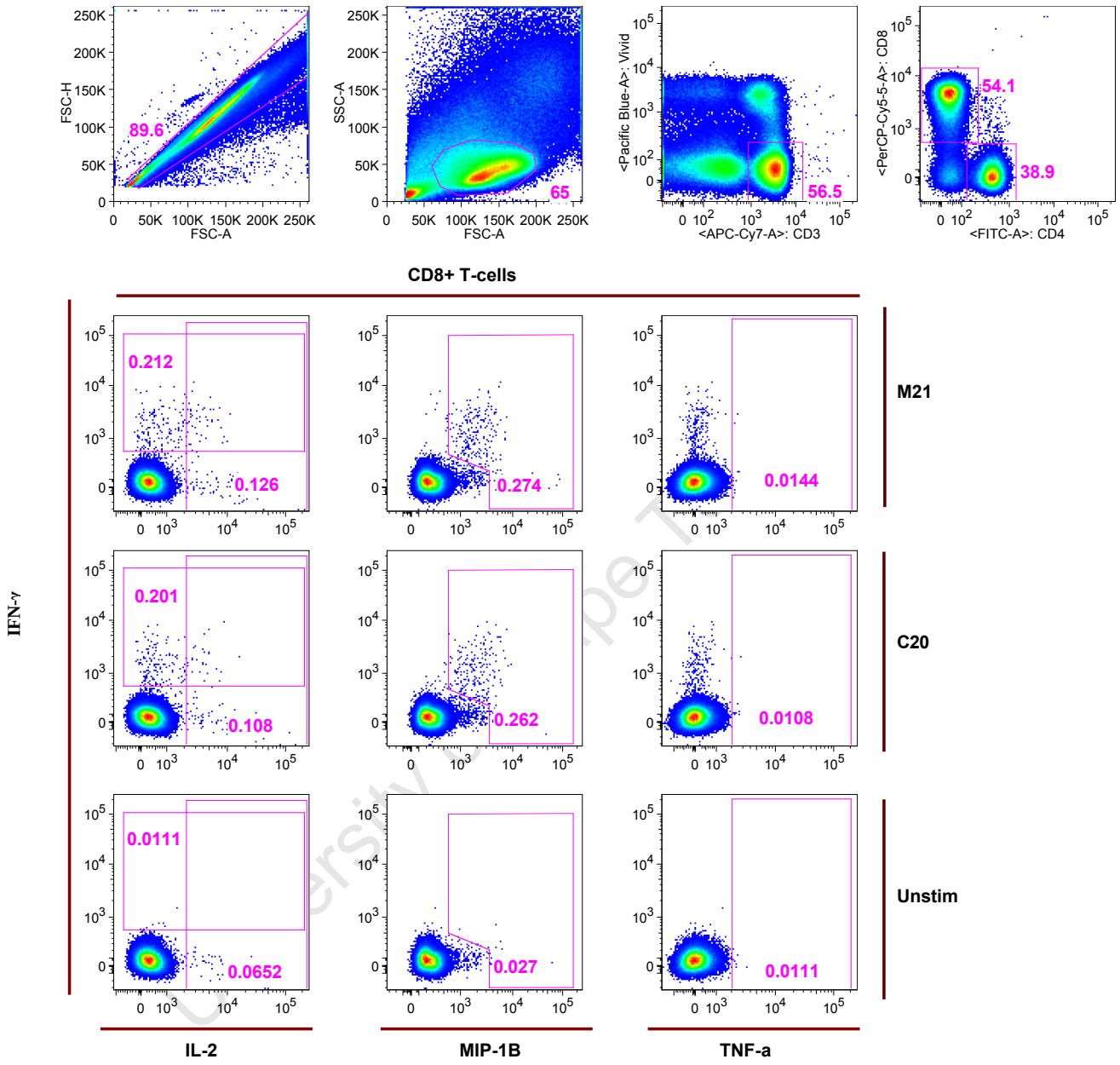


Figure E1.2. Cytokine profile of CD8+ T-cells from study individual ZA033 after stimulation with for peptide variants M21 and C20.

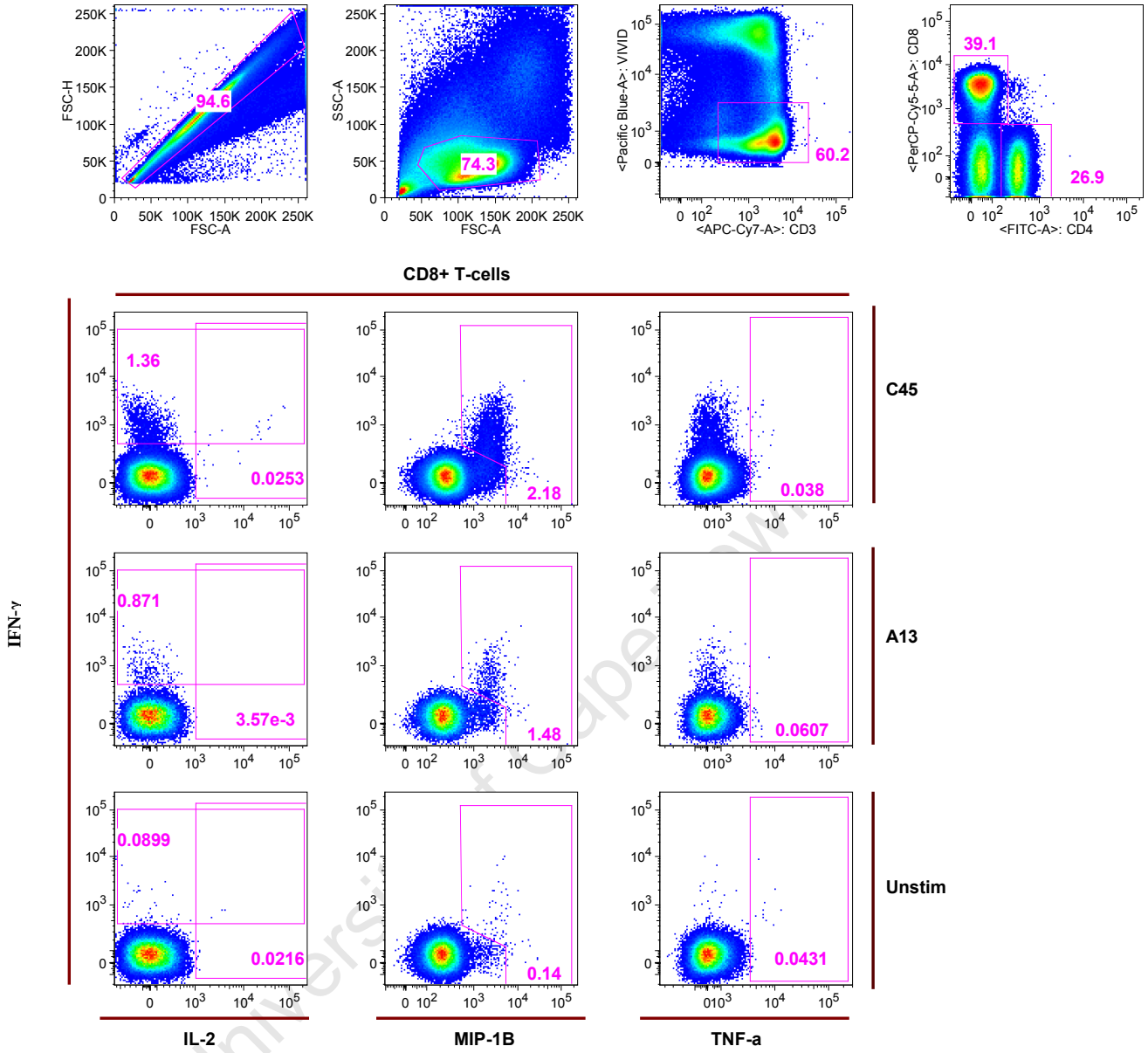


Figure E1.3 Cytokkine profile of CD8+ T-cells from study individual ZA006 after stimularion with peptide variants C45 and A13.

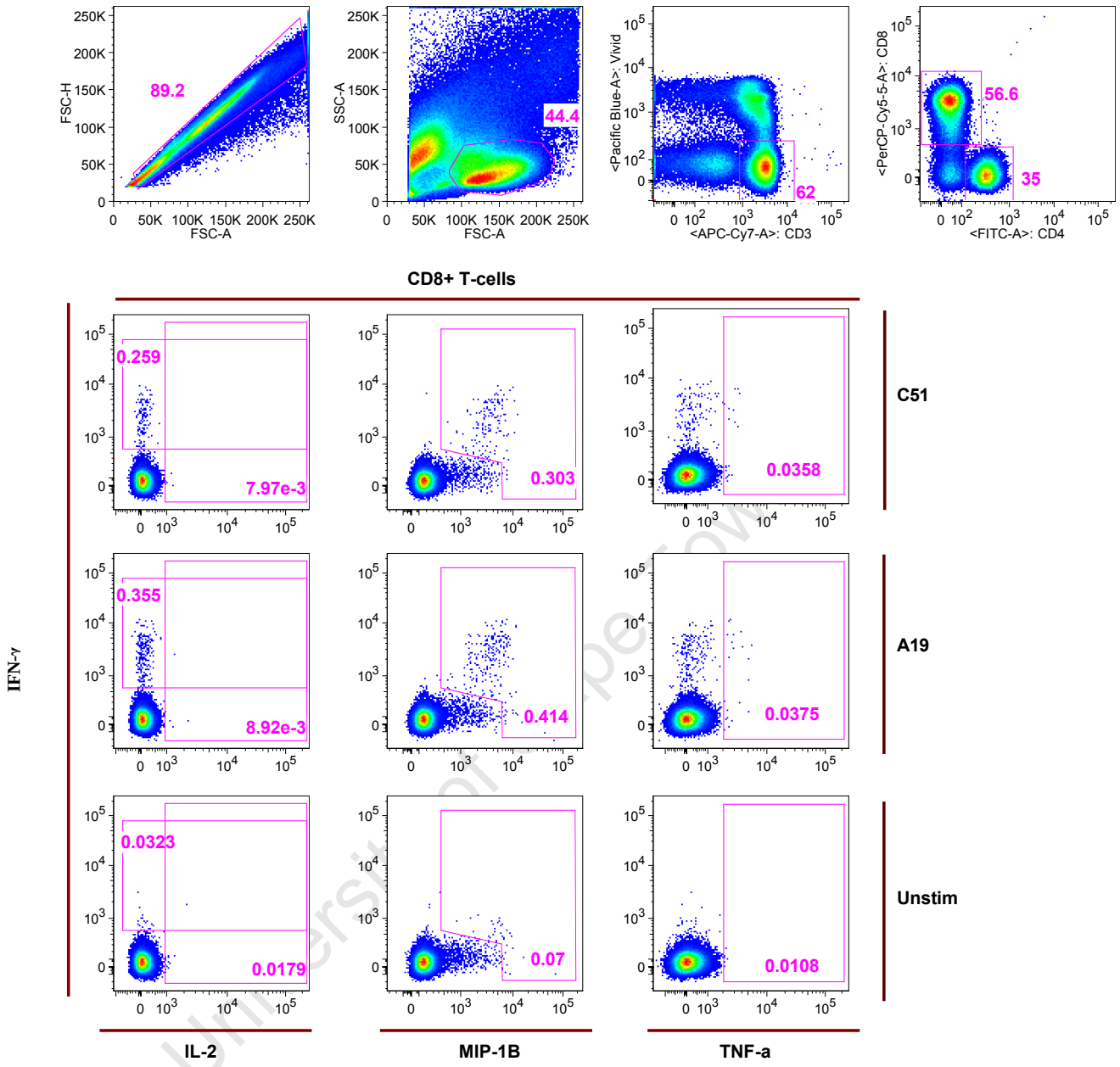


Figure E1.4. Cytokine profile of CD8+ T-cells from study individual ZA021 after stimulation with peptide variants C51 and A19.

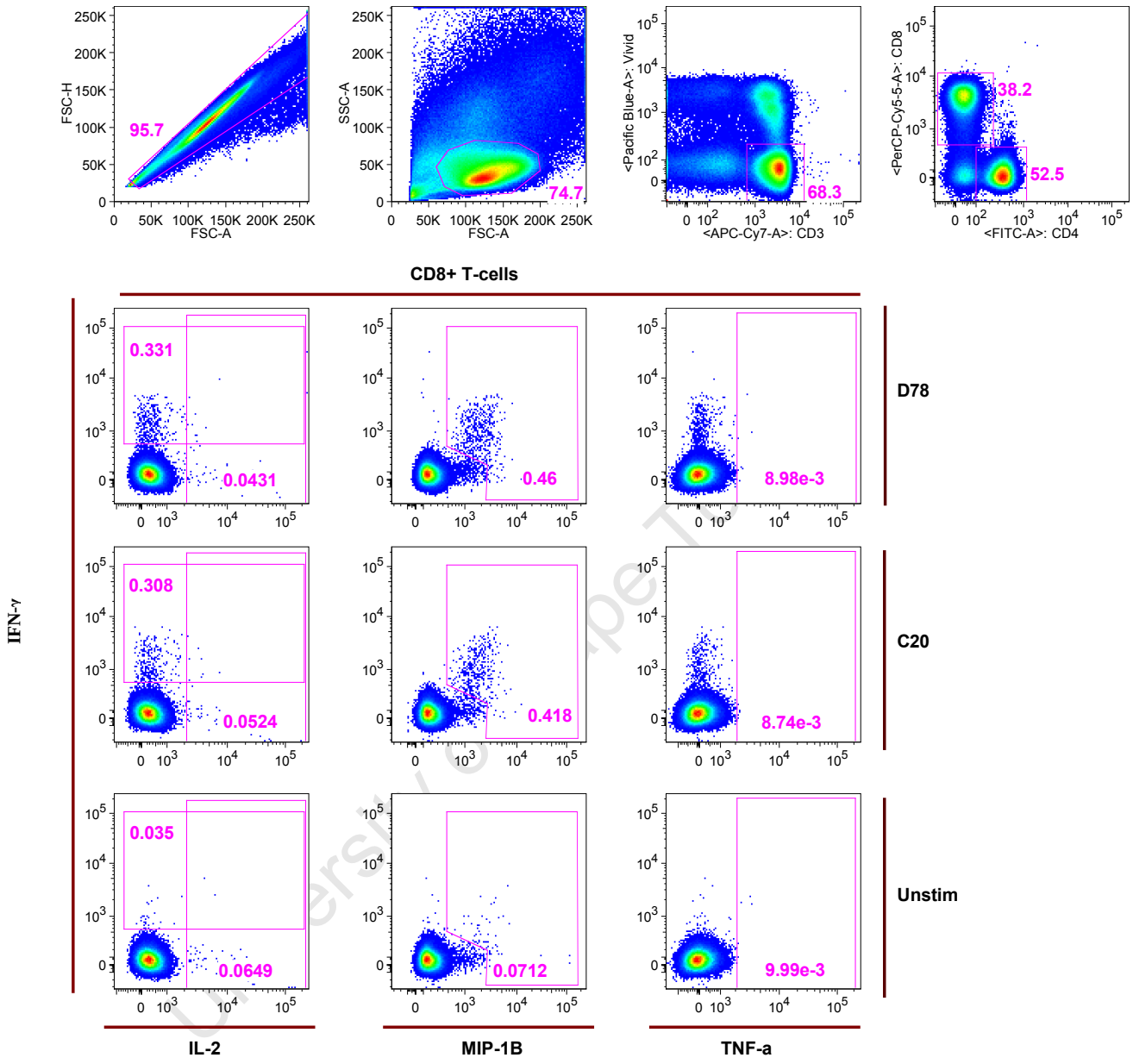


Figure E1.5. Cytokine profile of CD8+ T-cells from study individual ZA028 after stimulation with peptide variants D78 and C20.

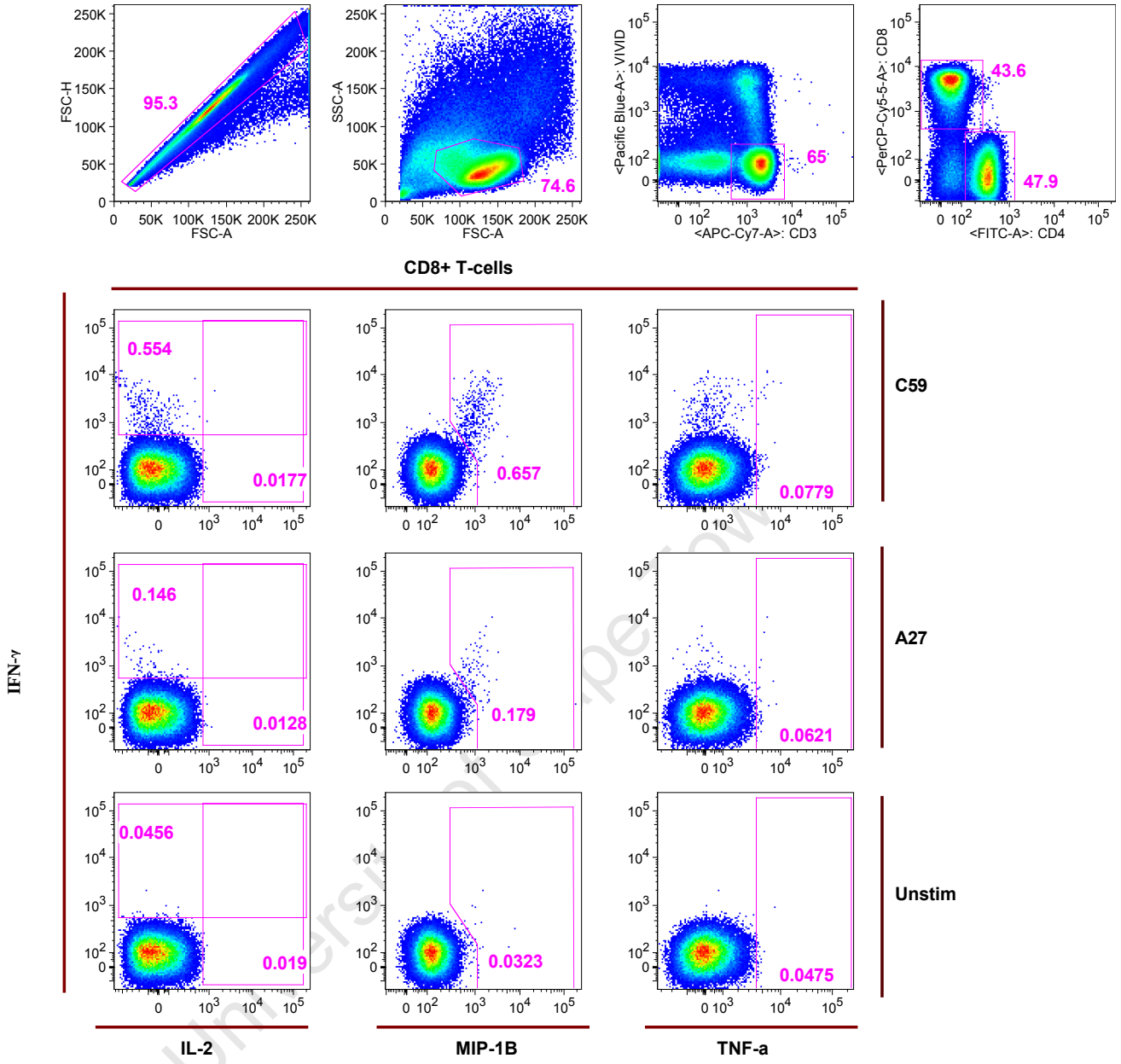


Figure E1.6. Cytokine profile of CD8+ T-cells from study individual ZA030 after stimulation with peptide variants C59 and A27.

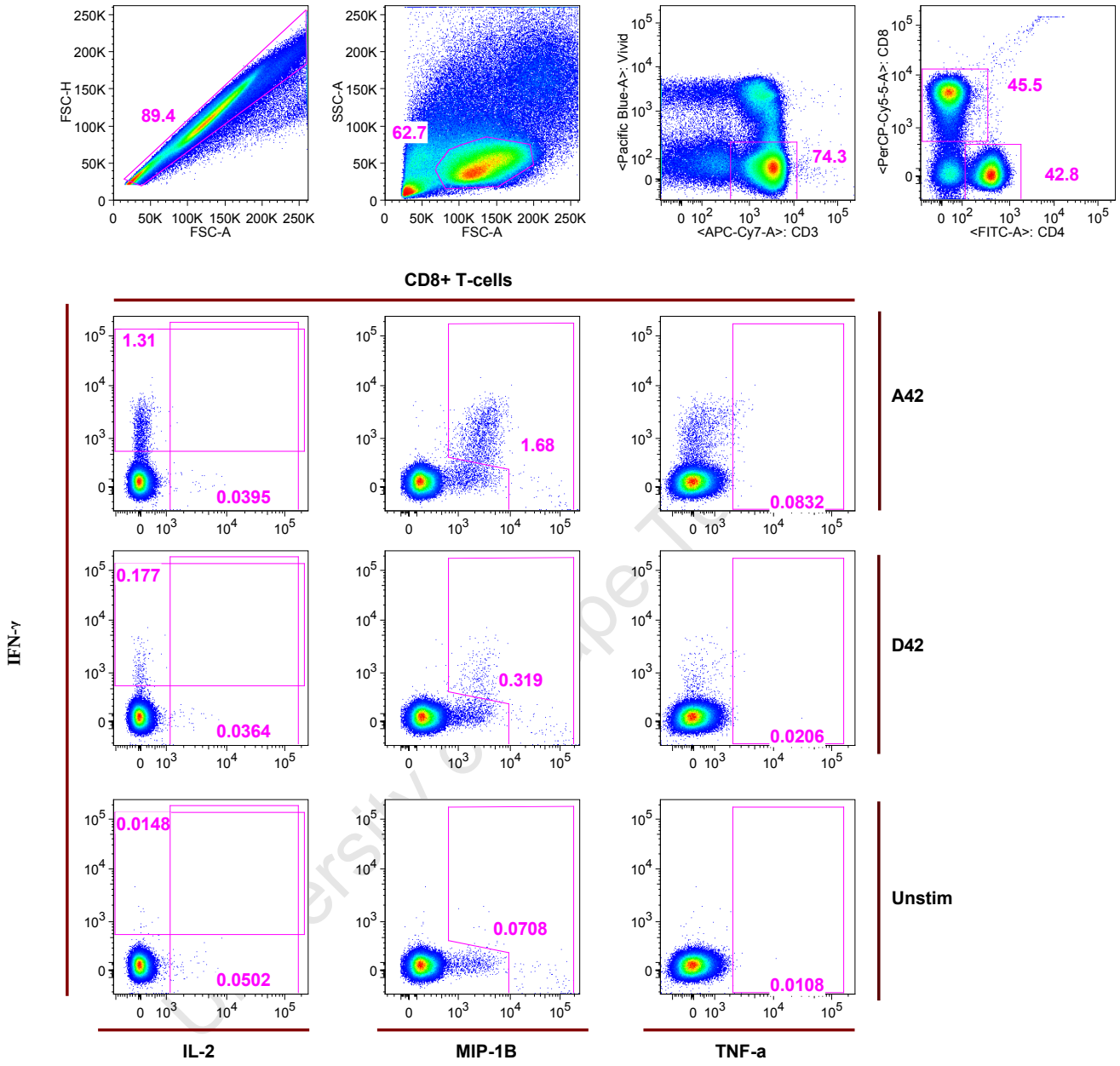


Figure E1.7. Cytokine profile of CD8+ T-cels from study individual ZA029 after stimulation with peptide variants A42 and D42.

Appendix E2. Cytotoxic panel

Table E2.1. Cytotoxic panel and antibodies.

Marker	Type/function	Antibody	Laser
Viability marker			
Vivid	Identification of live/dead cells	CD14/CD19 Pac Blue	violet
CD4	T-cell marker	PE Cy7	blue
CD8	T-cell marker	PerCP Cy5.5	blue
CD3	T-cell marker	APC Cy7	red
IFN-γ	Cytokine	Alexa Flour 700	red
CD107a	Degranulation marker	FITC	blue
Perforin	Cytotoxic molecule	PE	blue
Granzyme B	Cytotoxic molecule	Alexa Flour 647	blue

Staining protocol

The same protocol for the cytokine panel (Appendix E1) was followed for the cytotoxic panel except that CD107a was included during the stimulation step at the start of the experiment.

Flow cytometry results layouts for each individual participant

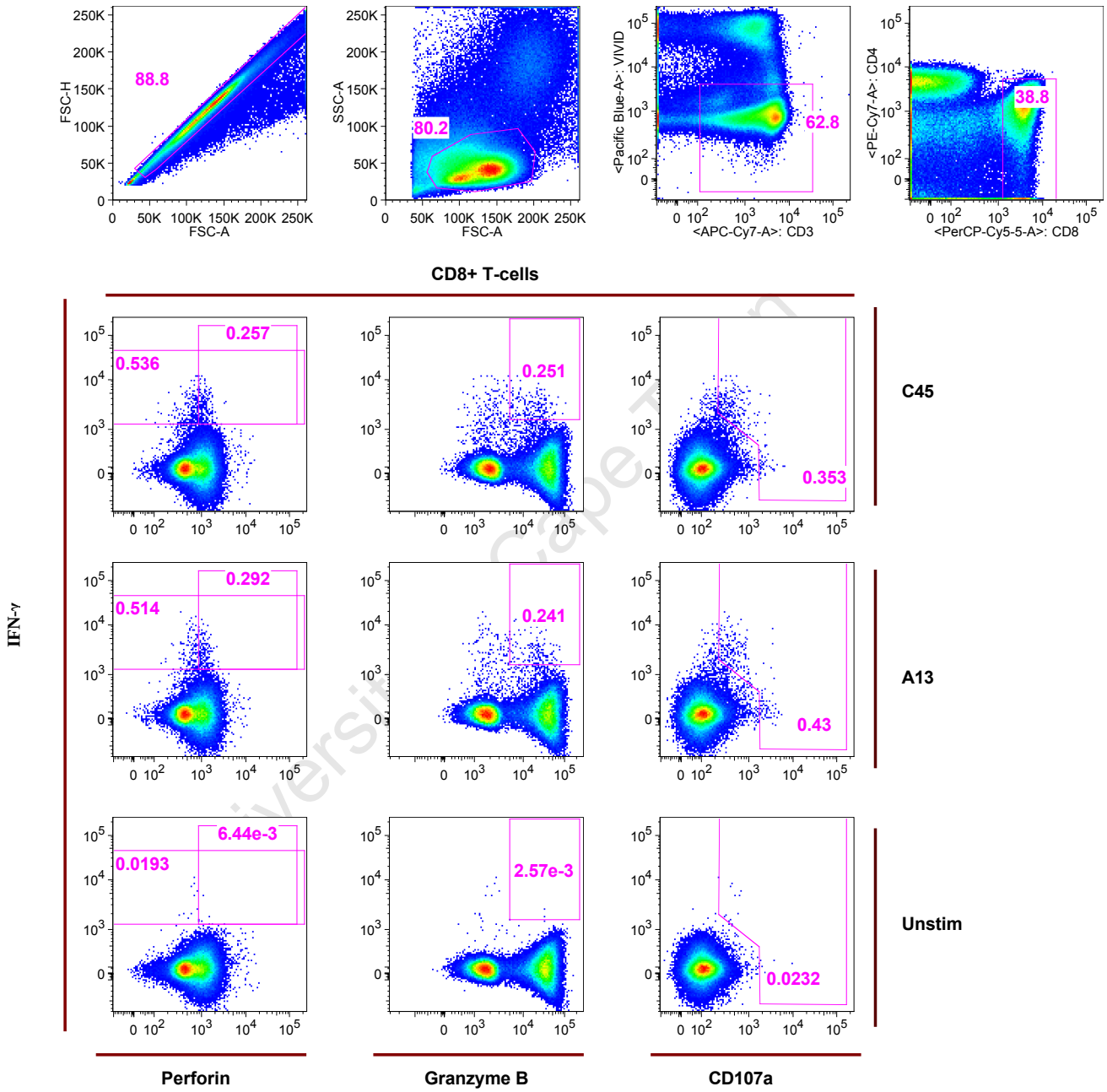


Figure E2.1. Cytpotoxic potential of CD8+ T-cells recognizing peptide variants C45 and A13 in study individual ZA006.

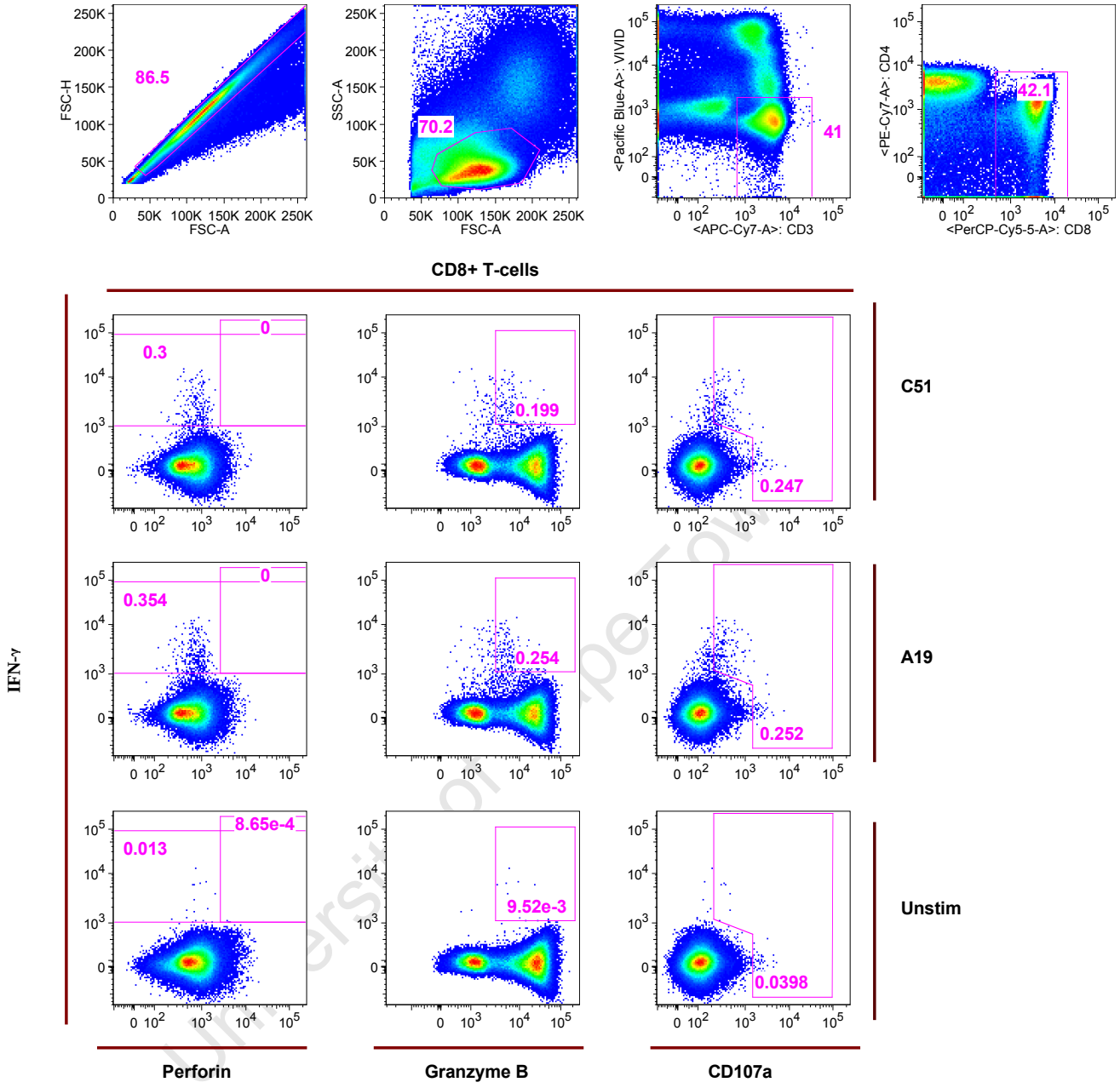


Figure E2.2. Cytpotoxic potential of CD8+ T-cells recognizing peptide variants C51 and A19 in study individual ZA021.

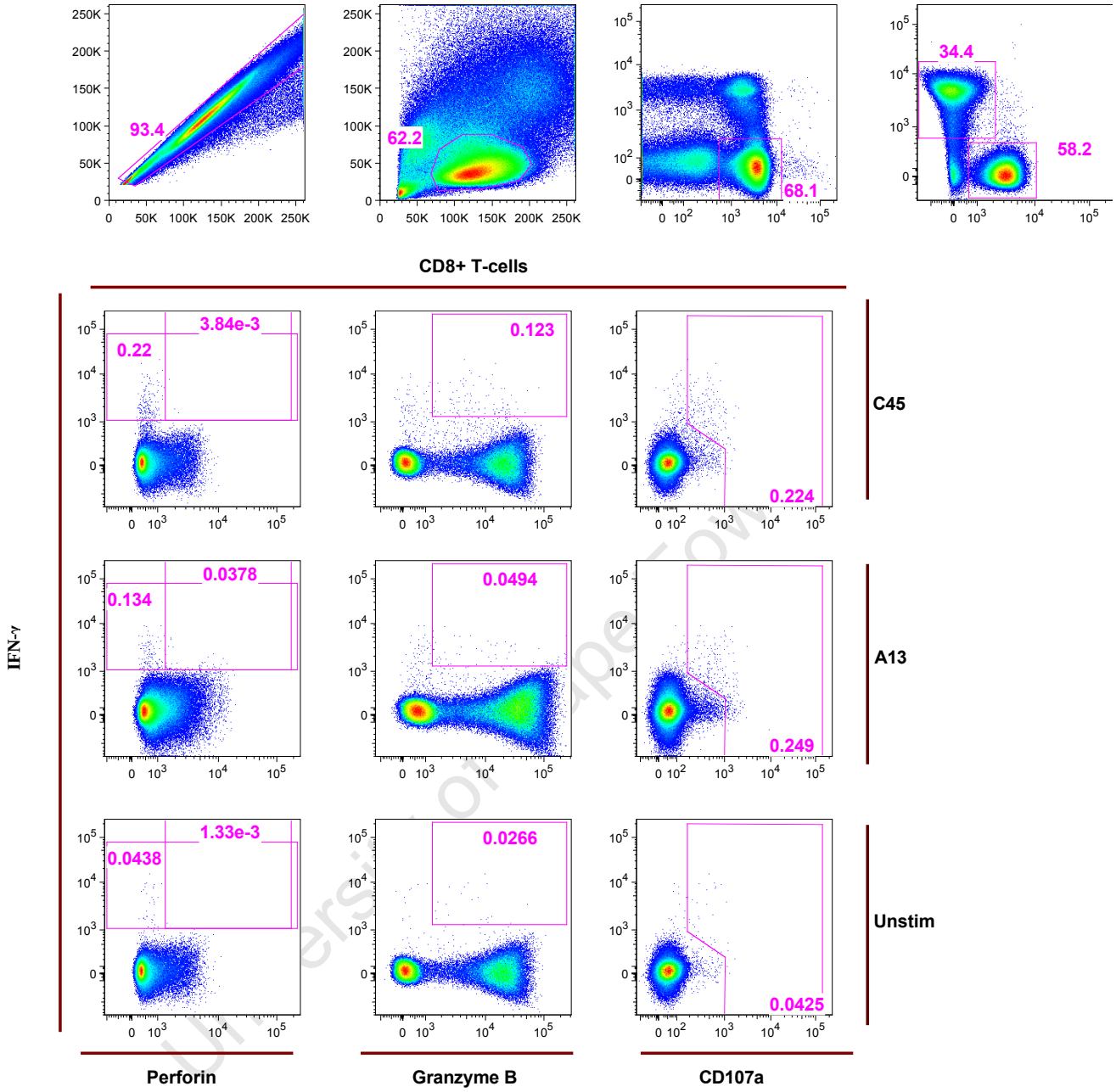


Figure E2.3. Cytpotoxic potential of CD8+ T-cells recognizing peptide variants C45 and A13 in study individual ZA023.

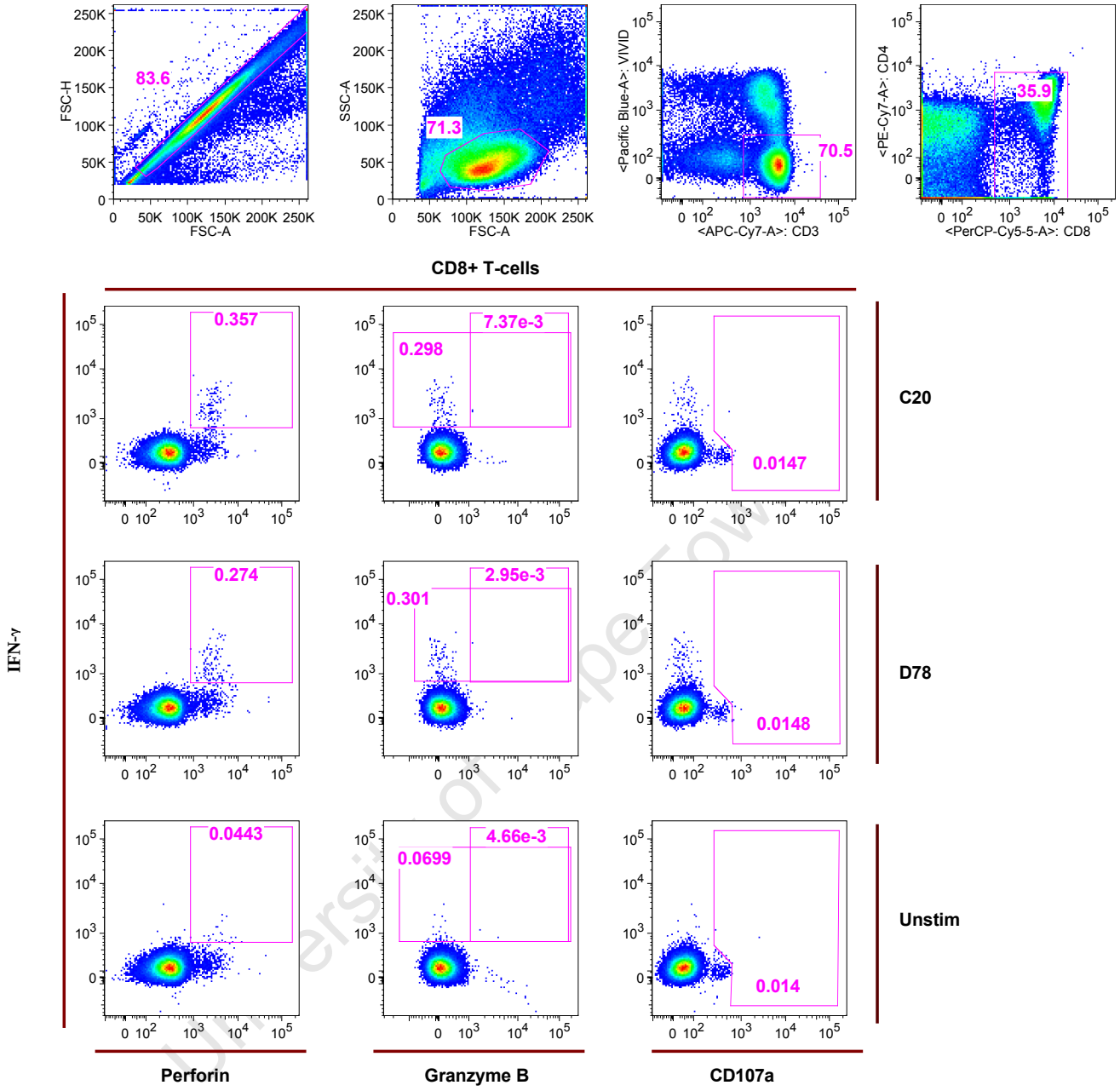


Figure E2.4. Cytpotoxic potential of CD8+ T-cells recognizing peptide variants C20 and D78 in study individual ZA028.

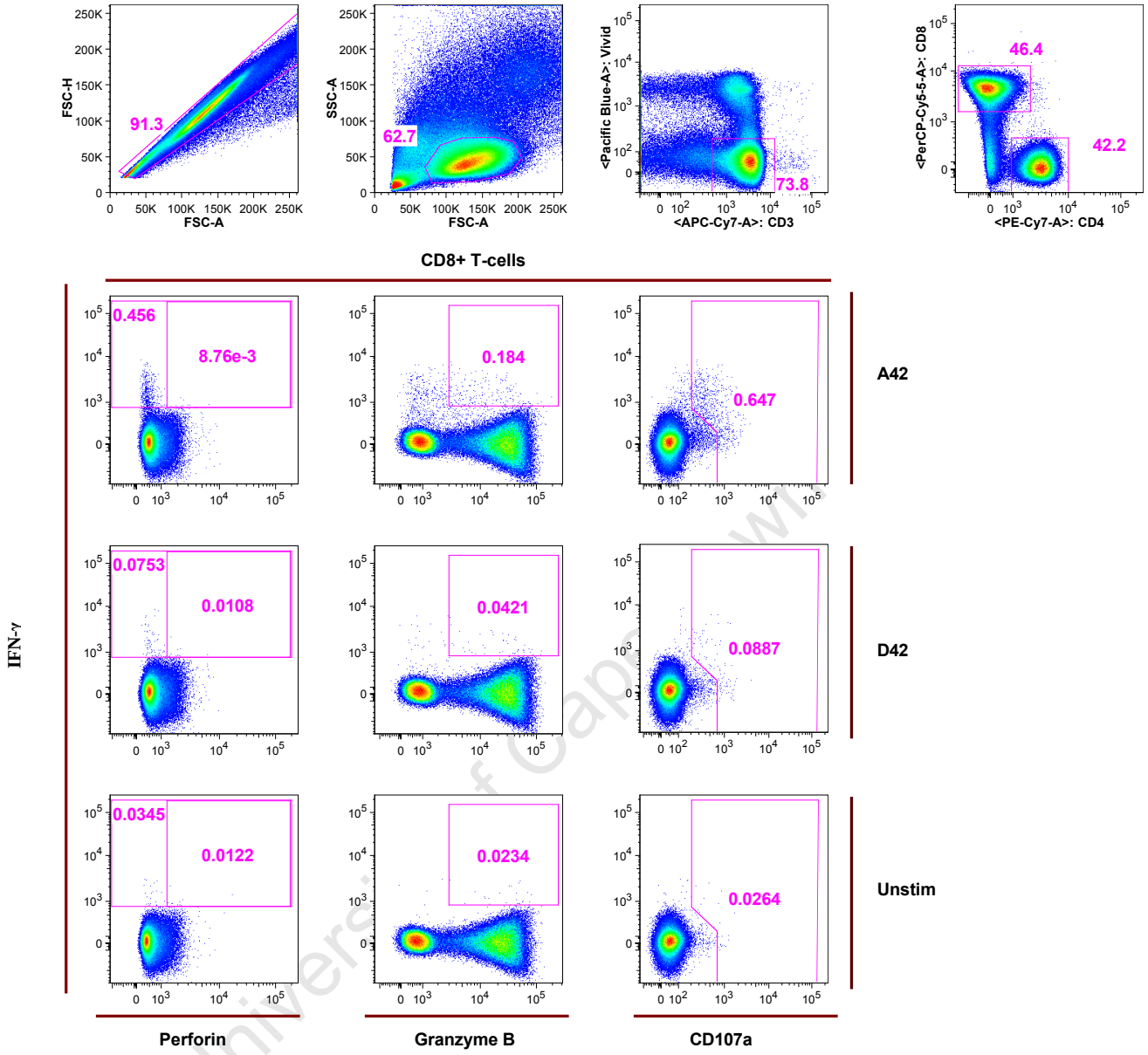


Figure E2.5. Cytpotoxic potential of CD8+ T-cells recognizing peptide variants A42 and D42 in study individual ZA029.

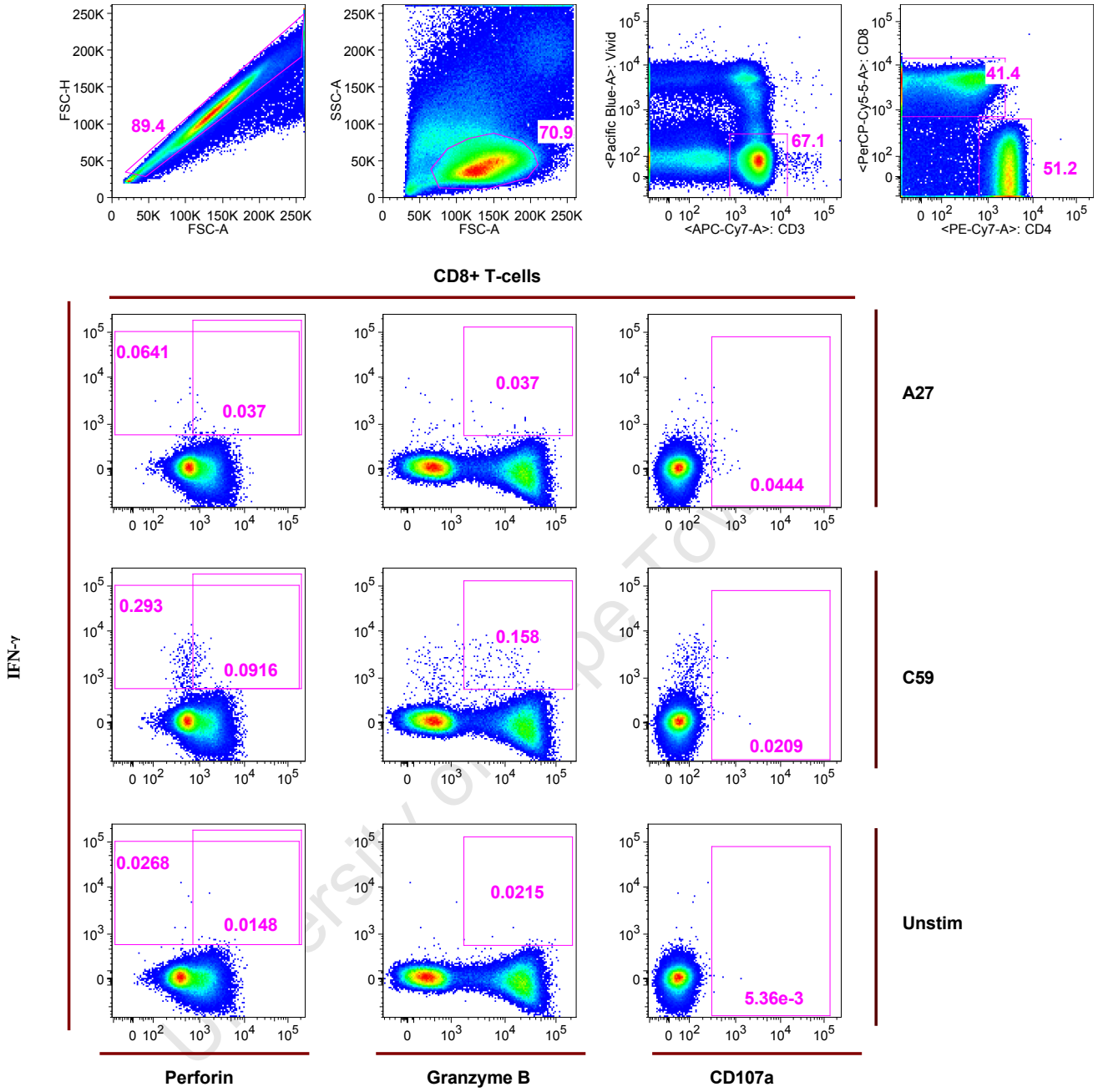


Figure E2.6. Cytpotoxic potential of CD8+ T-cells recognizing peptide variants A27 and C59 in study individual ZA030.

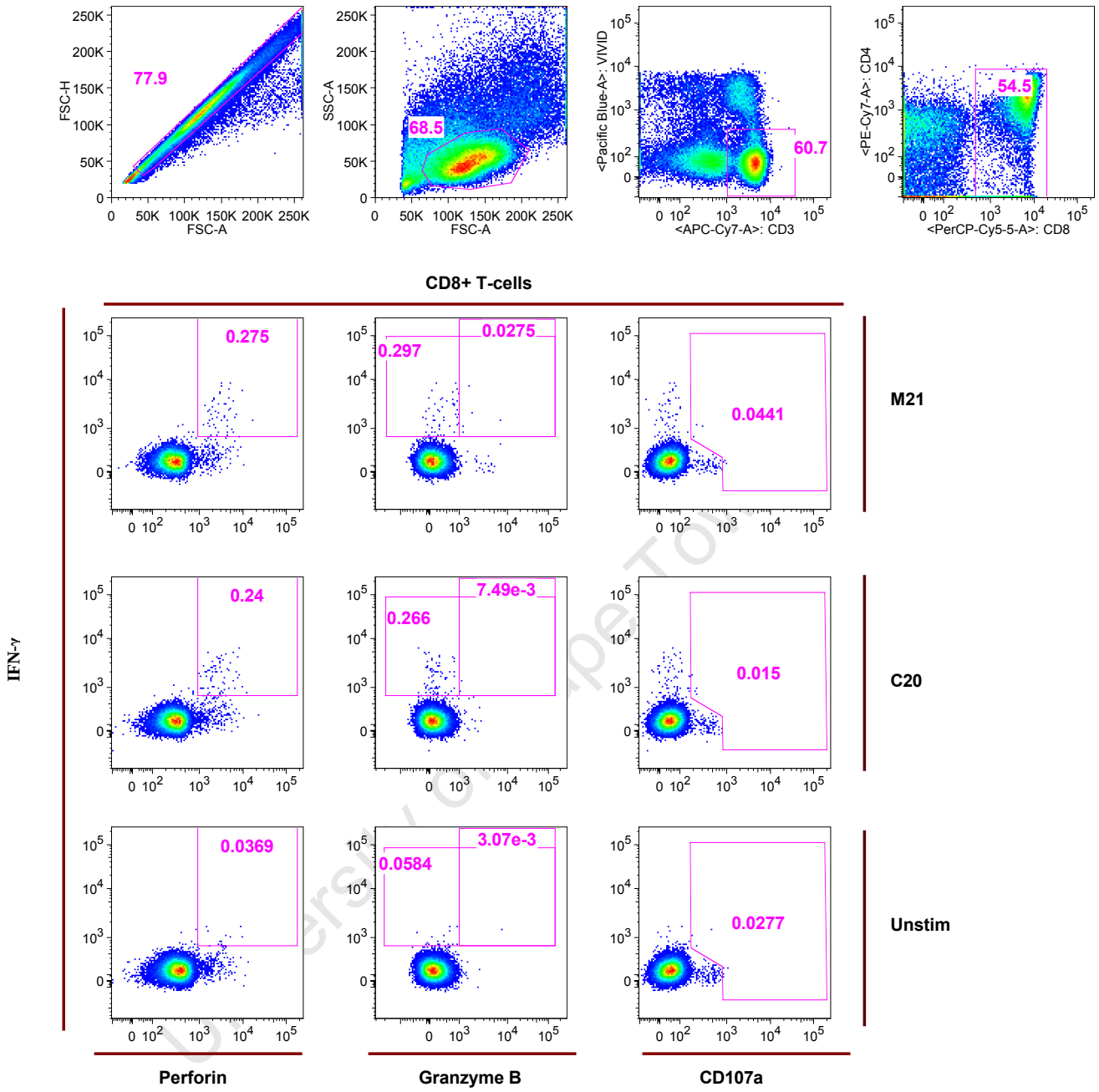


Figure E2.7. Cytpotoxic potential of CD8+ T-cells recognizing peptide variants M21 and C20 in study individual ZA033.

Appendix E3. Oregon green proliferation assay

Protocol

1. Required cells were calculated, approximately 200000 cells/well (four wells/antigen) and one extra unstimulated well for Oregon green compensation control.
2. Fresh RPMI with Penicillin/Streptomycin (5mL) was used.
3. RPMI with penicillin/streptomycin (200mL) was aliqouted and an aliquot of L-Glutamine and Fungin (2mL L-glutamine, 200µl Fungin) was added. This was used to make R10 (RPMI with 10% FCS or RPMI with 10% Human AB serum).
4. Cells were washed twice in PBS (no serum).
5. Oregon Green was diluted 1:10 (add 18µl to a 2µl aliquot).
6. Oregon Green (10µl) was used for 1-10 x 10^6 cells (in 1mL). If more cells, it was increased proportionally.
7. PBS (990µl) was added to 10µl of Oregon Green from step five above
8. This was then added to 1ml of cells (cells resuspended in R10, RPMI with 10% FCS) and incubated at RT for 4 mins in the dark.
9. The cells were mixed and incubated for another 3 mins afterwhich they were vortexed on high speed for 10sec.
10. PBS (1mL) was added and incubated for 3 mins, and then topped up to 15mL with PBS.
11. Cells were centgrifuged at 1200 rpm for 10 mins and resuspend in R10 (RPMI with 10% Human AB serum).
12. Cells were resuspended such that approximately 200000 cells/well (4 wells/antigen) and one extra unstimulated well for Oregon green compensation tube could be plated into a 96-well U-bottomed sterile suspension culture plate.
13. Cells (195µl) and Ag (5µl, final concentration 1ug/ml for each peptide variant). SEB was included as positive control and R10 (RPMI with 10% Human AB serum) as negative control). No BFA was added
14. The cells were incubated for 6 days at 37ºC afterwhich wells for each stimulation were combined in facs tubes and wells washed to make sure no cells were left.
15. This was centrifuged at 1800rpm for 5 min and resuspended in 500 µl R10.
16. Cells were stained as in Appendix E1 for the markers in Table E3.1.

Panel

Table E3.1. Proliferation assay panel and antibodies

Marker	Type/function	Antibody	Laser
Vivid	Identification of live/dead cells	CD14/CD19 Pac Blue	violet
CD4	T-cell marker	PE	blue
CD8	T-cell marker	PerCP Cy5.5	blue
CD3	T-cell marker	APC Cy7	red
Oregon Green	Dye taken by diving cells	FITC	blue