The Impact of Host and Therapy Mediated Selection on HIV-1 Evolution

A thesis submitted in partial requirement for the degree of

Doctor of Philosophy in Clinical Medicine

by

Kuan-Hsiang Gary Huang

Green Templeton College

University of Oxford

Trinity Term 2010

ABSTRACT

The impact of host and therapy mediated selection on HIV evolution

Kuan-Hsiang Gary Huang **D. Phil.**
Green Templeton College **Trinity Term, 2010**

The Human immunodeficiency virus (HIV) pandemic has resulted in a heavy global disease burden, and clinically causes Acquired Immuno-Deficiency Syndrome (AIDS). The development of highly active antiretroviral therapy (HAART) has achieved remarkable control of the rapidly evolving HIV. However, HIV remains neither curable nor preventable by vaccine, and in the developing regions worst affected by HIV, HAART remains inaccessible to most patients. Furthermore, the change in both immunology and viral evolution during chronic HIV infection and its relation to AIDS pathogenesis remains unknown. Following the failure of recent HIV vaccines, it is believed that a better understanding of host-pathogen interaction is vital to advance therapeutic (vaccine and drug) design. In this thesis, I have performed an investigation of viral adaptation in response to different selection forces during advanced HIV infection and AIDS.

The thesis first examined a case study that reveals the potential role of B cell-mediated neutralising antibody (NAb) in chronic HIV infection through the unexpected effect of B cell depletion agent, anti-CD20 (Rituximab). Here, longitudinal results have shown that viral load (VL), *env* gene diversity, and NAb sensitive strains increased during B cell and NAb depletion as a result of Rituximab administration, and reversed as B cells recovered. The study provides preliminary evidence to support the idea that NAb may be effective at suppressing HIV.

The rest of the thesis focused on the cross-sectional cohort at Bloemfontein, South Africa (n=1491), a resource-limited region affected by the pandemic. Here, we used methods that include molecular and pretherapy drug resistance epidemiology, mathematical modelling,

phylogenetically adjusted bioinformatics analysis and *in vitro* viral replication capacity (VRC) assay to study materials including cohort demography, plasma samples, CD4 cell count, VL, viral genetic sequences and host human leukocyte antigen (HLA) tissue types. Our analysis was further augmented by the additional data kindly contributed by our neighbouring Durban cohort collaborators (n=775), which also includes an IFNγ ELISPOT assay that measures cytotoxic T lymphocyte (CTL) responses.

Using the HIV *pol* sequencing data and phylogenetic analysis we confirmed that the local molecular epidemiology is similar to the circulating strains documented in the regional database. However, the pretherapy drug resistance mutation screening results have revealed an unexpected high incidence of drug-induced viral mutants in the AIDS patients with CD4 counts <100 cells/μl. According to mathematical modelling, this finding is attributable to additional sources of antiretroviral therapy exposure, which warrants public health caution. The investigation then focused on studying the changes in HLA class I mediated CTL selection and viral evolution as CD4 counts are reduced in AIDS. Interestingly we have noted evidence that suggest weakening CTL immune selection against *gag* during AIDS is associated with increased viral fitness (measured by VRC) and reversion of previous immune-escape mutations which conferred high fitness costs.

In conclusion, this thesis compared different sources of host and drug mediated HIV selection and its implication for viral evolution. The identification of more bottleneck sites conferring high fitness costs to the selection of escape mutants is expected to be helpful in the design of future therapeutics (via vaccine, drug, immune therapy, or public health strategy). As we have learnt from the principle of combinational ARV, it would be desirable for a vaccine to select HIV at multiple sites of high escape-mutation fitness cost, hence offering protective effect.

ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisors, Professor Paul Klenerman, Dr John Frater, and Professor Rodney Phillips, for allowing me to work in their laboratories, offering of unreserved support, provision of generous funding and for their excellent advice and critical appraisal at every stage of my research. This thesis would not have been possible without their mentoring and support.

I would also like to acknowledge all those who have directly contributed to this thesis, most of whom are detailed in the "Contribution to Thesis Section". Specifically, I would like to thank all of the patients, clinicians, researchers and funders once more. In specific, my D.Phil. program was only made possible following the generous support received from The Rhodes Trust, The James Marin 21st Century School, Medical Research Council, Green Templeton College and Wellcome Trust. Equally, I would like to thank the colleagues and friends that have offered me mixture of mentorship, enlightenment, partnership and friendship essential to the perfection of this thesis: Professor Philip Goulder, Professor Angela McLean, Dr Ellie Barne, Dr Dominique Goedhals, Dr Aris Katzourakis, Dr Mark Brockman, Dr Zabrina Brumme, Dr David Heckerman, Dr Jonathan Carlson, Professor Todd Allen, Professor Bruce Walker, Professor Shie-Liang Edmond Hsieh, Professor Siamon Gorden, Professor Jon Austyn, Professor Helen Chapel, Professor Nick Willcox, Professor Lars Fugger, Dr Oliver Pybus, Dr Tulio De Oliveira, Dr Chris Seebregts, Dr Cloete van Vuuren, Stephen Hickling, Helen Brown, Nicola Robinson, Swati Mishra, Dr Matthias Hoffman, Dr Helen Fryer, David Bonsall, Dr Emma Thomson, Dr Philippa Matthews, Dr Anna Duda, Dr Tsan Yuk Tommy Lam, Dr Vicki Fleming, Nicole Yager,

Christopher Tang, Dr Suzanne English, Dr Stephen Leslie, Chia-Ming Leo Chang, Jessica Mitchell, Isla Humphrey, Hayley Crawford, Dr Tien-Yu Owen Yang, Dr Ching-Chi-Chi, Dr Chih-Hao Lucas Chang, Dr Pao-Yang Chen, Dr Dong Tao, Dr Lucy Dorrell, Dr Nilu Goonetilleke, Dr Yu-Hoi Kang, and Dr Rebecca Payne.

Beyond those involved in the work itself, I would like to thank all my friends and family who have always been loving, caring, supportive and encouraging. I am indebted to so many life mentors and friends who have fulfilled my inspiring pursuit for D.Phil. and Oxford experience: Marian, Ryan, Johnny and family, Yung-Tai, Leo, Corrine, Harry, Prof Oettle, Su-Ching and family, Chris, Tafadzwa, Bruce, Ioannis, Adrian, Owen, Bryan, Sophie, Anthony, Wen, Ye, Mo, Dong-Dong, Ed, Charlene and family, Marcy, Ching-Chi and family, Alan, Bean, David, Yu-Fen, Pao-Yang, Crystal, Mya, Brian, Arthur, Henry, Jason, Hao Tong and family, David, Hsiao-Fen, Jason, Yao-Hung, Evelyn, Ariel, George, Jasmine, Gorden, Iris, Ko-Yun, Vivian, Laurel, Howie, Ken, Jonathan, Oscar, James, Julie, Anita, Seung-Hye, Derek, Nelson, Prof Veriava, Mavis and family, Jean and family, Aunties, Charlene, Carol, Penny, Zaheer and family, Ishaaq and family, Prof Laburn, Prof Mitchell, Prof Wadee, Dr Radev, Annette, Justice Edwin Cameron, Mary, Warden and everyone at Rhodes House.

Above all, I would like to thank my wife, Sherry, who is a constant source of strength to me. Her patience, love and kindness have been crucial in keeping me going, and I owe her more than I can express. Jason, William and Leona, thank you for always been such lovely brothers and sister. Finally, thank you, dads and moms, for all of your unreserved love, care, support and inspiration: Sherry and I are grateful to be your children, and wish we can continue to cherish each other in the future.

CONTRIBUTIONS TO THIS THESIS

All of these projects have depended on the supervision, advice and input of my supervisors, Professor Paul Klenerman, Dr John Frater and Professor Rodney Phillips.

Chapter 3:

This retrospective clinical case study with longitudinal sampling was made possible because the case subject came from a larger short course antiretroviral therapy (SCART) trial at St Mary's Hospital, London, United Kingdom. The clinicians and researchers at St. Mary's Hospital, led by Dr Graham Crooke, referred the case to us for further investigation. Phillips' group laboratory is an active collaborating group of the St Mary's Hospital SCART trial, where Nicola Robinson and Helen Brown performed most of the host HLA class I genotyping and longitudinal ELISPOT assays to measure CTL responses for the entire cohort. I have performed the longitudinal, clonal sequencing of HIV *env* especially surrounding the timing of Rituximab therapy. I have also conducted the phylogenetics and evolutionary biology analysis under expert advice received from Dr Aris Katzourakis and Dr Oliver Pybus (Department of Zoology, University of Oxford). The collaborators from Myra McClure's group (Imperial College, London), particularly David Bonsall and Emma Thomson, have carried out antibody neutralisation assays.

The project detailed in Chapter 3 has also been re-written into manuscript format, and published in Nature Communication (2010). Current list of co-authors to the manuscript includes: Kuan-Hsiang G. Huang and David Bonsall (equal contribution as first author), Aris Katzourakis, Emma C. Thomson, Sarah J. Fidler, Janice Main,

David Muir, Jonathan N. Weber, Alexander J. Frater, Rodney E. Phillips, Oliver Pybus, Philip Goulder, Myra O. McClure, Graham S. Cooke (equal contribution as the last author) and Paul Klenerman. We wish to also thank Quentin Sattentau for his advice and support. Antibodies b12 and 4.8d were kind gifts from Dennis Burton (Scripps Research Institute) and James Robinson (Tulane University), respectively. We are grateful for the funding support received from Wellcome Trust, NIHR Biomedical Research Programme (Oxford and Imperial College NHS Trust), Rhodes Scholarships (for Kuan-Hsiang Gary Huang) and the James Martin School for 21st Century.

Chapter 4, 5 and 6:

The Bloemfontein (BFN) cohort in Free State Province, South Africa, forms the core of this thesis investigation. We are grateful to the patients, clinicians and researchers at the governmental antiretroviral (ARV) clinics, University of Free State, and Medical Research Council of South Africa, as well as our researcher colleagues at Department of Zoology and Clinical Medicine in University of Oxford. In Chapter 5, our collaborator, Professor Philip Goulder and his research consortium has kindly contributed some of the relevant research information from their Durban (DBN) cohort, which neighbours the BFN cohort geographically (see Chapter 2 and 5 for more details). The DBN cohort has provided augmented power in our bioinformatics analysis, as well as important knowledge regarding HLA class I mediated CTL selection in advanced AIDS.

In Chapter 4, Nicola Robinson, Helen Brown, Dr Dominique Goedhals and Stephen Hickling have conducted part of the viral mRNA extraction and cDNA templates

transcription processing. We are also grateful to Dr Aris Katzourakis, Dr Oliver Pybus, Dr Tulio De Oliveira, Los Alamos National Library database, Stanford Drug Resistance database and BioAfrica database for the advices and supporting of data analysis regarding phylogenetics, molecular epidemiology and pretherapy drug resistance mutation analysis. Professor Angela McLean and Dr Helen Fryer from the Department of Zoology have also contributed their expertise in mathematical modeling that infers disease spread and progression in Chapter 4. Dr Owen Yang, Ching-Chi Chi and Dr Pao-Yang Chen have provided valuable advice regarding statistic analysis of the cohort. The project reported in Chapter 4 was published in the peer review journal, Antiretroviral Therapy during 2009, with the following co-authors: Kuan-Hsiang Gary Huang, Dominique Goedhals, Helen Fryer, Cloete van Vuuren, Aris Katzourakis, Tulio De Oliveira, Helen Brown, Sharon Cassol, Chris Seebregts, Angela McLean, Paul Klenerman, Rodney Phillips, John Frater, and the Bloemfontein-Oxford Collaborative Group (Huang, et al. 2009).

In Chapter 5, Dr David Heckerman and Dr Jonathan Carlson from Microsoft Institute at Seattle, United States, have conducted a Phylogenetic Dependency Network and jointly developed the phylogenetically corrected odds ratio analysis with us that detects changes in HLA class mediated associations in different CD4 count strata. Most of HIV *gag* and *nef* sequencing was conducted by Dr Dominique Goedhals and Stephen Hickling, respectively. The combination of BFN cohort (established by Phillips' group) with DBN cohort (contributed by researchers at Goulder's group, University of Oxford) has improved the statistical power significantly. The DBN cohort has further provided ELISPOT information that help with the understanding of changing CTL function in advanced AIDS. We are also grateful for Chia-Ming Leo

Chang's contribution regarding the correction in the counting of overlapping peptide in the analysis of ELISPOT results.

In Chapter 6, Dr Mark Brockman, Dr Zabrina Brumme, Dr Toshiyuki Miura and Professor Bruce Walker have provided us with technical training and support regarding construction of chimeric virus for the measurement of viral replication capacity (VRC). Swati Mishra and Christopher Tang have helped considerably in the construction of chimeric viruses, and the measurement of VRC. Nicola Robinson and Nicole Yager have provided valuable support to ensure our methods are carried out safely and efficiently in the Category III laboratory environment. Nicole Yager has also kindly shared her experience and so contributed to the optimization in the methodology of Chapter 6.

The results reported in Chapter 5 and 6 are currently in the process of re-writing into manuscript for peer-reviewed journal publication (currently reviewed by the journal PLoS Medicine). The current proposed list of co-authors includes: Kuan-Hsiang Gary Huang and Dominique Goedhals (equal contribution as first author), Jonathan Carlson, Cloete van Vuuren, Mark Brockman, Swati Mishra, Zabrina Brumme, Stephen Hickling, Christopher Tang, Chris Seebregts, Dewald Steyn, David Heckerman, Paul Klenerman, Philip Goulder, Rodney Phillips, Bloemfontein-Oxford Collaborative Group and John Frater. The projects in Chapter 4, 5 and 6 are jointly funded by generous supports provided by James Martin 21st Century School, Rhodes Trust, Medical Research Council (UK) and Wellcome Trust.

For

Mums, Dads and Sherry

TABLE OF CONTENTS

TABLES AND FIGURES

TABLES

FIGURES

ABBREVIATIONS

℃	Degrees centigrade
∞	Infinity
AA	Amino acid
AIDS	Acquired immunodeficiency syndrome
Amp	Ampicillin
ANC	Antenatal clinic
ARV	Anti-retroviral therapy
BFN	Bloemfontein
CD	Cluster of differentiation
cDNA	Complimentary DNA
CO_2	Carbon dioxide
CRF	Circulating recombinant forms
CTL	Cytotoxic T cell
DBN	Durban
DNA	Deoxyribonucleic acid
DRC	Democratic Republic of the Congo
E. Coli	Escherichia coli
ELISPOT	Enzyme linked immunospot
env	Envelope Glycoproteins (gp120 and gp41)
FACS	Fluorescence activated cell sorter
FCS	Foetal calf serum

FS	Free State province, South Africa
g	Gram
gag	Capsid Proteins
GALT	Gastrointestinal associated lymphoid tissues
HAART	Highly active antiretroviral therapy
HIV	Human immunodeficiency virus
HLA	Human leukocyte antigen
INT	Integrase
KZN	KwaZulu Natal
LB Broth	Luria-Bertani Medium
Log	Logarithm
LTR	Long terminal repeats, containing regulatory elements
MAb	Monoclonal antibody
MCMC	Markov chain Monte Carlo
MHC	Major histocompatibility complex
μg	Microgram
Mg	Milligram
μL	Microlitre
mL	Millilitre
μM	Micromolar
ML	Maximum likelihood
Ms	Millisecond
MTCT	Mother to child transmission
Mut	Mutants
OR	Odds ratios

N	Number
NAb	Neutralising antibodies
NaCl	Sodium chloride
nef	Negative Factor Protein
ng	Nanogram
nM	Nanomolar
NRTI	Nucleoside reverse transcriptase inhibitor
NNRTI	Non-nucleoside reverse transcriptase inhibitor
OLP	Overlapping peptides
PBMC	Peripheral blood mononuclear cells
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
pH	Potential hydrogen
PhyloOR	Phylogenetically adjusted OR
PI	Protease Inhibitor
pol	Viral enzymes (protease, reverse transcriptase and integrase)
PR	Protease
pVL	Plasma viral load
rAd5	Recombinant adenovirus serotype 5
RE	Restriction enzyme
RNA	Ribonucleic acid
rev	The second regulatory factor
rpm	Revolutions per minute
RT	Reverse Transcriptase
RT-PCR	Reverse transcriptase polymerase chain reaction

SA	South Africa
SCART	Short course antiretroviral therapy
SFU	Spot forming units
SIV	Simian immunodeficiency virus
tat	Transactivator of HIV gene expression
U	Units
V	Volts
vif	Viral Infectivity Factor
VL	Viral load
vpr	Viral Protein R
vpu	Viral Protein U
VRC	Viral replication capacity
WT	Wild type

CHAPTER 1:

Introduction

1.1 Global pandemic

Since its identification as the pathogen responsible for Acquired Immuno-Deficiency Syndrome (AIDS) over 25 years ago, human immunodeciency virus (HIV) infection has emerged as a major global pandemic (Weiss 2008). In 2008, 33.4 million of the global population were living with HIV infection, 2.7 million of new cases were diagnosed and 2.0 million people died from HIV related causes. Although the disease is widespread globally, sub-Saharan Africa is the epicentre of the pandemic, with 22.4 million infected individuals. In 2007, the prevalence of HIV infection amongst pregnant women in the worst affected region such as Free State (FS) and KwaZulu Natal (KZN) provinces of South Africa were estimated to be 33.5% and 37.4%, respectively (AVERT.org 2009). With 10.8% of South African population infected with HIV, the life expectancy at birth has reduced from its peak of 61.50 years in 1991 to 51.48 years in 2008 (Bank 2009, UNAIDS/WHO 2009).

HIV infection is arguably the most serious infection to have affected humankind (Rambaut, et al. 2004). The development and introduction of rapid diagnostic tests, antiretroviral therapy (ARV, and its optimised combinational regimen of highly active antiretroviral therapy, HAART) and effective public health campaign over the past three decades have remarkably improved the quality of life and reduced disease progression. However, current HAART regimens are unable to clear established HIV

infection and HAART remains not universally accessible (or affordable) to every patient in need. The rapid emergence and transmission of HIV mutants resistant to HAART and ARV-related drug toxicity further complicates the clinical application of ARVs. Other major challenges to the HIV pandemic include: a struggle to develop an effective vaccine (discussed in Chapter 1.7), social stigmatisation and denial (Levy 2007, Weiss 2008).

1.2 Natural disease progression

HIV is transmitted via exchange of body fluids containing infectious viral particles, typically through the one of the three routes: sexual contact, parenternal transmission, or mother to child transmission (MTCT). As Figure 1.1 demonstrates, the viral infection often manifests clinically in three stages: a transient acute seroconversion illness, a variable period of latency and eventually AIDS (Levy 2007). HIV specifically infects CD4 subsets of T cells and macrophages, replicating constantly to maintain or increase its viral load (VL), causing continual CD4 T cell decline and the ultimate failure of immune competence.

Following transmission, HIV replicates rapidly amongst CD4 T cell subsets across the body, especially within the gastrointestinal associated lymphoid tissues (GALT), resulting in peaked plasma viraemia, an approximate 50% decrease in the total CD4 T cell count and a moderate decrease in the circulating CD4 T cell count in the blood (Brenchley, et al. 2004, Fiebig, et al. 2003, Kaufmann, et al. 1998). The acute infection typically last between 3-12 weeks and the patients often experience mild flu-

like symptoms, fever and a characteristic rash (referred to as seroconversion illness); the diagnosis is often missed during this period.

The pathogenesis occurring during acute HIV infection reflects complex host-pathogen interactions, and sets an important scene for the subsequent chronic disease progression: it is apparent that in approximately 76.47% of infections (n=102) through the mucosal route are established from a single strain of viral founder, with the remainder of infections resulting from two, three or four viral founder strains (Keele, et al. 2008). The gp120 subunit of the HIV envelope glycoprotein binds specifically to surface $\alpha 4\beta 7$ integrins of gut-homing CD4 T cells, thereby facilitating specific entry into the GALT – the largest tissue lymphocyte reservoir of the body (Arthos, et al. 2008).

During acute infection, over 80% of the GALT are destroyed (Brenchley, et al. 2004, Li, et al. 2005). This massive cell death is the combined result of direct viral cytopathic killing and indirect killing of both infected and uninfected CD4 T cells due to the activation induced cell death (Hellerstein, et al. 2003, Xu, et al. 1997). Following the peak viraemia of acute infection, the VL typically decreases significantly to a lower level that varies between individuals, possibly due to the substrate exhaustion (loss of GALT) and concurrent emergence of CD8 T cell responses (discussed further in Chapter 1.5) (Borrow, et al. 1994, Koup, et al. 1994).

The patients are mostly asymptomatic during the clinical latency period, despite continual viral evolution and CD4 T cell decline. Although the plasma viral load (pVL) during latency appears relatively constant, the chronically translocated enteric

bacteria and their pro-inflammatory products (such as lipopolysaccharides) further facilitate activation of CD4 T cells that fuels the viral infection (Brenchley and Douek 2008, Douek 2007, Grossman, et al. 2006). Indeed, the level of translocated bacterial products, immune activation and "set point" level are predictive of the rate of CD4 T cell decrease and disease progress to AIDS (Geskus, et al. 2007, Mellors, et al. 1996). The tropism of the virus has often been noted to shift over the course of chronic infection: HIV evolves from nearly exclusively CCR5+ tropic strains in its co-receptor usage during early infection, towards to either dual-tropic (CCR5 and CXCR4) or CXCR4-tropic strains in 50% of the HIV-1 subtype B infections. This has been hypothesised to be the results of substrate exhaustion over chronic infection (in CCR5+CD4+ cells) (Dittmar, et al. 1997).

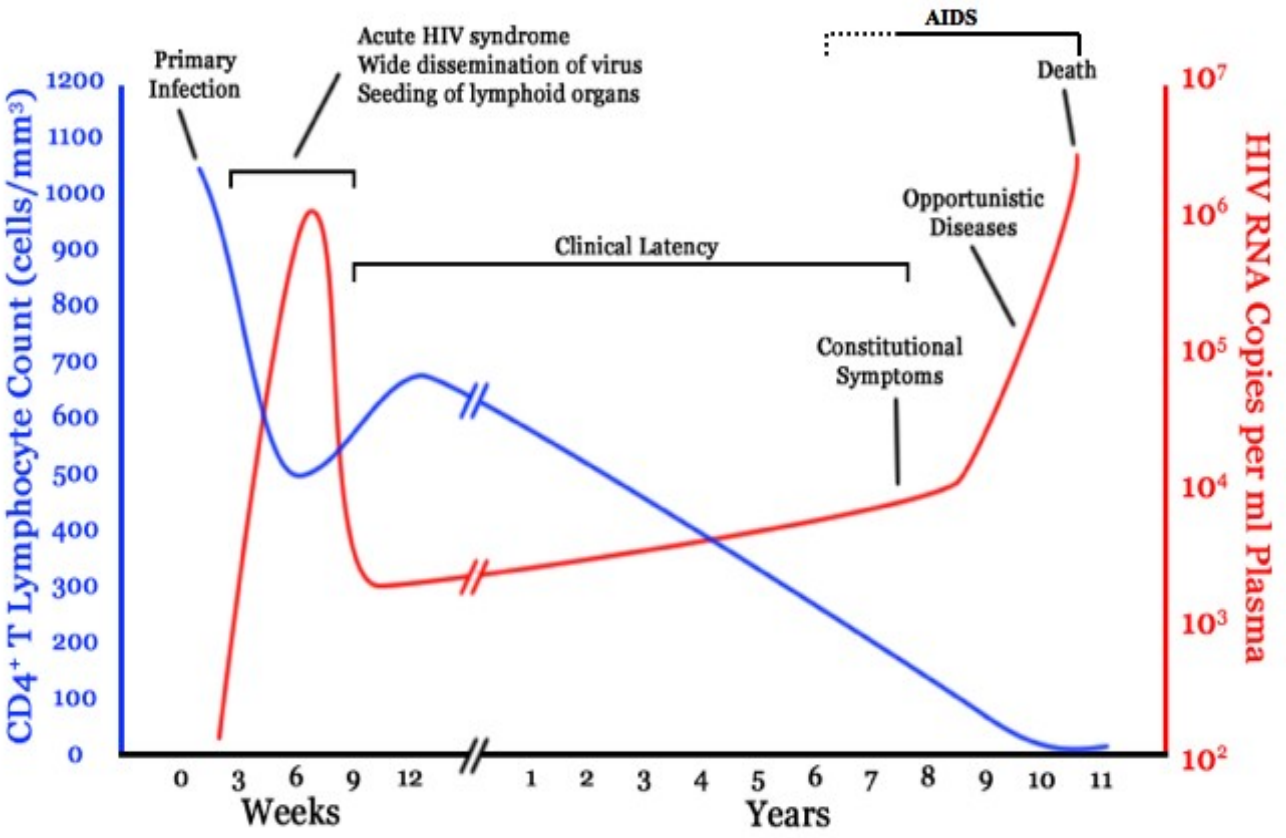


Figure 1.1: Cartoon showing a typical course of HIV infection

(Image modified from Pantaleo G. et al, 2003 (Pantaleo, et al. 1993))

Unless treated with antiretroviral therapy (ARV), the immune system continues to deteriorate during HIV infection. AIDS is the syndromic end stage presentation of HIV infection. It occurs when the immunodeficient hosts become vulnerable to opportunistic infections and tumours, causing debilitation and death (Stevenson 2003). Clinical symptoms of immunodeficiency and declining CD4 T cell counts of patient are two core clinical markers that define AIDS (HIV/AIDS, et al. 1993). The likelihood of AIDS increases significantly when the plasma CD4 T cell counts become less than 350 cells/mm^3, and is clinically defined when CD4 T cell counts become less than 200 cells/mm^3.

Several factors have an important impact on the rate of natural disease progression and clinical outcome. This can be broadly classified into factors relating to the pathogen, host, environment, and the interplay between them (Levy 2007, Stevenson 2003, Weiss 2008). In terms of pathogen, the quantity of viral inoculum, adaptive fitness of the virus(es) that establish the infection, the route of transmission and VL during latency period (the "VL set point", discussed in Chapter 1.5) are some recognised factors that can affect disease progression (Brumme, et al. 2009, Frater, et al. 2006, Hecht, et al. 2010, Mellors, et al. 1997). Environmental factors include, but are not limited to, the sexual or drug use behaviours, the existence of other sexual transmitted infections, nutritional status, translocation of enteric bacteria and their pro-inflammatory products, use of other medication with immunosuppressant effects, and concomitant infections that enhance HIV infection or reduce immune control (Cohen 2004, Melchior, et al. 1999, Nagot, et al. 2007). In terms of host factors, the rate of CD4 T cell loss, the changes in CD4 cell immune quality and constitution, the level of immune activation in host, the efficiency of the antiviral immune response,

the host tissue type (especially the human leukocyte antigen (HLA) class I profile) and other human genetic polymorphisms are some key factors that relate to disease progress, or control (Brumme, et al. 2009, Fellay, et al. 2007, Frater, et al. 2007, Kaslow, et al. 1996). The study of mechanisms leading to effective control over HIV infection and clinical progression can potentially inform successful therapeutic design. This thesis will aim to achieve better understanding of the complicated host-pathogen interaction that mediates viral evolution and impact on disease progression.

1.3 Structure, Function and Replication Cycle

HIV is a lentivirus from the *Retroviridae* family. As shown in Figure 1.2, HIV contains two copies of single stranded positive sense RNA genome complexed with viral enzymes. Accessory and regulatory protein are stabilised inside the bullet shaped protein core (formed from capsid protein, p24). The core is surrounded by a spherical protein matrix layer derived from capsid protein, p17, which is in turn enveloped inside a lipid bilayer derived from host cell membrane (Levy 2007, Stevenson 2003). The viral surface envelope glycoprotein gp120 protrudes on the virion surface to identify the target cell receptor CD4 and co-receptors, whereas the transmembrane gp41 facilitates fusion between the viral and target cell membranes (Kowalski, et al. 1987, Kwong, et al. 1998).

The HIV-1 genome is approximately 9.7 kilobases long (inclusive of long terminal repeat, LTR, sequences), containing 9 genes that encode for 15 proteins (Figure 1.3). The three principal genes are *gag* (encodes for core structural capsid protein p7, p17 and p24, and p6 protein), *pol* (encodes for protease, reverse transcriptase and

integrase enzymes essential for viral replication and processing) and *env* (encodes for envelop glycoprotein gp120 and gp41). In addition, HIV-1 possesses 6 accessory genes, *vif*, *vpr*, *vpu*, *tat*, *rev*, and *nef* that regulate the replication cycle, interact with the host and modulate host responses (Greene and Peterlin 2002).

The life cycle of HIV is summarised in Figure 1.3. Briefly, after HIV binds to CD4 cell surface molecules, entry into the cell requires binding to co-receptors (CXCR4 and CCR5) (Choe, et al. 1996, Dragic, et al. 1996, Feng, et al. 1996). HIV is uncoated inside the cell and reverse transcriptase (RT) copies genomic RNA into cDNA. Viral DNA utilises the enzyme integrase (INT) to integrate into host DNA and become a part of the cellular genome. The virus uses cellular machinery to synthesize viral proteins (Gomez and Hope 2005, Stevenson 2003). Several of these are long amino acid chains that must be cleaved by a specific viral protease (PR) before new viral particles can become active(Klotman, et al. 1991). The virus is then assembled into mature HIV virions that bud off from the host cell with host cell membrane (Stevenson 2003). The distinct phases of viral replication life cycle provide potential targets for the design of antiretroviral therapies (Figure 1.4).

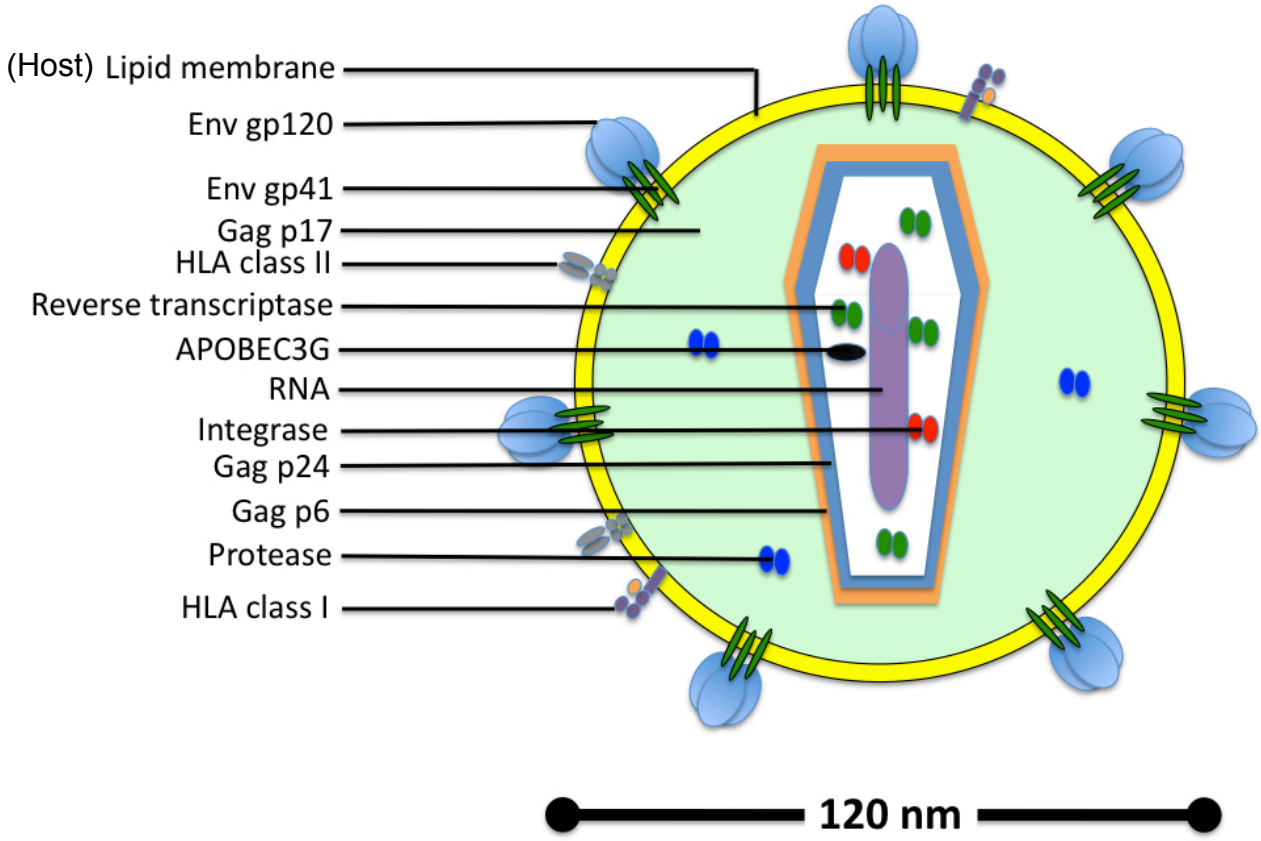


Figure 1.2: A schematic representation of the HIV virion

(Original sketch adapted from the image of Los Alamos HIV Database (Los Alamos National Laboratory 2010)).

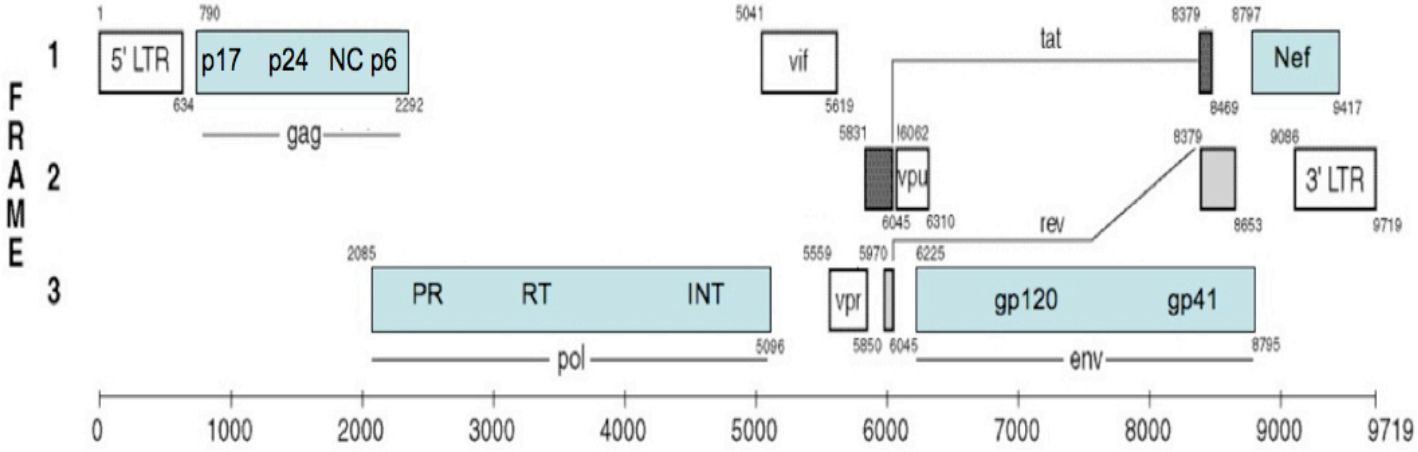


Figure 1.3: A schematic organisation of the HIV-1 genome

(Adapted from: Los Alamos database(Los Alamos National Laboratory 2010)).

Key abbreviations: **LTR –** Long terminal repeat sequences, **gag –** Capsid proteins gene encoding viral matrix (p17), capsid (p24), nucleocapsid (NC) and p6, **pol** - Viral enzymes (Protease (prot), Reverse Transcriptase (RT) and Integrase (INT)), **tat -** Transactivator of HIV gene expression, **rev –** The second regulatory factor, **vif -** Viral Infectivity Factor, **vpr -** Viral Protein R, **vpu -** Viral Protein U, **env -** Envelope glycoproteins gp120 and gp41, **nef -** Negative Factor Protein. This thesis will focus on the genes highlighted in turquoise.

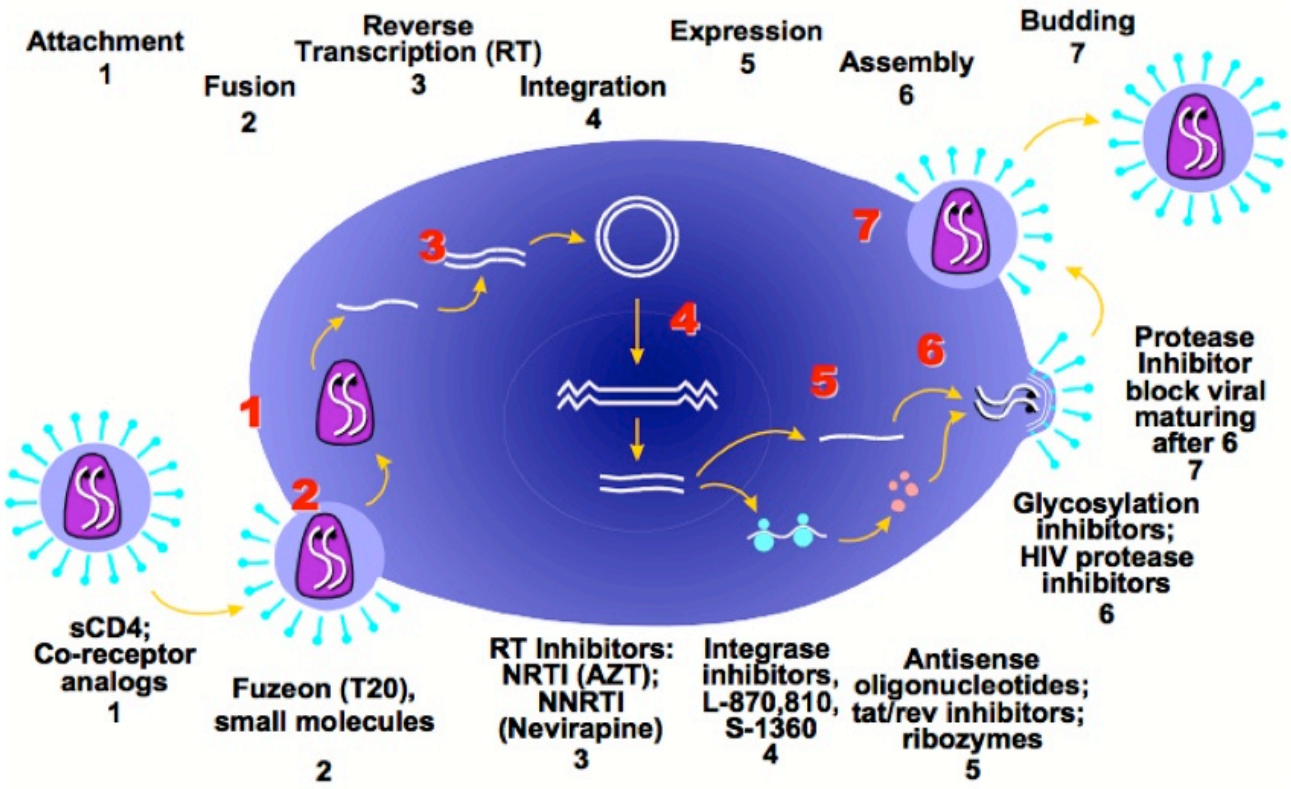


Figure 1.4: Life cycle and potential therapeutic targets of HIV

(Modified from http://www.dukehealth.org/Services/DART/ (S. Georgia Centre for AIDS Research 2005)).

1.4 HIV Origin, Divergence and Diversity

HIV-1 is most likely to have originated from Central Africa as a primate lentivirus, simian immunodeficiency virus (SIV), through several episodes of zoonotic transmission over early 1930s (Korber, et al. 2000, Lemey, et al. 2003). The rapid and variable mutation rate of the virus, complicated human population growth and globalisation movement during the 20th century together with a lack of case definition, diagnostic methods and clinical sample storage have contributed to the difficulties in dating the exact timing of HIV transmission to humans. Before 1976, there exist only two preserved samples that successfully yielded HIV-1 RNA and sequences, both found in Léopoldville of the then Belgian Congo (now Kinshasa, Democratic Republic of the Congo, DRC): one is a subtype D HIV-1 found in plasma sample stored in 1959, and another subtype A viral was very recently extracted from the preserved paraffin fixed lymph node tissue block dated back to 1960 (Worobey, et al. 2008, Zhu, et al. 1998). The diversity existed between these two earliest samples suggests that the original event of HIV cross-species establishment may date back to as early as 1890-1920s (Worobey, et al. 2008).

There are two major sub-species of HIV. HIV-1 is thought to originate from SIV of chimpanzees (*Pan troglodytes troglodytes*) in Central Africa, whereas HIV-2 is thought to originate from SIV of sooty mangabey monkeys (*Cercocebus atys*) of West Africa (Hahn, et al. 2000, Lemey, et al. 2003, Worobey, et al. 2004). HIV-1 can be further divided into three groups, M, N, and O. HIV-1 group M is the pandemic viral strain with extensive diversity and near global distribution. The group M is subdivided phylogenetically into at least 11 subtypes and 45 circulating recombinant

forms (CRF), which predominate in different geographical regions (Los Alamos National Laboratory 2010, Robertson, et al. 2000).

Recent studies suggest that bottlenecks exist during HIV transmission; in the majority of cases (76.47%, n=102) only one founder virus is responsible for the establishment of early infection. HIV is capable of generating very high intrapatient diversity (Keele, et al. 2008, Korber, et al. 2000). The genetic diversity of HIV is extremely high compared to other infections or population models, such as influenza viruses. On average, the error prone viral reverse transcriptase introduces one nucleotide mutation per replication cycle, whilst template switching and recombination offers further mechanisms to generate highly diverse variants (collectively known as "quasispecies"). Although random mutations are usually detrimental, the rapid and high output of diverse HIV quasispecies offer a highly effective adaptation strategy against pressures from host defence, vaccine challenges and therapeutic targeting (Arien, et al. 2007, Korber, et al. 2001, Malim and Emerman 2001, Shankarappa, et al. 1999, Stevenson 2003, Weiss 2008).

This thesis will focus on the most prevalent subtypes: subtype C (HIV-1C, accounts for over 52% of global cases) that predominates in Sub-Saharan Africa/Asia, and subtype B (HIV-1B, accounts for 10% of global cases) that predominates in America/Europe (Figure 1.5). We aim to explore how HIV-1 evolves in response to two major selection forces, namely drug therapy and adaptive immunity.

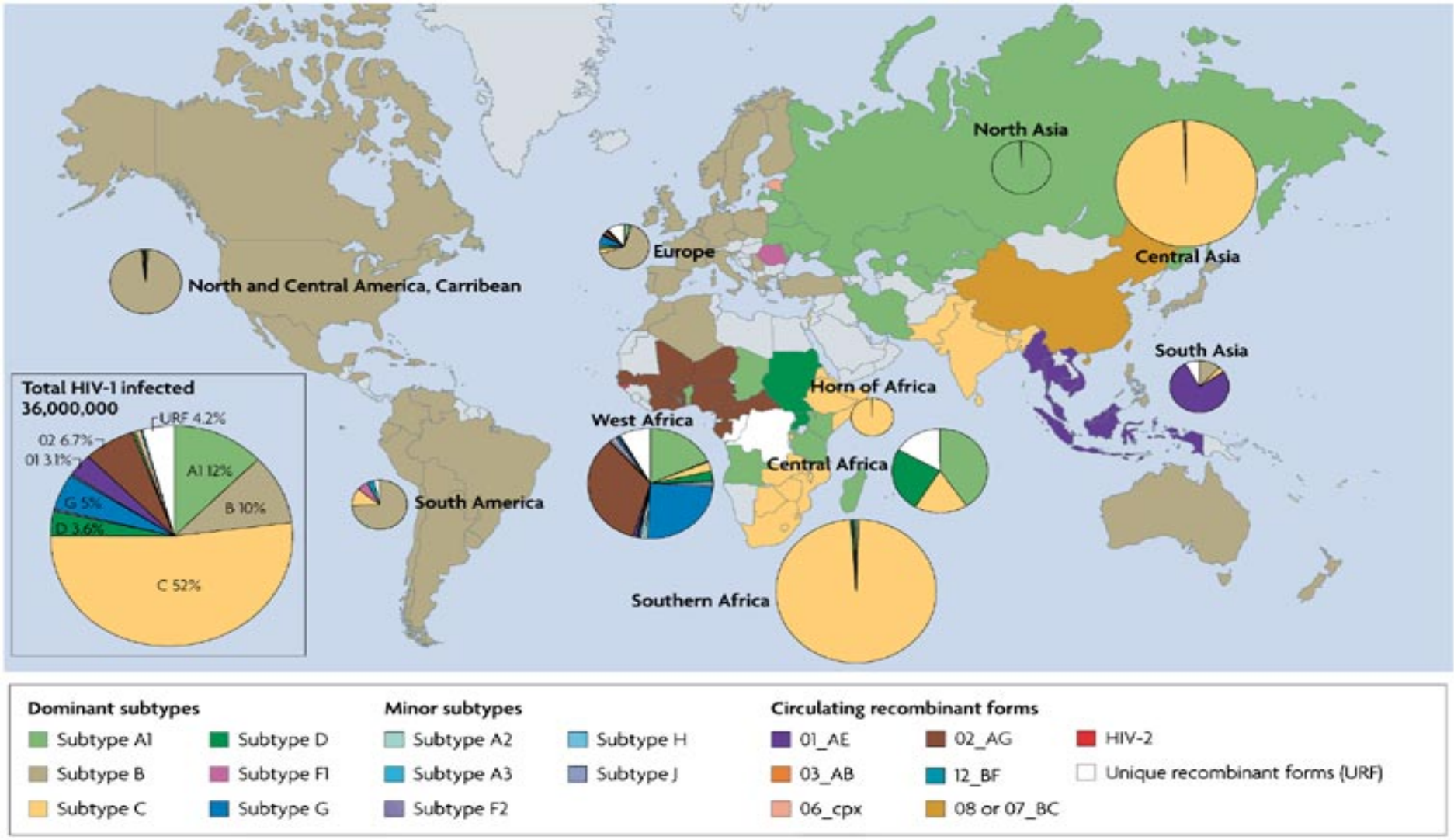


Figure 1.5: Global distribution and impact by different HIV-1 subtypes

(Image source: Arien, K.K. et al. 2007(Arien, et al. 2007))

1.5 Natural Selection: Immune Response

Over the course of natural infection, the host immune response is important in determining HIV disease progression and outcome. It is likely that both humoral and cellular components of the immune response are required for efficient viral control (Klenerman and Hill 2005, Korber, et al. 2001, Letvin 2006, Weiss 2008).

1.5.1 CD8 T cell selection and viral evolution

The cellular arm of the immune response is important in containing an intracellular pathogen like HIV. CD8 T cells, also known as cytotoxic T cells (CTL), are one of the key players with well-defined antiviral roles (Klenerman and Hill 2005, Letvin 2006). CTL recognise HIV infected cells that process the linear viral peptides (termed 'epitopes', often 8 to 11 amino acids in length) intracellularly for presentation on the cell surface using human leukocyte antigen class-I molecules, as depicted in Figure 1.6 (Zinkernagel and Doherty 1997). The three major HLA class-I gene alleles (HLA-A, HLA-B and HLA-C) are highly polymorphic, inherited in a polygenic manner and expressed co-dominantly to generate extensive host population diversity. Different HLA class-I alleles have different preferences for epitope binding, but are essentially capable of sampling intracytoplasmic HIV proteomes (Frater, et al. 2007, Kiepiela, et al. 2007, O'Brien, et al. 2001, Zinkernagel and Doherty 1997). As many as a third of HIV genomes demonstrate a significant footprint of CTL selection (Asquith, et al. 2006, Moore, et al. 2002). CTL recognition of the HLA class-I epitope complex is restrictive, permitting immune targeting of genuinely infected cells in a host-specific manner.

Taken together, the CTL targeting of HIV is proteome-wide using different HLA class-I epitope combinations amongst different individuals (Carrington, et al. 1999, Frater, et al. 2007, Gao, et al. 2001, Kiepiela, et al. 2004, Kiepiela, et al. 2007, O'Brien, et al. 2001, Tang, et al. 2002, Zinkernagel and Doherty 1997). Clinically, the appearance of CTL in acute infection overlaps with the sharp VL decrease to the "set-point" (often a 2 to 5 $\log_{10}$ decrease from VL peak) (Borrow, et al. 1994, Koup, et al. 1994). The SIV/macaque model has confirmed that CTL blockade via monoclonal antibodies is associated with sustained viraemia and rapid progression to AIDS (Jin, et al. 1999, Matano, et al. 1998, Schmitz, et al. 1999). Furthermore, certain HLA class I alleles are over-represented in HIV infected long term non-progressors (LTNP), and associated with lower VL, lower rate of CD4 T cell decline and therefore, better clinical outcome (Kiepiela, et al. 2004, O'Brien, et al. 2001). Examples of these beneficial molecules in the setting of South African population include HLA-B*57, HLA-B*5801 and HLA-B*8101.

The diversity of the immune response, both within the host and at the population level, poses strong selection forces on HIV replication (Frater, et al. 2007). The virus escapes CTL selection by mutating the epitope or its flanking regions, so thereby alters CTL-HLA class-I epitope specificity, epitope processing or presentation by HLA molecules (Goulder, et al. 1997, Korber, et al. 2001, Letvin 2006, Tenzer, et al. 2009). The HLA-driven immune escape hypothesis characterises and predicts patterns of HIV mutation, leaving behind distinctive "HLA footprints" on viral genetic sequences even after transmission into a new HLA-mismatched hosts (Bhattacharya, et al. 2007, Frater, et al. 2007, Leslie, et al. 2005, Moore, et al. 2002). Although population specific HLA diversity is a possible contributing factor to the intra-subtype

genetic differences and even potentially the inter-subtype differences, the extent of this effect is difficult to interpret in the presence of other complicating factors, namely: frequent human population movement, the presence of random mutations, recombination between circulating viral forms, mutations driven by other selection forces (other immune factors or ARV), viral founder effects, complex transmission networks, paucity of historical samples for comparative viral evolution studies, epitope clustering in immunogenic regions, and complex dynamics that exist between variable selection pressures imposed by different HLA molecules on the viral genes with variable plasticity to mutate (thereby mediating different rates that lead to fixation of escape mutations, occurrence of secondary/compensatory mutations, and even reversion of mutations to wild type upon transmission of variant virus into a HLA mismatched hosts). (Asquith, et al. 2006, Bhattacharya, et al. 2007, Duda, et al. 2009, Goulder and Watkins 2008, Kawashima, et al. 2009, Leslie, et al. 2005, Worobey, et al. 2008).

Mutations within certain HIV genes and at amino acid sites may confer higher fitness costs to the virus, leading towards better clinical control (Goepfert, et al. 2008, Martinez-Picado, et al. 2006, Rousseau, et al. 2008). The viral genes that encode proteins of important structural integrity (such as *gag*) or essential enzymatic activities (such as *pol*) tend to be more conserved (Frater, et al. 2007, Wang, et al. 2009). As shown in Figure 1.6, studies have suggested that the protective effect of the beneficial HLA class I molecules may have been exerted through effective CTL selection upon conserved viral genetic regions. This controls viruses bearing wild type sequences in the epitope, and positive selection of epitope mutants may potentially reduce the viral fitness of the immune escape mutants (Goulder and

Watkins 2008, Kiepiela, et al. 2007). This fitness impairment through CTL escape is further supported by reversion events observed in viruses upon transmission to an HLA-mismatched recipient. The reversion events also correlate with the rate of clinical progress and outcome (Duda, et al. 2009, Martinez-Picado, et al. 2006). Another pathway also exists for the immune escape viruses to regain fitness. Secondary mutations may also arise following emergence of viral variants that contain costly escape mutations, thereby partially restoring fitness (Brockman, et al. 2007, Crawford, et al. 2007).

Currently, CTL mediated selection is best described in the acute and latency phases of HIV infection (Baker, et al. 2009, Weiss 2008). The results of these findings have been informative in CTL vaccine design, despite failures in the first attempted CTL vaccine, the STEP trial (discussed in Chapter 1.7) (Baker, et al. 2009, Weiss 2008). Little is known regarding the roles played by CTL in progressive disease and the advanced stages of AIDS. It is not clear whether CTL continue to effectively control viral replication in AIDS, when the blood CD4 T helper cell counts are depleted. Does the immunodeficiency during AIDS impair CTL functions? Can the clinical immunodeficiency during AIDS impair CTL selection control over virus bearing mutations with significant fitness costs and create a suitable environment for the mutant to revert back into wild type form?

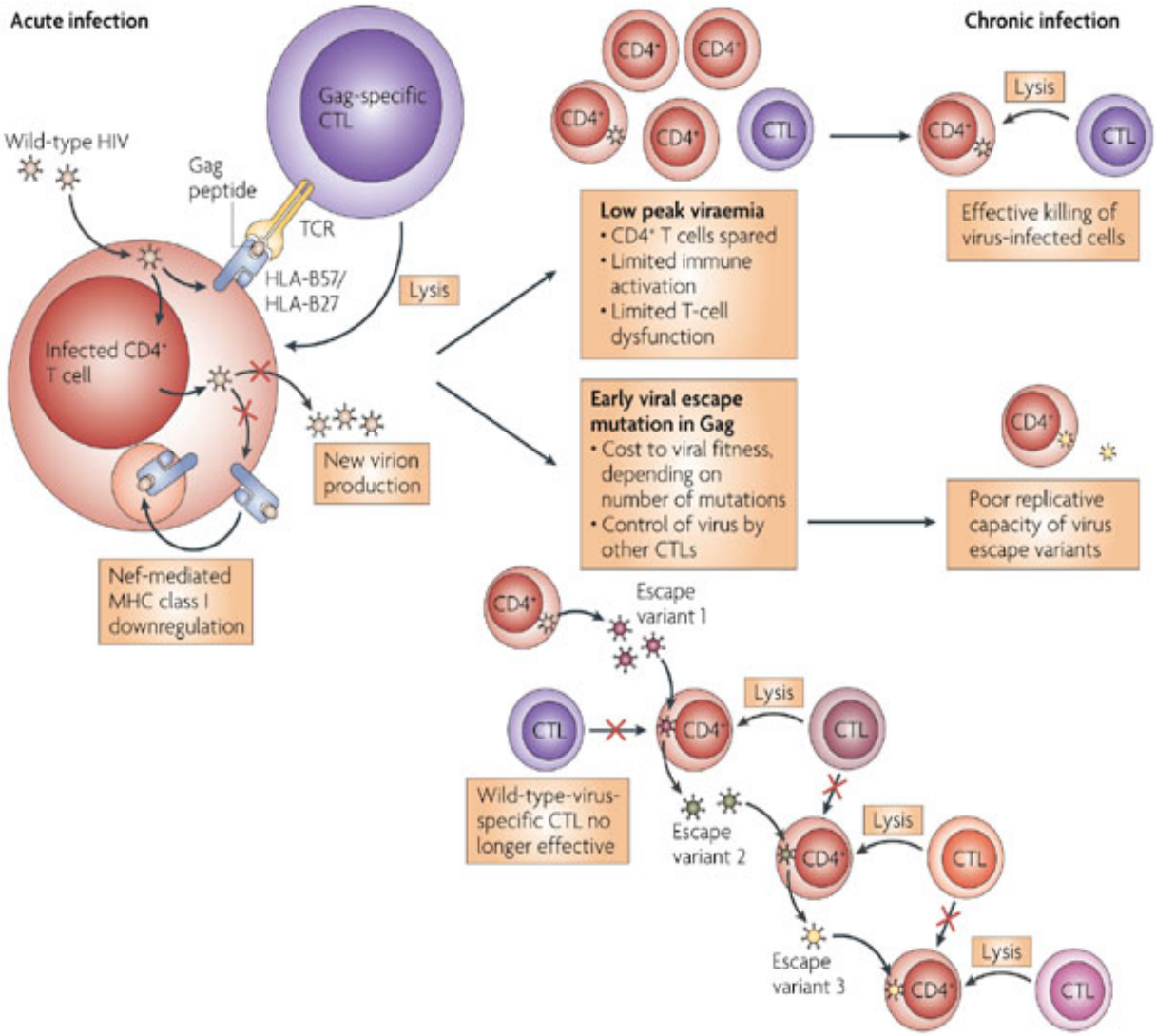


Figure 1.6: Schematic model of immune control of HIV mediated by HLA alleles presenting multiple Gag epitopes

(Image source: Goulder, P. et al. 2008, (Goulder and Watkins 2008))

1.5.2 Neutralising Antibodies (NAb)

Neutralising antibodies are immunoglobulins produced by B-lymphocytes that defend then host from pathogens by specific binding against antigenic active sites functional in infection, hence inhibiting the pathogen (Janeway, et al. 2005). HIV-specific NAb produced are readily detectable in patients' plasma, offering humoral defence against extracellular virions (Sarngadharan, et al. 1984). NAbs target glycoproteins gp120 and gp41 (encoded by the *env* gene) and block target cell binding and entry. The success of vaccine-elicited antibodies in preventing various infections has motivated development of anti-HIV NAb vaccines. The VaxGen AIDSVAX HIV-1 vaccine trial in 2003 attempted to induce protective HIV-specific NAbs, but failed to provide significant protection (McMichael and Hanke 2003, Montefiori, et al. 2007, VaxGen 2003). The partial success of the more recent trials in Thailand which include a gp120 component, have however, re-ignited interest in Nab-mediated protection from HIV infection.

Unlike CTL, NAb appear only 3 to 6 months after HIV infection, and do not appear to be involved in the resolution of primary viraemia (Ariyoshi, et al. 1992). HIV escapes from NAb rapidly through 4 key means: firstly, the *env* gene encoding glycoprotein is highly variable, particularly in domains targeted by NAb, with point mutations sufficient to permit escape (Korber, et al. 2001, McKeating, et al. 1989). Secondly, heavy glycosylation of envelope glycoproteins renders the virus poorly immunogenic during antigen processing (Mascola and Montefiori 2003). Thirdly, glycoprotein gp120 has flexible tertiary folds and a glycan shield that conceals functional or receptor-binding epitopes (Mascola and Montefiori 2003, Wei, et al. 2003). Lastly, viral debris is a powerful decoy to NAb, shifting selection pressure away from live

virus (Parren, et al. 1997). Antibody "escape" occurs constantly in natural infection, demonstrated by the observation that NAb isolated from a given patient are often effective against virus isolated from earlier time points, but rarely neutralise contemporaneous viral envelope (Richman, et al. 2003, Wei, et al. 2003).

Neutralising monoclonal antibodies (MAb) have shown some promise in delaying viral rebound in individuals coming off antiretroviral therapy (Trkola, et al. 2005). Furthermore, some antibodies do not directly inhibit viral entry into target cells, but exhibit antiviral activity through the Fc region of the antibody molecule. This leads to the complement binding induced viral lysis, phagocytosis of antibody coated virions, and antibody-dependent cellular cytotoxicity (Holl, et al. 2006, Holl, et al. 2006). There still exists a large knowledge gap in the HIV-NAb relationship and the role of B-lymphocytes in HIV infection (Montefiori, et al. 2007). The failure of a CTL-based vaccine, the STEP trial, has encouraged revitalised interest in antibody-based vaccines (Fauci, et al. 2008). Future study in this area will aim at deciphering mechanisms that will successfully elicit production of broad and potent anti-HIV antibodies in all the vaccine recipients (Scheid, et al. 2009, Walker, et al. 2009).

Further understanding into the roles played by NAb and B-lymphocytes in HIV control is important, and inducing effective NAb remains a central goal in preventive HIV vaccine design.

1.6 Pharmacological Selection: Antiretroviral Therapy (ARV)

Significant progress has been made in the development of ARV. The first antiretroviral drug, azidothymidine (zidovudine, originally developed to treat cancer), was discovered to inhibit the RT enzyme of HIV in 1986 (Yarchoan, et al. 1986). Since then, more inhibitor classes to other essential steps of viral replication have been discovered, as summarised in Figure 1.4 (Barbaro, et al. 2005). However, unlike the host immune selection targeting multiple residues, each ARV agent targets specific viral enzymes or structures. Therefore, viral escape mutations emerge rapidly whenever a single agent ARV is used, given the rapid mutational capability of HIV (Aboulker and Swart 1993). In 1996, the introduction of combination ARV, often termed highly active antiretroviral therapy (HAART), introduced a dramatic treatment response to patients. HAART delays disease progression, reduces AIDS mortality (by up to 70% annually) and partially restores CD4 T cells, but cannot clear HIV infection (Palella, et al. 1998).

Under multiple ARV drug selection (usually 3 agents, balancing clinical efficacy and patient tolerability), HIV is successfully suppressed (Larder, et al. 1995). However, given time or opportunity (such as when therapy is interrupted or not adhered to), HIV can acquire resistance to one or more agents in HAART regimens, requiring changing of failed ARV components. Drug toxicity is another major health concern (Carvajal-Rodriguez, et al. 2007, Larder, et al. 1993, Larder, et al. 1995). Although some drug escape mutants are impaired in replication fitness, others seem unaffected or may even have increased fitness (Armstrong, et al. 2009, Garcia-Lerma, et al. 2004, Prado, et al. 2002). Similar to escape seen in natural selection, reversion to wild type virus occurs rapidly following treatment cessation. Drug compensatory mutations also

arise secondarily to restore fitness in the mutants harbouring costly escape (Gandhi, et al. 2003, Stanford University 2007). Furthermore, intra-subtype and inter-subtype differences in HIV genetic sequences evoke different pathways leading to escape (Pieniazek, et al. 2000). Therefore, concerted efforts, surveillance programs and comprehensive databases have been set up to describe subtype-specific drug resistance mutations and their accessory mutations (which may either have predisposing, compensating or augmenting effects to drug concurrent resistance mutations) (Bennett, et al. 2009, Gifford, et al. 2007, Stanford University 2007).

ARV and HAART regimens were originally designed to inhibit subtype B HIV-1, but are efficacious against other subtypes, including C - the dominant subtype infecting South Africa (Frater, et al. 2001, Gordon, et al. 2003, Kantor, et al. 2005). In the *pol* gene subtype C varies from subtype B by 10-12% at the nucleotide level, yet these genotypic differences do not appear to confer major pre-therapy drug resistance, although may be associated with more accessory resistance mutations and different molecular pathways to resistance (Frater, et al. 2002, Gordon, et al. 2003, Pieniazek, et al. 2000, Robertson, et al. 2000, Sanches, et al. 2007, Velazquez-Campoy, et al. 2001). Ahead of the mass dispensing of HAART in South Africa, it is important to study the molecular characterisation and baseline pre-therapy drug resistance profiles of subtype C HIV-1 (Department of Health 2003, Gilks, et al. 2006). This will also contribute to the long-term follow up of resistance development.

1.7 Vaccine Design

Attempts to design satisfactory vaccines against HIV have been disappointing to date. Both major clinical trials aimed to elicit anti-HIV NAb (through HIV gp120 based antigen developed by VaxGen) or antiviral CTL response (through adenovirus vector conveying HIV *gag*/*pol*/*nef* antigens) have failed to provide significant protective or therapeutic effect (Letvin 2006, McMichael and Hanke 2003, Report 2007, VaxGen 2003). Nevertheless, these trials have yielded important results and lessons for next generation vaccine design. Given the above HIV escape strategies: rapid HIV mutation, extensive intra-subtype and inter-subtype genetic diversity amongst circulating forms, and inability of vaccines to elicit broad and potent responses (NAb or CTL) are potential factors leading to both vaccines' failure (Barouch and Korber 2010, Pantaleo 2008). Furthermore, immune escape mechanisms of envelope glycoproteins pose additional challenges to antibody based vaccine (Mascola and Montefiori 2003, Poignard, et al. 1999).

The CTL based vaccine faces different complicated challenges in its antigen, adjuvant and vector designs. In 2007, the STEP trial was terminated prematurely, after seeing failure to meet its primary endpoint and presence of a trend to increased in HIV infection amongst the vaccinated arm. Several ongoing investigations continue to assess attributing factors that may have caused the increased harm (Barouch and Korber 2010, Pantaleo 2008, Report 2007). Pre-existing NAb against vector adenovirus (recombinant adenovirus serotype 5, rAd5), circumcision history and HSV co-infection have been implicated as potential factors linked to the increased HIV infection amongst vaccinated arm of trial (Buchbinder 2009). The exact mechanism whereby pre-existing immune response towards rAd5 may enhance HIV infection

remains inconclusive, but its answer is important to the conceptual design of vaccine vectors (Hutnick, et al. 2009, O'Brien, et al. 2009).

Currently, both B cell and T cell vaccine initiatives are searching for means to elicit potent, broad and durable protection in every tissue compartment. It is also likely that the two vaccine mechanism need to be combined to achieve effective synergistic control over HIV (Barouch and Korber 2010, Letvin 2006, Weiss 2008). Further study into HIV escape strategies in response to different selection pressures should enhance understanding regarding viral evolution relating to vaccine candidates, whilst study of CTL selection in advanced HIV disease will assess the durability of CTL in a CD4 T cell depleted environment.

1.8 Summary and Research Questions

The HIV pandemic is a major global health crisis. The virus mutates fast to diversify and escape selection rapidly. Subtype C and B of HIV-1 are the most prevalent strains, and Southern Africa is challenged with the highest disease burden. Despite advances in the development of ARV and HAART that prolong survival, there remains no vaccine or cure after 25 years of intense research (Weiss 2008). The failure in both antibody and CTL based vaccine initiatives has revealed important lessons, and suggest better understanding is required regarding antiviral immunology and viral evolution. AIDS continues to be prevalent in resource-limited areas not accessible to HAART, and is the end stage disease for patients who fail available HAART regimens. The pathogenesis of AIDS remains unclear.

Viral evolution mediated by different co-existing selection forces (namely: antibody, CTL and ARV) during advanced HIV infection is complex. However, insights gained from each selection process may yield common insights into viral evolution, which could potentially enhance therapeutic design. This thesis therefore seeks to investigate key aspects of various selection forces and their related HIV evasion, including:

- The role of NAb and B-lymphocytes in chronic HIV infection (case study).
- In Free State province of South Africa:
    - To better define the molecular epidemiology of HIV variants
    - Investigate the prevalence and implication regarding pre-therapy ARV resistance.
    - Characterise the features of CTL selection in the HIV found in patients suffering from AIDS.
    - Describe HLA class I-mediated viral evolution in viral isolates found in patients suffering from AIDS.
    - Investigate viral replication fitness (VRC) of HIV isolated from patients of different clinical disease stages and HLA class I alleles.

CHAPTER 2:

Patients, materials and methods

2.1 Patients, cohorts and stratification

This thesis draws its findings based on the investigation done relating to one UK based clinical case study and two South African clinical cohorts.

2.1.1 Clinical case study (Chapter 3)

The patient who inspired our study in Chapter 3 (B cell depletion in vivo reveals a role for antibody control of HIV) is a Caucasian male attending St Mary's Hospital, London, United Kingdom. He was diagnosed with HIV seroconversion illness aged 58, in 2002. Three years earlier he was diagnosed with low-grade lymphoplasmacytoid lymphoma for which he did not receive treatment. His only other medical history of note was an acute hepatitis B infection some thirty years earlier. Following HIV-1 seroconversion he was recruited into a prospective, non-randomised observational study of early treatment and received three months of short course antiretroviral therapy (SCART) (Fidler, et al. 2002), and remained well until 2005 with a stable CD4 T cell counts and plasma HIV VL, as shown in Figure 3.1. Thirty months later he developed a rising paraproteinaemia, attributed to his lymphoma and received thalidomide treatment without clinical effect. He initiated Rituximab therapy 1075 days after HIV seroconversion and received four weekly

doses (375mg/m^2). He felt unwell 20 weeks after the first dose of Rituximab with malaise and fever associated with a marked rise in HIV pVL (increase of 1.7log_{10}, Figure 3.1). Additionally, he developed a biochemical hepatitis (attributed to reactivation of hepatitis B) which resolved with only supportive treatment. Antiretrovirals were initiated 1392 days after seroconversion.

Plasma viral samples are available from various important clinical event time points, labelled in days post-seroconversion (Figure 3.1): 1 month before rituximab therapy (d1042, T1), 6 weeks after rituximab (d1135, T2), immediately following viral load peak (d1270, T3), remission (d1295, T4), new baseline (d1353, T5) and after commencement of highly active anti-retroviral therapy (HAART, d1392, T6). The patient's previous commitment to SCART study offers additional proviral samples from peripheral blood mononuclear cells (PBMC) from seroconversion up to day 1042.

2.1.2 Clinical cohorts in South Africa (Chapter 4-6)

Two South African cohorts were studied: both the Free State province based Bloemfontein cohort (BFN) and the Kwazulu Natal province based Durban cohort (DBN) are established cohorts described previously (Huang and Frater et al, 2009, Kiepiela et al, 2007). In summary, each of the cohorts were comprised of pre-therapy adult patients chronically infected with HIV-1 subtype C (cohort summary and numbers are available in Table 5.1). The two cohorts neighbour geographically in the central east region of South Africa (Figure 2.1), and are both constituted with Xhosa and Zulu majority in terms of population ethnicity.

The Bloemfontein cohort was established in 2006 between the University of Free State and our group at University of Oxford, at the beginning of the provincial dispensing of HAART program – a part of governmental stepwise national provision of HAART therapy to eligible patients in primary care sector (Department of Health 2003). All patients had been identified through voluntary screening programmes in the Free State province and following diagnosis with HIV infection at local primary care clinics, were referred to district or regional centres. Here, the patients were counseled by trained HIV nurses and asked directly whether they had previously received PMTCT, mono-, dual-, or triple- ARV therapy. Patients who admitted previous drug exposure were excluded from this study. For all other patients, an extra 10ml of blood was taken at the same time as routine bloods for CD4 T cell counts and other clinical investigation. The University of the Free State and ACADIA consortium approved the ethics of this study (ETOV10/04 and ETOVS 206/05). The Department of Health of Free State Province granted approval to conduct the study.

In total, 1491 HIV-infected pre-therapy adult patients attended 7 of 15 government ARV clinics in Free State province during February and September 2006 have consented to the study enrolment. According to availability of CD4 T cell counts, CD4 T cell counts stratification, viral load, additional plasma availability for viral RNA extraction, success in viral genetic sequencing and HLA typing, different number of patients are included into the analysis across Chapter 4 to 7 (and described separately). The Bloemfontein cohort forms the core of analysis within these chapters, however, supplementation of Durban cohort (in Chapter 5) have improved the power of analysis significantly.

Figure 2.1: Map of South Africa and its provinces

Figure adapted from Wikimedia Commons (Commons 2010)

Left panel: South Africa (in red) at the south of Sub-Saharan Africa (boxed and enlarged from the world map in left lower corner).

Right panel: Free State province (in blue, capital: Bloemfontein) and KwaZulu Natal province (in purple, capital: Durban) are two of the nine provinces in central east region of South Africa.

The Durban cohort analysed in Chapter 5 was first established in 2002 by our collaborators, Philip Goulder (University of Oxford, UK), Bruce Walker (Harvard University, US) and Hoosen Coovadia (University of KwaZulu Natal, RSA). The cohort has since expanded vastly in scale to include over thousands of chronic HIV patients and described important findings regarding CTL mediated HIV selection (Frater, et al. 2007, Kawashima, et al. 2009, Kiepiela, et al. 2004, Matthews, et al. 2008). Our collaborators have kindly supplied us information regarding 775 pre-therapy chronic adult patients who possess near-complete details regarding HLA types, HIV *gag*/*pol*/*nef* sequences, clinical viral load, CD4 T cell counts and ELISPOT responses against consensus overlapping peptides. Nearly one third of the patients (n=219) in Durban cohorts were tested for HIV infection during pregnancy visit to antenatal clinics (ANC), the remaining of the subjects (n=556) were recruited from outpatient follow-up clinics. This latter group was mostly diagnosed following clinical presentation suggestive of HIV infection. This study was approved by institutional review boards, and all subjects gave written informed consent.

2.1.3 Cohort stratification by CD4 T cell count

The studies in Chapter 3 through to Chapter 6 focus on the study of viral molecular epidemiology and evolution amongst patients suffering advanced AIDS. Therefore, we defined "low" CD4 T cell counts in our study as <100 CD4 cells/μl, significantly lower than the conventional clinical AIDS definition (HIV/AIDS, et al. 1993) to emphasise the severity and duration of advanced HIV infection. Other CD4 T cell counts strata are used (and named) in relation to "low" CD4 T cell counts to offer comparison in event observations, including: 'intermediate' (200-400 CD4 cells/μl) and 'high' (>500 CD4 cells/μl) CD4 cell counts stratification groups.

2.2 Preparation of samples

2.2.1 Preparation of patient blood samples

For all the patients in Bloemfontein cohort, an extra 10ml of blood was taken at the same time as routine bloods for CD4 T cell counts and other necessary investigation. According to governmental HAART provision guideline, VL is not routinely counted for every patients, and is only requested where clinically indicated or essential to our analysis. The VL measurement was undertaken using the Roche Amplicor version 1.5 assay, where applicable.

The remaining (additional) blood stored in EDTA tubes is centrifuged at 3000 rpm for 15 minutes at room temperature. The supernatent was removed, transferred to a sterile tube for a further 15 minutes at 2000 rpm, to remove potential cellular contamination. The supernatant (containing HIV-1 RNA) and whole blood cells (containing host DNA and cells for HLA typing) are aliquoted into 3 separate cryotubes and stored at -80 ℃ facilitie.

The blood samples of the Durban cohort were processed separately by our collaborators, and have been extensively described in other literatures (Leslie, et al. 2004). In additional to Bloemfontein cohort, the Durban cohort has additional frozen PBMC available for ELISPOT assay.

2.2.2 Preparation of DNA for HLA typing

In the Bloemfontein cohort, only the individuals from 'high' (n=110) and 'low' (n=195) CD4 T cell counts categories were studied for their HLA types. Patients' HLA type was determined to the oligo-allelic level using Dynal RELITM Reverse Sequence-Specific Oligonucleotide kits for the HLA-A, -B and –C loci (Dynal Biotech). To obtain four-digit typing, Dynal Biotech Sequence-Specific priming kits were used, in conjunction with the Sequence-Specific Oligonucleotide type. In total, 260 of 305 patients with adequate and sufficient samples were successfully typed (Table 5.1). 71 HLA class I molecules (27 HLA-A, 27 HLA-B and 17 HLA-C, Table 3-2) expressing at ≥0.5% phenotypic frequency in the cohort were included in the analysis.

In Chapter 5, the HLA typing of the Durban cohort (n=775) were carried out separately by our collaborators, and have been extensively described in other literatures (Kiepiela, et al. 2007).

2.2.3 Media

<u>R10 medium (also known as RPMI/P/S/FCS meduium)</u>

R10 medium was mainly used for the culture of CEM-GXR cell line. The RPMI 1640 medium (Moore, et al. 1967), supplied by Sigma-Aldrich was supplemented with L-glutamine (146mg/L), Penicillin G (100IU/ml) and Streptomycin sulphate (100μg/mL). Fetal calf serum, inactivated by incubation at 56℃ for 60 minutes, was added to a final concentration of 10% by volume.

Luria-Bertani Medium (LB broth)

The LB broth was used mainly for the culture of *E. Coli* bacteria in liquid medium. The LB broth tablet (Sigma-Aldrich) is concentrated, sufficient to prepare 50mL of medium. Each litre of LB broth comprises of 10g enzymatic digest of casein, 5g yeast extract, 5g of NaCl and 2g of inert agent necessary for tableting process. The tablets are dissolved in required amount of deionised water and autoclaved. If required, Ampicillin was added to a final concentration of 100μg/mL (LB/Amp broth).

LB/Amp Agar bacterial culture plates

The LB/Amp agar plate was used mainly for the culture of *E. Coli* bacteria successfully transformed with anti-ampicillin resistant plasmid (carrying gene insert of interest) on solid medium on Petri dish. The LB agar tablets (Sigma-Aldrich) are concentrates, each sufficient for the preparation of 50mL of agar medium. Each litre of LB agar comprises of 9.14g enzymatic digest of casein, 4.57g yeast extract, 4.57g of NaCl, 13.72 bacteriological agar and 1.6g of inert agent necessary for tablet-making process. The tablets are dissolved in required amount of deionised water and autoclaved. Ahead of plating, agar are dissolved in microwave, and cooled to 56℃. Ampicillin was then added to a final concentration of 100μg/mL, and then 25-30ml of medium are poured into 100mm culture dishes, which were then labeled and stored at 4℃.

2.2.4 Competent bacteria cell lines

TOP10 chemically competent *E. Coli*

One Shot TOP10 (Invitrogen) *E. Coli* was used for clonal sequencing study of the HIV *env* in Chapter three. The bacteria grows in LB broth and agar, and with successful transformation with PCR amplified product insertion into the plasmid encoding Ampicillin resistance (See below at Chapter 2.3.4), the transformed TOP10 *E. Coli* clones grow steadily in LB/Amp broth and agar.

Stbl3 chemically competent *E. Coli*

Stbl3 (Invitrogen) *E. Coli* was used for the amplification of plasmids encoding HIV (or its component) necessary for the methods in Chapter six. Compared to TOP10, the Stbl3 bacteria have reduced frequency of unwanted homologous recombinations of long terminal repeats (LTR) in retroviral vectors, greater stability and increased DNA yield. The competent cells grow in LB broth and agar, and with successful transformation with plasmids encoding DNA of HIV (or its component) and Ampicillin resistance gene, the *E. Coli* then grows steadily in LB/Amp broth and agar. The LB/Amp broth prepared with glycerol (at a volume ratio of 2:1) is suitable for storage at -70℃.

2.2.5 Plasmids and preparation

Plasmid pNL4-3

Plasmid pNL4-3 was obtained through our collaborator, Zabrina Brumme and Mark Brockman, originally acquired from the AIDS Research and Reference Program, Division of AIDS, NIAID, NIH, from Dr Malcolm Martin (GenBank ref AF324493). It is a full-length replication competent DNA of HIV-1 subtype B derived from NY5 (5') and LAV (3') (Adachi, et al. 1986). The Ampicillin resistant plasmid was

provided as DNA (100µl of 630µg/mL) in TE buffer, transformed into Stbl3 *E. Coli* and grown in LB/Amp Agar and broth.

Modification of this well described laboratory strain virus has led to the development of its related plasmid DNA construct, pNL4-3Δ*gag-protease*, permitting generation of chimeric viruses (see below and Chapter six) (Miura, et al. 2009). A diagram of the plasmid, identifying relevant restriction sites is presented in Figure 2.2. The original plasmid do not contain any *BstEII* restriction enzyme (RE) sites (purple dotted line, Figure 2.2), and mutagenesis are required to engineer and introduce the two *BstEII* RE sites into pNL4-3 genome and create deletion over *gag-protease* gene fragment (see below).

Plasmid pMJ4

Plasmid pMJ4 was obtained through our collaborator, Zabrina Brumme and Mark Brockman, originally acquired from the AIDS Research and Reference Program, Division of AIDS, NIAID, NIH, from Dr Thumbi Ndung'u (GenBank ref AF321523). It is a full-length replication competent DNA of HIV-1 subtype C molecular clone derived from a Southern African (Botswana) isolate (Ndung'u, et al. 2001). The Ampicillin resistant plasmid was provided as DNA (100µl of 420µg/mL) in TE buffer, transformed into Stbl3 *E. Coli* and grown in LB/Amp Agar and broth. The virus grows considerably slower in *in vitro* viral growth assays when compared with subtype B laboratory strains (like NL4-3), and previous attempts to engineer HIV-1 subtype C chimeric virus such as pMJ4Δ*gag-protease* have not been successful (Ndung'u, et al. 2009).

Plasmid pNL4-3Δ*gag-protease*

The pNL4-3Δ*gag-protease* was obtained through our collaborator, Zabrina Brumme and Mark Brockman, who have contributed to the engineering of the plasmid construct (Miura, et al. 2009). The pNL4-3Δ*gag-protease* is essentially a pNL4-3 plasmid deleted with its original *gag-protease* gene, and is replication incompetent unless recombined with correctly amplified *gag-protease* PCR product (from autologous product, for example).

The pNL4-3Δ*gag-protease* is engineered based on serial mutagenesis of pNL4-3 that leads to introduction of two unique restriction enzyme (*BstEII*) sites: one at the 5' end of the *gag* gene and another one at 45 bases downstream from the 3' end of the *protease* gene of pNL4-3. The *BstEII* enzyme digest results in deletion of HIV *gag-protease* and self ligation of the remaining plasmid (pNL4-3Δ*gag-protease*). The circular pNL4-3Δ*gag-protease* plasmic with self-ligated *BstEII* site can be digested (and therefore linearised) during future transfection experiment, and recombine with co-transfected HIV *gag-protease* PCR product containing *BstEII* sites at its end, through homologous recombination. The detail of these methods are described in previous study and Chapter 2.3.6 below (Miura, et al. 2009).

Plasmid DNA extraction and purification

Plasmid DNA was extracted using Qiagen Miniprep Kit or Qiagen Maxiprep Kit, according to manufacturer's instruction (Qiagen). The quality and yield of the purified plasmid DNA are assessed by nanodrop, and stored at -20℃ follow satisfactory measurements.

Restriction enzyme digestion of plasmid

The pNL4-3 and pNL4-3Δ*gag-protease* differ in their genetic sequences surrounding *gag-protease* region, and each plasmid possess different fragmentation pattern following restriction enzyme digestion using *Hind III* and *BstE II* (New England Bio-Lab, UK, according to manufacturer's instruction) (Figure 2.2). Unique restriction sites in DNA were checked using Sequencher software (version 4.8, Gene Codes Corporation and its licensors), which detects all restrictions sites of restriction enzymes in a given uploaded sequence.

The circular pNL4-3 contains 5 *Hind III* sites (in red solid line) and no *BstE II* sites, in contrast to the circular pNL4-3Δ*gag-protease* that contains 4 *Hind III* sites and 1 *BstE II* site (in purple solid line). The differences in restriction enzyme sites is due to the deliberate modification of pNL4-3 that derived pNL4-3Δ*gag-protease*, two *BstE II* sites were introduced via mutagenesis at either end of *gag-protease* genes (in dotted purple line, Figure 2.2, also discussed in Chapter 2.2.5). The subsequent *BstE II* digestion then removes the *gag-protease* gene from pNL4-3, leading to self ligation and hence formation of pNL4-3Δ*gag-protease*.

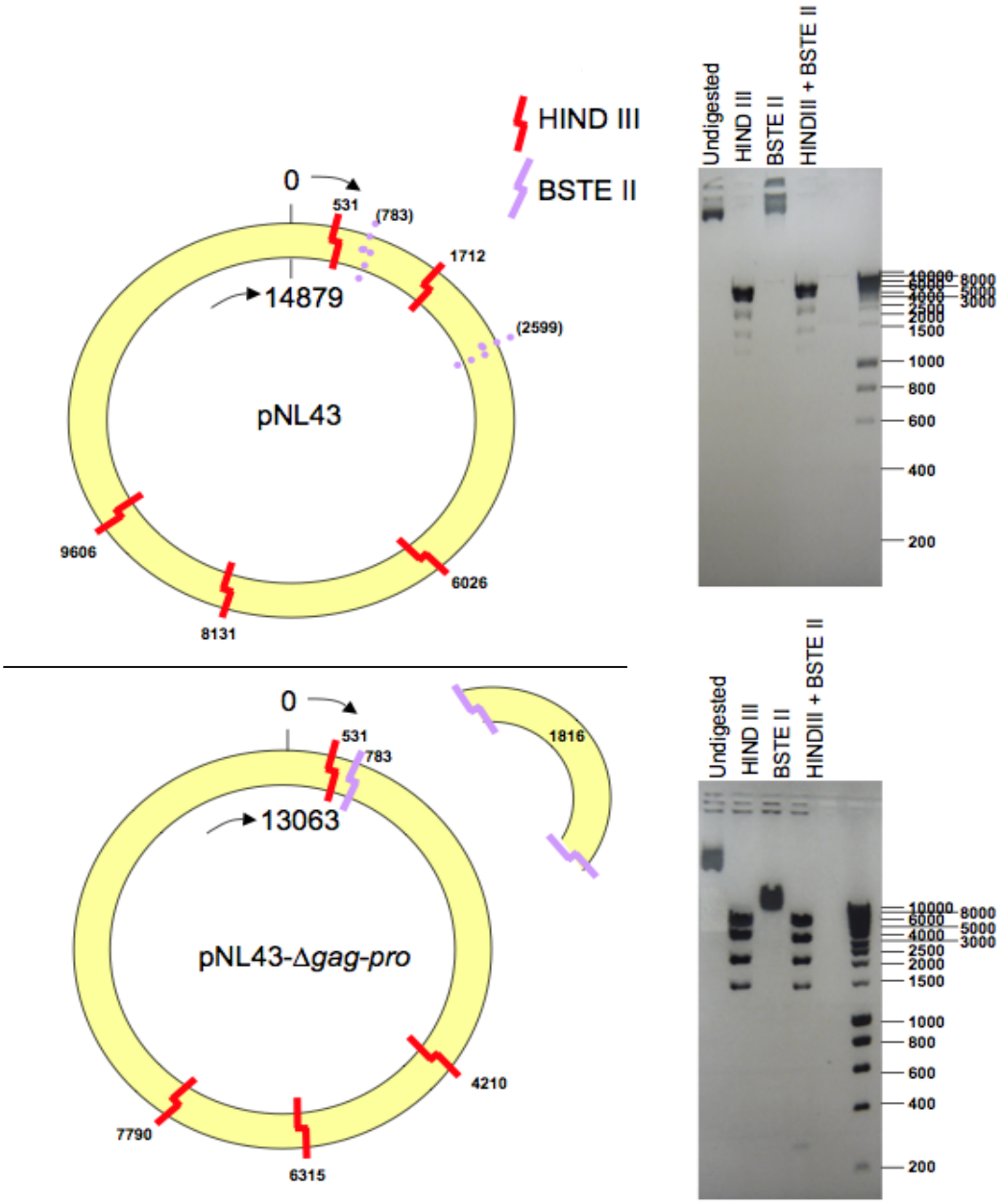


Figure 2.2: Schematic map of plasmid pNL4-3 and pNL4-3Δ*gag-protease* and restriction enzyme digestion sites

Upper panel: circular plasmid, pNL4-3, with *HIND III* restriction enzyme (RE) sites (in red). The numbers demarcate the nuclotide position of RE digestion. The size of RE digested product can be confirmed on gel electrophoresis (right). In ordinary pNL4-3, there are no *BstE II* RE sites. However, mutagenesis method can introduce novel *BstE II* sites (dotted purple lines) at either end of *gag-protease* region of pNL4-

3. The subsequent *BstE II* digestion then removes the *gag-protease* gene from pNL4-3, leading to self ligation and hence formation of pNL4-3Δ*gag-protease* (Lower panel). The circular pNL4-3Δ*gag-protease* contains different number of *HIND III* sites (red) and *BstE II* site (in purple solid line), and therefore different length polymorphism on gel electrophoresis.

2.2.6 Visualisation of DNA by agarose gel electrophoresis

Plasmid and PCR DNA products were visualized by agarose gel electrophoresis. Typically, amplified DNA products were mixed at a 1:1 ratio with a gel loading buffer (Bioline UK). This was then loaded onto a 1% agarose gel (Sigma UK) in TBE buffer (Tris-borate-EDTA, Promega, UK). Ethidium bromide solution (10 mg/ml in H_2O) was added at 5μl per 100ml of gel. Samples were run in parallel with Hyperladder I (Bioline UK) to allow an estimation of the size and concentration of the DNA fragments. The gel was run by electrophoresis at 70V for 1.5 hours. The DNA in the gel was then visualized under UV light using a Gel Doc 2000 (Biorad).

2.2.7 DNA extraction from agarose gel

Purification of DNA products by gel extraction was performed using a Gel Extraction Kit (Qiagen) according to the manufacturers instructions. DNA was run on a 1% agarose gel at 70V for 1.5 hours and extracted manually under UV light to visualize the correct bands. Gel cutting under UV light was performed swiftly to prevent UV-damage to the DNA. DNA was eluted in RNase/DNase-free H_2O (Sigma) and either stored at -80^{o}C or used immediately.

2.2.8 CEM-GXR reporter cell lines for HIV infection

The human-derived stable CD4 T cell line (CEM) that permit HIV infection *in vitro* was modified to report HIV infection status through Tat-driven GFP production (CEM-GFP cells) (Gervaix, et al. 1997). Our collaborators have made further modifications to the CXCR4 receptor positive CEM-GFP cells lines (via transfection) that allow co-expression of CCR5 receptors (CEM-GXR cells) (Brockman, et al. 2006). Therefore, the CEM-GXR cells are permissive to infection by both CXCR4

and CCR5-tropic HIV strains. In Chapter 6, the GFP fluorescence is used to indicate infection status of the CEM-GXR cells in the viral replication capacity (VRC) assay.

2.3 Virology

2.3.1 HIV-1 viral RNA extraction

HIV virions were pelleted from 500μL patient plasma by centrifugation at 23,500g for 1 hour at 4℃. 360μl of supernatant was then removed, leaving the remaining solution (140 μL) containing enriched viral particles. The viral RNA was then isolated from the enriched plasma using the QIAamp Viral Kit (Qiagen, UK), as per the manufacturers instructions. The protocol adopted also includes the optional addition of RNA-free DNAse (Qiagen, UK) to eliminate DNA contaminants and RNAase inhibitor RNAsin (Promega, UK). Extracted RNA was immediately converted to cDNA by RT-PCR, with the remaining RNA stored at -80℃ ready for repeated RT-PCR, if required.

2.3.2 Synthesis of complementary DNA (cDNA) by Reverse Transcription (RT)

The extracted RNA was reverse transcribed using the *Superscript RT III* enzyme (Invitrogen, UK), as per manufacturer's protocol. Random decamers was preferentially used to prime the synthesis of *gag* and *pol* cDNA, whereas oligoDT was the preferred method used to prime the synthesis of *env* and *nef* cDNA. However, the cDNA product synthesised from the two priming methods served as useful backup resource when cDNA synthesis failed in either of the methods. Some of the *gag*-

protease cDNA was synthesised using specific outer downstream (antisense) primer of the subsequent nested PCR (primer name: 2cRxOP of Table 2.1).

2.3.3 PCR of HIV gag, pol, gag-protease, env and nef

Amplification of *gag*, *pol*, *gag-protease, env* and *nef* by PCR was carried out using a two-stage nested reaction with two pairs of primers (termed outer primer pairs, OP, and inner primer pairs, IP). The primer sequences and references are detailed in Table 2.1. Two primer sets were designed for the most of the genes, which were used on samples not amplified by the first set of primers. All primers were unmodified oligonucleotide primers purchased from MWG Biotech, Germany.

The PCR reaction was performed using the Platinum *Taq* DNA polymerase enzyme (Invitrogen, UK) in a preparation of 50 μL:

38.3 μL	H_2O (DNAse/RNAase free; Sigma)
5 μL	10x PCR buffer (Invitrogen)
1.5 μL	$MgCl_2$ (50mM; Invitrogen)
1 μL	DNTP (25mM; Invitrogen)
1 μL	5' primer (20 pmol/μL)
1 μL	3' primer (20 pmol/μL)
0.2 μL	Platinum *Taq* (Invitrogen)
2 μL	cDNA template (for first round of nested PCR), or first round PCR DNA product (for second round of nested PCR)

The PCR reaction was performed on a Peltier Thermal Cycler PTC-200 (MJ Research, UK) using the following two-step cycling ("touch down") conditions:

Step 1	Denaturation	94℃	2 min
Step 2	Denaturation	94℃	15 sec
Step 3	Annealing	X℃+3℃	30 sec
Step 4	Elongation	72℃	(1 min per kb of expected product size)
Step 5		Repeat steps 2-4, for 19 times	
Step 6	Denaturation	94℃	15 sec
Step 7	Annealing	X℃	30 sec
Step 8	Elongation	72℃	(1 min per kb of expected product size)
Step 9		Repeat steps 6-8, for 19 times	
Step 10	Elongation	72℃	7 min
Step 11	Storage	4℃	

The temperatures of the annealing steps (3 and 7) are calculated according to the melting temperatures of the specific primer pairs being used. Hence the X℃ marked above can be calculated by taking the average of two melting temperature (Tm) amongst the primer pairs (Table 2.1), with the exception of *gag-protease* amplification (where the annealing temperatures of step 3 and 7 are 58℃ and 55℃ for

OP, and 60℃ and 57℃ for IP, respectively). The purpose of "touch down" thermocycling setting (where the annealing temperature of earlier cycles, step 3, is 3℃ higher than that of later cycles, step 7) is attempted to increase binding specificity during early cycles of amplification.

Table 2.1: Primers used in the PCR of HIV *gag, gag-protease, pol, env* and *nef*

Nested PCR primer name and pairs	Primer sequence	HXB2 nucleotide position	Tm (℃)	HIV Clade
Gag	(Leslie, et al. 2004)			
(Chapter 5 and 6)				
F1OLD5OP	CTAGCAGTGGCGCCCGAACA	629 – 648	64	C
F1OLD3OP	GCAGTCTTTCATTTGGTGTCCTTC	2067 – 2044	59	C
F1OLD5IP	TCTCTCGACGCAGGACTC	682 – 699	58	C
F1OLD3IP	TTTCCACATTTCCAACAGCC	2042 – 2023	55	C
F1NEW5OP	CTCTAGCAGTGGCGCCCGAA	627 – 646	58	C
F1NEW3OP	TCCTTTCCACATTTCCAACAGCC	2045 – 2023	55	C
F1NEW5IP	ACTCGGCTTGCTGAAGTGC	696 – 714	53	C
F1NEW3IP	CAATTTCTGGCTATGTGCCC	2003 – 1984	52	C
Gag-Protease	(Miura, et al. 2009)			
(Chapter 6)				
623OP	AAATCTCTAGCAGTGGCGCCCGAACAG	623 – 649	68	B/C
2cRxOP	TTTAACCCTGCTGGGTGTGGTATYCCT	2851 – 2825	66	C
5'UTR2OP	CACTGCTTAAGCCTCAATAAAGCTTGCC	512 – 539	65	C
GR2635OP	TCTTCTGTCAATGGCCATTG	2635 – 2616	55	C
GagProChimUnivIP	GACTCGGCTTGCTGAAGCGCGCACGGCAAGA GGCGAGGGGCGGCGACTGGTGAGTACGCCAA AAATTTTGACTAGCGGAGGCTAGAAGGAGAG AGATGGG	695 – 794	>75	B/C
GagProChimClaCIP	GGCCCAATTTTTGAAATTTTTCCTTCCTTTTCC ATTTCTGTACAAATTTCTACTAATGCTTTTATT TTTTCTTCTGTCAATGGCCATTGTTTAACTTTTG	2704 – 2605	>75	B/C
Pol	(Leslie, et al. 2004)			
(Chapter 4 and 5)				
F2OLD5OP	ATAGTAAGAATGTATAGCCCTGT	1606 –1628	55	C
F2OLD3OP	ATGGAGTTCATACCCCATCC	3254 – 3235	57	C
F2OLD5IP	GAGTGGGAGGACCTAGC	1844 – 1860	58	C
F2OLD3IP	TACAAGTCATCCATATATTGATAG	3112 – 3089	54	C

F2NEW5OP	GACATAAAACAAGGGCCAAAGGAA	1639 – 1662	59	C
F2NEW3OP	GGCAGCTGTATAGGCTGTACTG	3289 – 3268	62	C
F2NEW5IP	AGGACCTAGCCACAAAGC	1851 – 1868	56	C
F2NEW3IP	TTCTAAGTCAGATCCTACATACA	3131 – 3109	55	C
F3OLD5OP	GAAGGACAGTACTAAGTGGAG	2744 – 2764	58	C
F3OLD3OP	CTGGCTACATGGACTGCTAC	4471 – 4452	59	C
F3OLD5IP	ACTGCATTCACCATACCTAG	2931 – 2950	55	C
F3OLD3IP	TTATGTGCTGGTACCCATGA	4168 – 4149	55	C
F3NEW5OP	TATACTGCATTCACCATACC	2928 – 2947	53	C
F3NEW3OP	CCACTGGCTACATGGACTGCTA	4474 – 4453	62	C
F3NEW5IP	TGGAAAGGATCACCAGCAAT	3006 – 3025	55	C
F3NEW3IP	TACTTGTCCATGCATGGCT	4391 – 4373	55	C
Env	(Oelrichs, et al. 2000)			
(Chapter 3)				
Env10F	CTATGGCAGGAAGAAGCGGAGAC	5968 – 5990	64	B
Env11F	GCATGAGGATATAATCAGTTTATGGG	6536 – 6561	60	B
Env12R	CAAATGAGTTTTCCAGAGCAACCC	8035 – 8012	61	B
Env13R	CTATCTGTCCCCTCAGCTACTGC	8707 – 8685	64	B
Nef	(Leslie, et al. 2004)			
(Chapter 5)				
F8OLD5OP	GATCCATTCGATTAGTGGA	8455 – 8473	55	C
F8OLD3OP	CTTATATGCAGCATCTGAGG	9517 – 9498	50	C
F8OLD5IP	TTCAGCTACCACCGATTGAGA	8520 – 8540	52	C
F8OLD3IP	TGAGGGTTGGCCACTCC	9502 – 9486	52	C
F8NEW5OP	ACTCACCTTTGTCGTTTCAGA	8359 – 8379	50	C
F8NEW3OP	AGCTGCTTATATGCAGCATC	9521 – 9507	50	C
F8NEW5IP	CCCTTACCCCAAACCCGAGG	8380 – 8399	58	C
F8NEW3IP	CACTCCCCAGTCCCGCCC	9491 – 9474	59	C

OP = outer primer (first round of nested PCR)

IP = inner primer (second round nested PCR)

2.3.4 Bacterial cloning

In Chapter 3, the diversity and selection dynamic within HIV *env* of the virus isolated from different clinical stages were studied by clonal analysis. Here, 5μL of gel purified *env* PCR products (from different time point) were cloned using the TOPO TA cloning kit for sequencing (Invitrogen, UK), as per manufacturers instructions, followed by transmission of sequence-containing vector into competent *E.Coli (TOP 10* that confers resistance to kanamycin) cells achieved by heat shock for 30 sec at 42 ℃. Clonal sequence material was isolated from E.Coli using 96 wells mini-prep plate (Millipore, UK). TOPO TA cloning kit provides sequencing primer pairs flanking the PCR insertion regions that permit sequencing of the PCR inserts.

2.3.5 Sequencing of HIV gag, pol, gag-protease, env and nef

Sequencing of purified PCR products and bacterial clonal amplicons were performed using six to eight overlapping sequencing primers that binding both the sense and anti-sense strands of DNA, for increased coverage and accuracy. The inner primers (IP, second round) of the nested PCR were included as two of the sequencing primers, with another four to six additional primers designed to each confidently sequence 700 bp whilst permiting at least 200bp of sequencing overlap (Table 2.2). Sequencing was performed using the Big Dye Terminator v3 reaction mix (Applied Biosystems, UK). The reaction mix was as follows; 3.7μL ddH_20, 1.5μL BigDye buffer, 2μL primer (0.625μM), 0.8μL BigDye Terminator v3 reaction mix and 2μL PCR template. The sequencing reaction was performed on a Peltier Thermal Cycler PTC-200 (MJ Research, UK) and the program was as follows: 30 cycles of [96℃ for 10 sec, 50℃ for 5 sec, and 60℃ for 2 min].

Sequencing products were then cleaned by sodium acetate precipitation as follows: 10μL of ddH_2O, 2μL of NaOAC (3M, pH 5.2, Sigma, UK) and 50μL of 100% ethanol were added to the sequencing reaction, vortexed and incubated at RT for 30 min. This was centrifuged at 4000rpm for more than 80 minutes at 4℃, the supernatant removed and 150μL of cold 70% ethanol added, and then centrifuged under the same conditions for 10min. This wash step was repeated twice and the pellet dried at 94℃ for 1 min. The cleaned sequencing reaction were analysed on the ABI 3700 DNA analyzer (Applied Biosystems, UK) by Zoology department, University of Oxford.

The Sequences were aligned manually using X11, Sequencher (version 4.8, Gene Codes Corporation) and Se-Al (version 2.0a11, A. Rambaut, Department of Zoology, University of Oxford) software.

Table 2.2: Primers used in the sequencing of HIV *gag, gag-protease, pol, env* and *nef*

Clade	Primer name	Sequence
Gag	(Chapter 5 and 6)	(Leslie, et al. 2004)
C	F1-1 (F1OLD5OP)	TCTCTCGACGCAGGACTC
C	F1-2	CTGCACTATAGGATAATTTTGAC
C	F1-3	GACACCAAGGAAGCCTTAG
C	F1-4	CTCCCACTGGAACAGGTG
C	F1-5	GGAACAAATAGCATGGATGAC
C	F1-6 (F1OLD3IP)	TTTCCACATTTCCAACAGCC
Gag-protease	(Chapter 4 and 5)	(Miura, et al. 2009)
B/C	GP-1F	TGACTAGCGGAGGCTAGAA
C	GP-2F	CAGCATTATCAGAAGGAGCCAC
C	GP3-F	AGAAATGATGACAGCATGTC
C	GP-4F	GGAAACCAAAAATGATAGGAGGAATTGGAGG
C	GP-1R (GR2635OP)	TCTTCTGTCAATGGCCATTG
C	GP-2R	TTCCCTAAAAAATTAGCCTG
C	GP-3R	TTCCTGCTATGTCACTTCC
C	GP-4R	TCTTGTGGGGGTGGCTCCTTCTG
Pol	(Chapter 4 and 5)	(Leslie, et al. 2004)
C	F2-1 (F2NEW5IP)	AGGACCTAGCCACAAAGC
C	F2-2	ATCATCTGCTCCTGTGTCTA

C	F2-3	AGGCCAGGGAATTTCCTTCA
C	F2-4	CATTGTTTAACCTTTGGGCCATC
C	F2-5	CCCATTGAAACTGTACCAG
C	F2-6 (F2OLD3IP)	TACAAGTCATCCATATATTGATAG
C	F3-1 (F3NEW5IP)	TGGAAAGGATCACCAGCAAT
C	F3-2	TGGGTCATAATATACTCCATGTA
C	F3-3	CAGACATAGTACCACTAACTGAA
C	F3-4	CATAGAAAGTTTCTGCTCCT
C	F3-5	CTCCCCTAGTAAAATTATGGTA
C	F3-6 (F3NEW3IP)	TACTTGTCCATGCATGGCT
Env	(Chapter 3)	(Oelrichs, et al. 2000)
B	ENV11F	GCATGAGGATATAATCAGTTTATGGG
B	ENV12F	AGGAGGGGACCCAGAAATTG
B	ENV11R	TGGGAGGGGCATACATTGC
B	ENV12R	CAAATGAGTTTTCCAGAGCAACCC
Nef	(Chapter 5)	(Leslie, et al. 2004)
C	F8-1 (F8OLD5IP)	TTCAGCTACCACCGATTGAGA
C	F8-2	ACCTGAGGTCTGACTGGAAAG
C	F8-3	AGATAAACATGGGGCACTTAC
C	F8-4	AGCAGTCTTTGTAATACTCCG
C	F8-5	ACCTCAGGTGCCTTTAAGAC
C	F8-6 (F8OLD3IP)	TGAGGGTTGGCCACTCC

2.3.6 Generation of chimeric NL4-3 viruses carrying autologous HIV gag-protease from specific hosts

Autologous *gag-protease* from specific patient was amplified using PCR method described in Chapter 2.3.4. The second round of the nested PCR was conducted using 100-mer primer pair (GagProChimUnivIP amd GagProChimClaCIP Table 2.1) that completely matched the pNL4-3 sequences. Chimeric viruses were constructed by recombining the *gag-protease*-deleted pNL43 (pNL4-3Δ*gag-protease*) with autologous *gag-protease* PCR product. The methodology was developed in earlier studies for subtype B chimeric viruses (Miura, et al. 2009). Here we have adopted the protocol that leads to the construction of subtype B/C recombinant, where the HIV-1 subtype B backbone (pNL4-3Δ*gag-protease*) recombines with autologous HIV *gag-protease* amplified from South African HIV-1 subtype C infected patients (Chapter 6). The protocols of chimeric generation is identical to that of original subtype B chimeric viruses generation, with exception of the primers used in the first round of nested PCR are adjusted to suit amplification of subtype C viruses.

For transfection, the pNL4-3Δ*gag-pro*tease was linearised by *BstE II* digestion for two hours at 60°C (10U/ µl of enzyme per 1000 µg/ mL of plasmid). For each sample, 10 µg of the linearised pNL4-3Δ*gag-pro*tease was co-transfected with 5µg of purified PCR product of autologous *gag-protease* from patients into 3.9 x 10^6 GXR-CEM reporter T cells in 300 µl of R10 Plus medium. Transfection was achieved by electroporation (250 V, 950 µF, for 30 – 40 msec). After five minutes rest the product was transferred to a T25 flask (pre-warmed at 37°C) containing 10 ml R10+, 10µl polybrene (4µg/µl) and 1 x 106 GXR cells and incubated at 37°C. Positive control samples were generated using co-transfection of pNL4-3Δ*gag-pro*tease with *gag-*

protease amplified from pNL4-3, pHXB2 (subtype B), and pMJ4 (subtype C). The negative control samples had only pNL4-3Δ*gag-pro*tease electroporated, without co-transfection with *gag-protease* PCR product. The efficiency of the transfection was 86.6% in average.

2.3.7 Viral replication capacity (VRC) assay and flow cytometry

After five days GFP expression was measured every 1~2 days using a FACSCalibur flow cytometer (FACSCalibur; BD Biosciences, San Jose, CA), as per manufacturer's instruction. Supernatant was harvested when >15% of GXR-CEM cells were GFP+, and stored at -80°C. The recombinant viral culture yielding less than 15% GFP after 28 days post-transfection were discarded. Viral titration was then performed over two days, by addition of 400µl of supernatant to 1 x 10^6 GXR cells (1ml R10 Plus medium was then added after 20 hours). GFP+ percentage was measured after 48 hours to determine the multiplicity of infection (MOI), this yields the titration factor that allow VRC to be performed at the MOI of 0.003 (however, for practical purpose, we accepted MOI range between 0.0015 and 0.0045).

The viral replication capacity (VRC) was then assayed: 400µl of thawed supernatant was titrated in R10 Plus medium accordingly before added to 100µl of R10 Plus medium containing 1 x 10^6 GXR cells on Day 0 (1ml R10 Plus medium was then added after 20 hours). VRC assays were measured at least in duplicate, using gating strategies as previously described. The slope of the natural log of percent GFP-expressing cells was calculated between days 2 and day 7. The growth of virus can be summarised using slope derived from natural log of daily GFP+ percentage, as appropriate for an exponential growth curve. The final VRC of each variant is

expressed as percentage growth rate (slope) of the variants compared to that of recombinant NL4-3 strains. In addition, for each VRC assay, viral supernatants are collected on day 7 for viral RNA extraction and sequencing to confirm the population sequences VRC measured is comparable to the pre-transfection sequences.

2.4 Phylogenetics, modelling, bioinformatics and statistics

2.4.1 Phylogenetic and evolutionary analysis of HIV env in longitudinal clinical case study (Chapter 3)

Phylogenetic analysis and maximum likelihood (ML) trees were constructed using the General Time Reversible substitution model with optimised proportions of invariable sites and gamma distribution using PAUP* (version 4.0 beta; Sinauer Associates, Inc. Publishers, Sunderland, MO, USA) and PhyML (version 2.4.4) software (Guindon and Gascuel 2003).

All the time-dependant evolutionary analyses are produced by using BEAST and TRACER software packages, using methods described elsewhere (Drummond and Rambaut 2007, Drummond, et al. 2005). Briefly, the evolution of HIV *env* gene (969 nucleotides) through time was modeled using Bayesian relaxed, uncorrected lognormal, molecular clock that accommodate rate variation among branches. Given the sampling sequences, 60 millions steps of Bayesian Markov chain Monte Carlo (MCMC) algorithm was used to estimate posterior distribution of a set of evolutionary parameters, the first 10% were discarded as burn-in, and coverage was checked using TRACER. The summary of MCMC yielded highest likelihood tree

topology, with the branch lengths proportional to the absolute ages of the subtending nodes. Similarly, the Bayesian skyline plot (m=10) calculated mean estimate of effective population size ($N\varepsilon\tau$, a product of the effective population size and the generation length in days) over time (in days, since seroconversion). The skyline plot summarise 150 million of MCMC generated after discarding first 10% as burn-in. All analyses were repeated to ensure convergence. A preliminary BEAST analysis was performed that included the proviral sequence data, as these samples spanned a longer period of time than just the sequences obtained from plasma samples. The rate of evolution and effective population size were estimated from these time-structured sequences, and used as a prior in an analysis including only sequences obtained from the plasma samples.

CODEML was used to identify sites that have been subjected to positive selection using site-specific codon models (Yang and Nielsen 1998). These models allow dN/dS to vary among sites, with values significantly above 1 indicating sites that had been under positive selection. We compared the fit of models M7 (which does not allow any sites to have a dN/dS of above 1) with M8 (which fits a class of sites with a dN/dS of above 1), using a likelihood ratio rest. The empirical Bayesian approach implemented by CODEML was used to identify codons under positive selection, using a cutoff of >90%. Attempts were also made to assess overall dN/dS changes over time, using a method developed by Dr Aris Katzourakis (Gifford, et al. 2008). This method makes use of the results generated by model M1 of CODEML software, which calculate a specific dN/dS for every branch on the time measured tree. The change in selection pressure change over time was then achieved by combining dN/dS of the branches present at each time point, taken at an interval of 50 days apart.

2.4.2 Phylogenetic and molecular epidemiology of Free State Cohort, SA (Chapter 4 and 5)

The *pol* gene sequences were submitted to the BioAfrica database and REGA HIV-1 subtype tool for subtype and locus identification (de Oliveira and Cassol 2007, de Oliveira 2007). Phylogenetic analysis and ML trees were constructed using methods described in Chapter 2.4.1. The *pol* sequences of the cohort were compared to 986 well-characterised isolates from the BioAfrica database and those available on Los Alamos database (http://www.hiv.lanl.gov/) to determine subtype and lineage relationships (de Oliveira and Cassol 2007, Los Alamos National Laboratory 2010). The Slatkin and Maddison test was used to assess clustering between Free State Province sequences and other South African sequences, using MacClade (version 4) (Gifford, et al. 2007).

The *gag* and *nef* gene sequences were also analysed phylogenetically using similar methods described in Chapter 2.4.1. In addition, we also attempted using phylogenetic methods to make correction for potential founder effects present in significant HLA class I related polymorphism studies (see Chapter 2.4.5 and Chapter 5).

2.4.3 Pre-Therapy Drug Resistance Analysis

The *PR* and *RT* sequences were submitted online to the Stanford drug resistance database (version 4.3.1) to identify subtype-associated polymorphisms and genotypic drug resistance mutations for protease inhibitors (PI), nucleoside reverse transcriptase inhibitors (NRTI) and non-nucleoside reverse transcriptase inhibitors (NNRTI)

(Stanford University 2007). The Stanford mutation scoring system that indicates significance of drug resistance was used to distinguish accessory mutations (mutation score 0-9, virus remains treatment susceptible) from clinically significant mutations, including all those with scores > 9 ('potentially low level' (score 10-14), 'low level' (score 15-30), 'intermediate level' (score 30-59) and 'high level' (score >= 60) resistance) (Stanford University 2007).

These mutations were then compared with the drug resistance mutation lists of the 2008 IAS-USA and 2009 WHO transmitted drug resistance mutations (TDRM) surveillance list (Bennett, et al. 2009, Johnson, et al. 2008). Each polymorphism was studied further using the online Mutation ARV Evidence Listing program (MARVEL version 0.8), which summarises the evidence for each polymorphism according to the subtype-specific drug inducibility, viral phenotype and clinical outcome (Stanford University 2007).

2.4.4 Mathematic modelling on ARV resistance prevalence

Most of the modelling works are conducted by our collaborators, Helen Fryer and Professor Angela McLean, from Department of Zoology, University of Oxford. An ordinary differential equation (ODE) model was devised to describe the dynamics of resistance to a single drug, such as Nevirapine, with a simple genetic resistance profile. The model encapsulates four main processes: selection of drug resistance in hosts treated in the roll-out program, selection of drug resistance prior to 'official' treatment by other means such as private prescriptions, transmission of drug resistant strains to new hosts, and reversion to drug-sensitive strains in untreated hosts.

In the model the population is broken up into susceptible and infected hosts and transmission of HIV takes place under the assumption of homogenous mixing. Infected hosts pass through three stages: acute, chronic and late, and can be infected with either drug-resistant virus or drug sensitive virus. A proportion, p of hosts receive treatment from the official 'roll out' program once they enter the late stage of infection and a proportion, q of hosts have selected the drug-resistant strain prior to receiving official treatment. Selection of drug resistance in drug-sensitive hosts treated on the main scheme occurs at a rate ϕ and reversion from a drug sensitive to a drug resistant strain occurs at rate ψ in untreated hosts. There is host turnover; births occur at a constant rate and hosts survive for different length of time depending upon their state. Hosts also transmit at different rates depending upon their infection state. Using statistical parameters provided by Actuary Society of South African (ASSA) (ASSA 2007), estimated infection and demographic parameters are tailored to the Free State local population.

2.4.5 Identification of HLA class I linked polymorphisms

The phylogenetic dependency network was applied to analyse Gag protein amino acid site polymorphisms that associate significantly with host HLA class I alleles, the analysis method was previously described elsewhere (Carlson, et al. 2008). Briefly, the analysis utilise phylogenetic correction method and statistic model of evolution to evaluate associations between hosts HLA class I alleles and viral amino acid site-specific polymorphisms. The analysis aims to adjust for confounding factors including potential founders effect, linkage disequilibrium in host HLA types, co-variation in HIV codons and multiple correction in statistic testing. The significance of the associations is expressed as the q-value, which conservatively estimate the false

discovery rate for each p-value, in this study an association is considered significant when $q<0.20$.

The HLA class I associated polymorphism observed in overall cohort, and differences noted when comparing between groups with "All", "High CD4" and "Low CD4" CD4 T cell counts groups, is expected to infer the biological process of escape and reversion:

- 'Escape' describes increased polymorphisms observed at a particular amino acid site in the presence of a specific HLA class I allele. In the overall cohort study, escape can be noted at specific sites that appear to have increased variant amino acids, or reduced consensus amino acids, associated with an HLA class I allele.
- 'Reversion' describes decreased polymorphisms observed at a particular amino acid site in the absence of a specific HLA class I allele. The rate of reversion may suggest selection pressure and replication fitness cost conferred by the amino acid site. In the overall cohort study, escape can be noted at specific sites that appear to have increased consensus amino acids, or reduced variant amino acids, associated with an HLA class I allele.

The strength of associations between HLA class I alleles and viral polymorphisms in the "total" cohort, "high CD4" group or "low CD4" group was derived using a logistic regression model that corrected for phylogeny, and is reported as a $\log_2$-adjusted odds ratio.

Identified polymorphisms were mapped to known cytotoxic T cell epitopes. Epitope maps were defined from experimental data from the Durban cohort (Kiepiela, et al.

2004), and from those listed in the A-list of Los Alamos Database (http://www.hiv.lanl.gov/content/immunology/pdf/2008/optimal_ctl_article.pdf). The epitope flanking region was also included in the analysis, defined according to the five amino acids neighbouring to the defined epitope in both C and N terminal directions.

2.4.6 Statistics

The statistical analyses were carried out using Prism4 for Macintosh (GraphPad Software, version 4.0c, La Jolla, CA, USA).

In Chapter 4, Fisher's exact test (two tailed) and chi-square test for trend were used to compare pre-therapy resistance findings between groups with different CD4 T cell counts.

In Chapter 5, Fisher's Exact test is also used to evaluate 2x2 contingency tables for non-random associations between two categorical variables (such as between host HLA status and amino acid status at any given HIV gene locus). The phylogenetic dependency network described in Chapter 2.4.5 utilises *q*-value to adjust for multiple comparisons by calculation of a false discovery rate (FDR) for each p-value (Carlson, et al. 2008). This approach is considered more suitable for large numbers of multiple comparisons such as performed in genome studies since in this instance the classical bonferroni correction may be over-cautious and many true positive associations may be missed.

In Chapter 6, Student t-tests (two tailed) and Mann-Whitney tests (two tailed) were used to compare various clinical parameters and VRC measurements.

2.5 Immunology

2.5.1 The CTL restricted epitopes and ELISPOT assays

The ELISPOT assay results were utilised to examine HIV-1 specific IFN-γ responses in the clinical case study (Chapter 2), and our collaborator's cohort in Durban (Chapter 6). All the assay works were conducted previously and have been reported (Duda, et al. 2009, Frater, et al. 2007, Leslie, et al. 2004).

Identified polymorphisms were mapped to known cytotoxic T cell epitopes. Epitope maps were defined from experimental data from the Durban cohort and from those listed in the A-list of Los Alamos Database (Leslie, et al. 2004, Los Alamos National Laboratory 2010). The epitope flanking region was also included in the analysis, defined according to the 5 amino acids neighbouring to the defined epitope in both C and N terminal directions.

2.5.2 Antibody neutralisation study

This work is done by our collaborators, Dr David Bonsall and Dr Emma Thomson, at St Mary's Hospital, London, UK. The experimental data further suggest the role played by neutralising anybody in the chronic stage of HIV. Briefly, full-length virus *envelope (env)* was PCR amplified from cDNA obtained by reverse transcription (RT) of plasma-derived viral RNA collected 336 d and 1042 d post seroconversion for

construction of pre-Rituximab Env-pseudotyped virus. Site directed mutagenesis was conducted in some laboratory strains to generate mutants harbouring specific site mutations of interest (specific mutations selected in patients over different time points that may alter recognition mediated by neutralising antibodies). Neutralization of pseudotyped virus was determined using previously described methods, in triplicates by different investigators (Wei, et al. 2003). The virus was incubated for 1 hour with a 3-fold dilution series heat inactivated serum, or broadly specific neutralizing antibodies (b12 or 2G12). After 48 hours, fixed and stained cells were manually counted and neutralization measured as the percentage reduction of focus forming units (FFUs) relative to non-neutralized virus-controls.

Lastly, competitive binding enzyme-linked immunosorbent assay (ELISA) was done to determine antibody specificity. Here, the high-binding microtitre plates (Greiner) were coated overnight with recombinant Bal-strain gp120 and blocked with 2% milk powder to limit non-specific antibody binding. Plates were washed and incubated with 50% saturating titres of biotinylated-NAbs (b12, 2G12, 447-52D). An additional plate was pre-incubated with the CD4-peptide mimic M33 prior to incubation with biotinylated-NAb 4.8d25. After a further hours incubation and wash step, each plate received 3-fold dilution series of patient sera. Bound NAb was detected with streptavidin horseradish peroxidase and TMB chromogen-substrate. Serum mediated inhibition of NAb-binding was determined as the percentage reduction in light absorbance (A_{450nm}) compared to wells that received dilution buffer (PBS 0.05% tween-20, 1% BSA) instead of serum.

CHAPTER 3:

B cell depletion in vivo reveals a role for antibody control of HIV

This work is based on a manuscript that is published in Nature Communication (Huang, et al. 2010).

3.1 Background

HIV-1 is readily able to evade adaptive immune responses and establish persistent infection, yet there is good evidence that CD8+ T cell responses play a role in viral suppression (Goulder and Watkins 2008). For example, there are strong HLA Class I associations with clinical outcome, and in the SIV model, macaques in whom CD8+ T cells are depleted show significant rises in their viral load. These findings have led to a major initiative to develop T cell based vaccines, and much work to define the correlates of protection based on assays of T cell function. However, since a leading HIV vaccine candidate failed to induce protective T cell responses (Buchbinder, et al. 2008) and recent trial data suggest partial protection from a vaccine including a gp120 component (Rerks-Ngarm, et al. 2009), the incentive to understand the role of neutralizing antibodies (NAbs) and of B cells in HIV has increased considerably (Moir and Fauci 2009, Sattentau 2008, Scheid, et al. 2009).

B cells play a multifaceted role in humoral and cellular responses to HIV, yet their impact in controlling an established infection is unclear, partly because of the

complex and resilient nature of the antigenic target (gp120). NAbs emerge in acute HIV infection, but do not lead to sterilising immunity as in many other infections. One reason for this is the unstable genome of HIV, which facilitates the emergence of escape mutants. The NAb response also evolves through somatic recombination and affinity maturation of proliferating B cells, but is often only partially effective against concurrently circulating viruses, while retaining potent activity against historic strains (Richman, et al. 2003). In contrast to the over-representation of specific HLA mediated CTL selections amongst the elite HIV controllers, little is known regarding the antibody correlates to the disease protection and control (Bailey, et al. 2006, Pereyra, et al. 2008, Saksena, et al. 2007). Quantitatively, the HIV elite controllers have similar (or even lower) NAb titre relative to the progressors, although this may be related to the level of circulating antigen (pVL) (Bailey, et al. 2006, Pereyra, et al. 2008).

However, it is important to note that several monoclonal antibodies (MAb) have been identified which both cross recognise and potently neutralize a variety of strains (Corti, et al. 2010, Gonzalez, et al. 2010, Koff and Berkley 2010, Ng, et al. 2010, Walker, et al. 2009, Walker, et al. 2010). Unlike CTL studies, the passive introduction of neutralising monoclonal antibodies (NmAb) confers both preventive and controlling effect over HIV transmission and infection (Baba, et al. 2000, Hessell, et al. 2010, Ng, et al. 2010). The targets for these NmAb are well defined, which include: the CD4 binding site and co-receptor binding sites (b12, VRC01, HJ16), the membrane proximal external region of Env gp41 fusion subunit (2F5, 4E10, Z13el), the glycan shield (2G12), V1-V2 loop and V3 loop (PG9 and PG16) (Hessell, et al. 2010, Koff and Berkley 2010, Ng, et al. 2010, Scheid, et al. 2009, Walker, et al. 2009,

Zhou, et al. 2010). Although potent monoclonal responses arise rarely in natural infection, the proportion of gp120-specific antibodies capable of neutralization may be greater than once thought (Scheid, et al. 2009), and currently the functional data to causally link such antibodies to the control of viremia is lacking.

Rituximab (Rituxan™, MabThera™, Roche) is a monoclonal antibody targeting the CD20 antigen expressed on most B cells, with the exception of terminally differentiated plasma cells. It has become widey used for treatment of lymphoma and also of antibody-mediated autoimmune diseases (Mease 2008). There are no published studies on the impact of Rituximab alone on HIV-viremia in humans. Studies of acute SIV_{mac} infection in macaques treated with Rituximab indicate a role for NAb responses in containing post-acute viral load (Miller, et al. 2007, Schmitz, et al. 2003). In contrast, a recent study using African Green monkeys did not reveal any impact from Rituximab treatment during acute SIV_{agm} infection (Gaufin, et al. 2009). Even though the difference in findings between animal models may be related to the differential pathogenesis process (the macaque infected with SIV_{mac} tend to progress rapidly in contrast to SIV_{agm} infected African Green monkeys) (Hirsch 2004), the case study data from human studies are important to provide proof of principle and resolve this issue.

In human cases, there lack definitive case progression reports from patients suffering from primary antibody deficiency disorders prior to HIV infection. Currently, only one case recorded longitudinal HIV progression in a patient suffering background combined variable immunodeficiency disorder (CVID) (Jolles, et al. 2001). The patient progressed to AIDS defining illnesses rapidly (non-Hodgkin's lymphoma and

CD4 T cell < 350 cells/μl) 21 months post-infection. Interestingly, the antibody levels of IgG and IgM appeared to have improved since HIV infection. In another less well defined clinical case reported recently, the case subject presented atypically following HIV infection: the patient suffered AIDS defining illness (*pneumocystis jirovecii* pneumonia) during the prolonged seroconversion period (Padiglione, et al. 2010). In this case study, the genetic investigation revealed that the patient carried disease susceptible genes for antibody deficiency disorder (despite been asymptomatic for CVID in clinical history), potentially explain the delayed seroconversion and unusual pathogenesis. There currently exist no studies reporting on the acquired inhibition of NAb production in human cases.

Here, we report a unique case in which Rituximab monotherapy was used in a patient with stable viremia, in the absence of antiretroviral therapy, allowing an analysis of the antiviral role of B cells in a direct fashion. We show reversible loss of HIV-1 control following Rituximab monotherapy, associated with reciprocal changes in NAbs.

3.2 Hypothesis

Durable B cell function may be more important in the control of HIV replication in chronic infection than has been suspected. Following the depletion of B lymphocytes by rituximab, the emerging viral population could potentially arise from

(i) Previous viral strains suppressed by NAb that can no longer be contained due to lost of NAb,

(ii) Contemporarily circulating viral population that are no longer suppressed, or

(iii) Newly evolved viral strains unable to be controlled by loss of B cells NAb production.

3.3 Aims

i) To describe viral dynamics and sequence evolution in the *env* gene, in the context of the patient's clinical progress.

ii) To functionally assess the evolving *env* gene following changes in NAb and B cells selection pressure.

3.4 Results

3.4.1 Clinical course of Study Subject

The patient, a Caucasian male, was diagnosed with HIV seroconversion illness aged 58. Three years earlier he was diagnosed with low grade lymphoplasmacytoid lymphoma for which he did not receive treatment. His only other medical history of note was an acute hepatitis B infection some thirty years earlier. Following HIV-1 seroconversion he was recruited into a prospective, non-randomised observational study of early treatment and received three months of highly active antiretroviral therapy (HAART)[9]. Thirty months later he developed a rising paraproteinaemia, attributed to his lymphoma and received thalidomide treatment without clinical effect.

He initiated Rituximab therapy 1075 days after HIV seroconversion and received four weekly doses (375mg/m^2). The patient's CD4 T cell counts was 530 cells/µl (24%) and VL 15,000 copies/ml immediately prior to administration of rituximab. He felt unwell 20 weeks after the first dose of Rituximab with malaise and fever associated with a marked rise in HIV pVL (increase of 1.7$\log_{10}$, Figure 3.1), non-productive cough, abdominal discomfort and weight loss of 7 kg. Additionally, he developed a biochemical hepatitis (attributed to reactivation of hepatitis B) which resolved with only supportive treatment. The VL subsequently returned to baseline levels (283 copies/ml) without any ART intervention. His cough was attributed to an interstitial pneumonitis, which resolved on subsequent imaging. Antiretrovirals were initiated 1392 days after seroconversion.

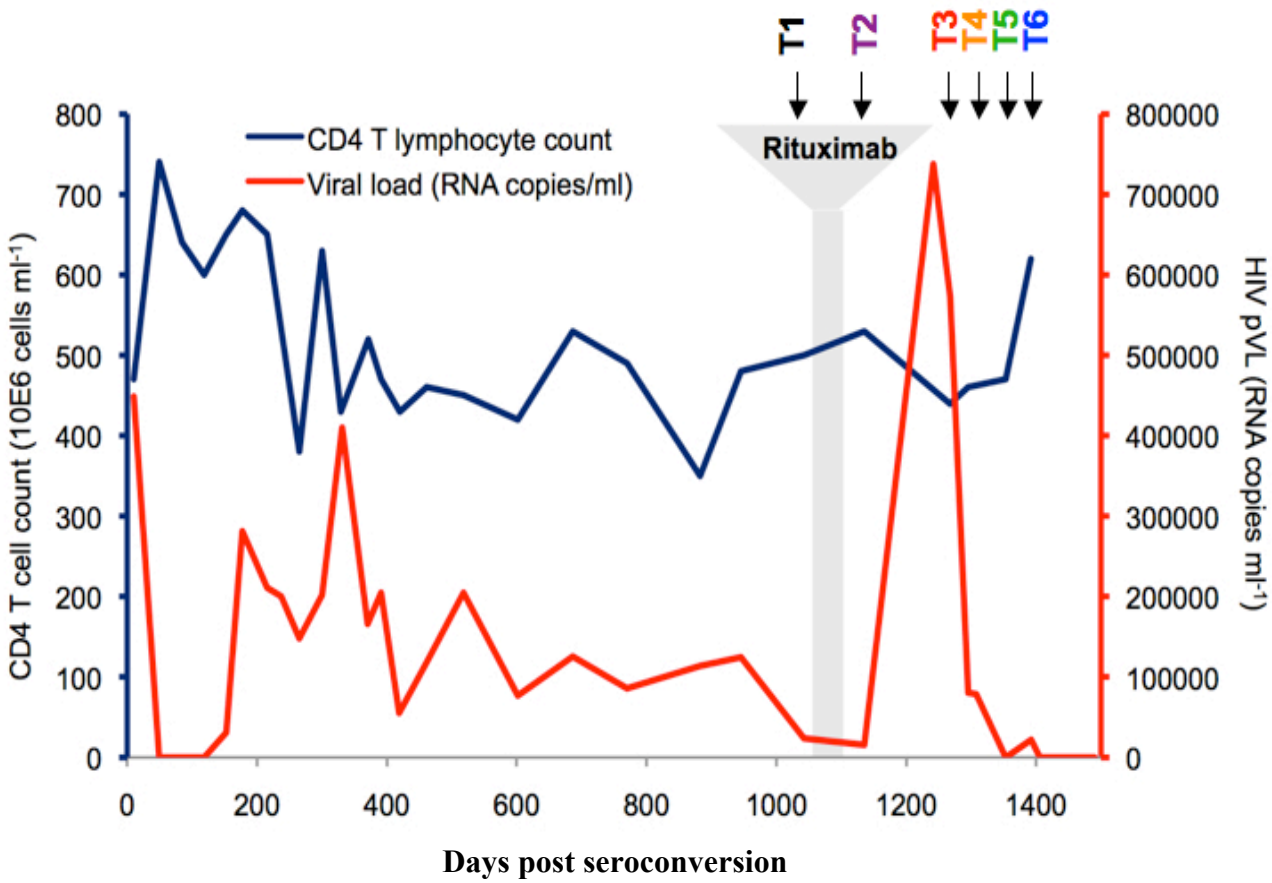


Figure 3.1: Clinical course of HIV in relation to rituximab therapy

The plasma viral load (red; right y axis) and CD4 T cell counts (blue; left y axis) of the individual studied are displayed over time (days since seroconversion). The brief dip in viral load over the first 100 days represents the period of initial ART as part of the SCART trial. Subsequently viral load was contained over approximately 1000 days without ART. The time of initiation of Rituximab therapy is indicated. Time points 1-6 (T1, T2, T3, T4, T5 and T6) are indicated and correspond to the time points at which plasma virus was sequenced and displayed in Figure 3.2.

3.4.2 Analysis of sequence evolution in HIV-1 envelope over time

Plasma viral samples are available from various important clinical event time points, labelled in days post-seroconversion (Figure 3.1): 1 month before rituximab therapy (d1042, T1), 6 weeks after rituximab (d1135, T2), immediately following viral load peak (d1270, T3), remission (d1295, T4), new baseline (d1353, T5) and after commencement of highly active anti-retrovirus (HAART, d1392, T6). The patient's previous commitment to the SCART study offers additional proviral samples from peripheral blood mononuclear cells (PBMC) from seroconversion up to day 1042. The V1 to C4 region of the *env* gene (969 base pairs) was successfully sequenced from the patient at 12 clones per time point. The virus is confirmed to be of HIV-1 subtype B in terms of molecular epidemiology classification.

To evaluate sequence evolution over this period *env* products were generated by RT-PCR from plasma viral RNA, prior to and following Rituximab therapy (d 1042-1392; Figure 3.2A). We identified clear sequence diversification as the virus peaked during Rituximab treatment (mean Hamming distance = 0.88 vs 12.86; $p<0.0001$, Figure 3.2B). This diversification manifested as an increase in effective population size as measured using BEAST (5-fold change; Figure 3-2B) (Drummond and Rambaut 2007). To assess whether these diverse strains represented re-emergence of archived provirus - potentially as a result of profound loss of neutralising antibody - we used an identical sequencing strategy to analyse stored PBMC samples available from the time of seroconversion up to day 1042; these sequences were placed on a time-structured tree (Figure 3.5). Viruses found post-Rituximab did not represent re-emergence of early archived strains; those strains present as pVL peaked were most

closely related to strains present in plasma at day 1042, with a subset related to those archived in PBMC at day 1042.

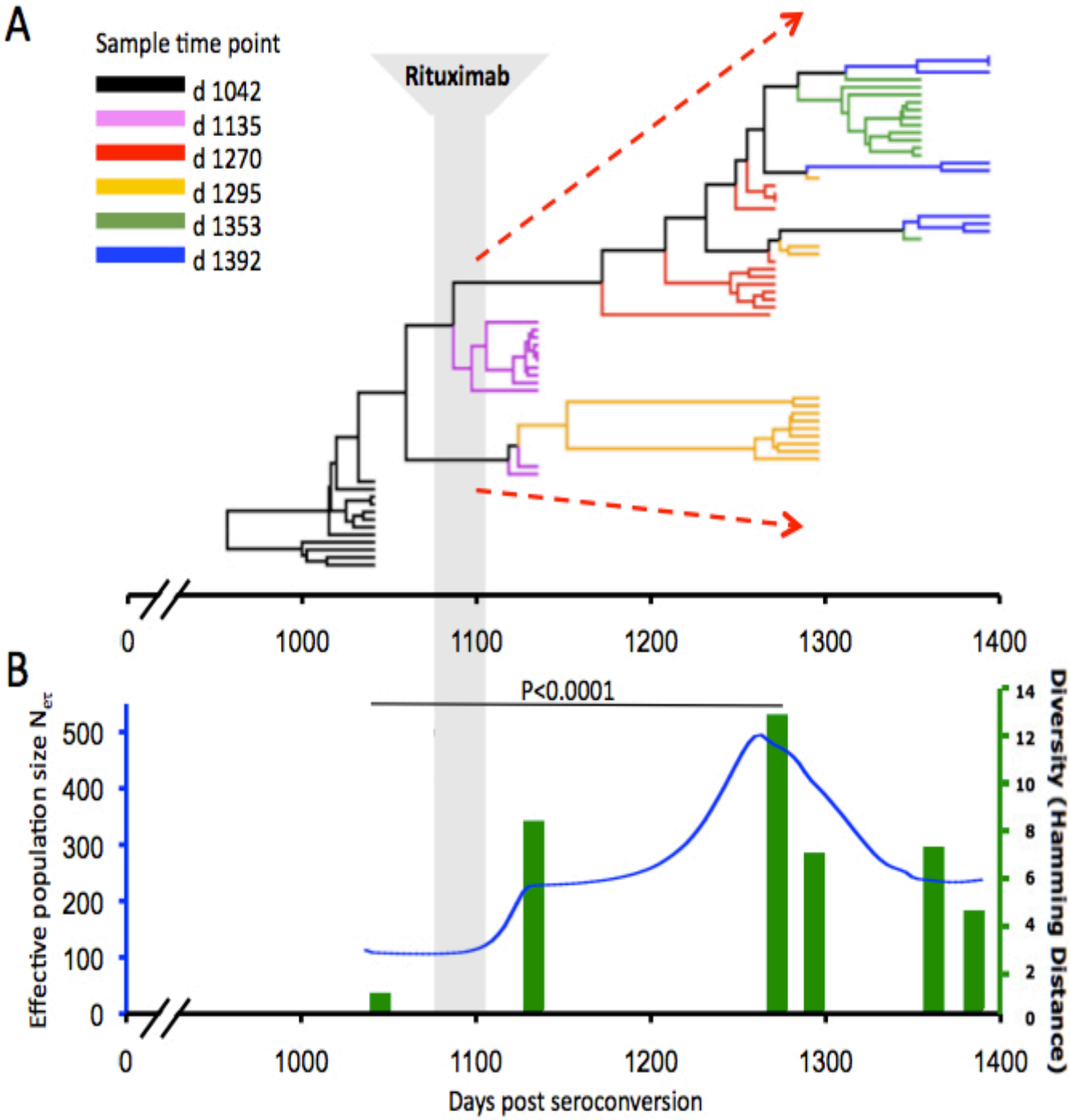


Figure 3.2: Diversification of HIV *env* sequences after rituximab therapy

a) Phylogenetic relationships of sequences over time: A time-structured tree, showing divergence time estimates among intrahost plasma HIV-1 clonal sequences before and after Rituximab therapy (shaded grey), was based on a Bayesian relaxed molecular clock applied to 969 nucleotides (Yang and Nielsen 1998). The highest likelihood tree

topology was calculated using BEAST software packages, with branch lengths proportional to the absolute ages (time estimate) of the subtending nodes (labelled by X-axis, summarizing 60 million Markov chain Monte Carlo (MCMC) simulations, generated after discarding 6 million as burn-in). All the parameters in the MCMC started from random values, except for substitution rates and population size, where the priors were imported from the combined proviral and plasma viral sequence data of the same host calculated by BEAST software. The red dashed arrows highlight the sequence diversification since Rituximab introduction.

b) Diversity over time (skyline plot): The left y-axis showed the mean estimate of effective population size ($N\tau$, a produce of the effective population size and the generation length in days, in blue) over time (in days, since seroconversion) of Bayesian skyline plot (m=10). The right y-axis depicts the Hamming distance measurement of mean viral Env genetic diversity (in green). The x-axis (days post-seroconversion) is the same as for Figure 3.2A. The plot is derived from intrahost plasma HIV clonal sequences measured before and after Rituximab therapy, based on a Bayesian relaxed molecular clock applied to 969 nucleotides. The plot is produced using BEAST and TRACER software packages (summarising 150 million of MCMC generated after discarding 15 million generations as burn-in). All the parameter MCMC started from random values, except for substitution rates and population size, where the priors were imported from the combined proviral and plasma viral sequence data of the same host calculated by BEAST software.

3.4.3 Impact of B cell depletion on NAb titres:

(The results in Chapter 3.4.3 derive from the work completed by our collaborators, see acknowledgement for more details)

Following Rituximab therapy the subject's HIV-1 pVL, previously well controlled, rose by over 1.7 log_{10}, peaking at 737,400 copies/ml four months after the final dose of Rituximab (Figure 3.1 and figure 3.3). The pVL subsequently returned to baseline levels in the absence of ART. We therefore analysed the impact of Rituximab therapy on NAb titres in relation to viral dynamics.

Rituximab therapy led to a 50% reduction in total IgG levels which reached their lowest point at the peak of virus load (10.8 mg/L, Figure 3.8C) and returned to pre-Rituximab levels thereafter. To evaluate NAb activity throughout this period we generated pseudoviruses derived from autologous *env* sequences. Serum obtained prior to the first dose of Rituximab potently neutralized virus pseudotyped with autologous envelope derived from plasma acquired 2 years previously (d 336 post serconversion). The same serum neutralized contemporaneous autologous Env-pseudotypes, albeit to a lesser extent, with 1:60 serum dilutions effectively blocking more than 90% infection. Heterologous NAb activity was also observed against a standardized panel of Tier 2/3 clade-B reference strains (NIH, AIDS) (Table 3.1). Following Rituximab therapy, we observed a marked decrease in neutralizing activity characterised by a drop of more than 3-fold in IC50 serum titres against d 336 Env-pseudotypes, which recovered over a four month period. Strikingly, these changes coincided with the increase in pVL. Recovery of NAb responses correlated with control of viremia (Figure 3.3).

To further define the site of binding of the patient's NAbs, we performed a direct competition experiment using well-defined monoclonal antibodies and recombinant clade B envelope as the antigenic target (Figure 3.4). This experiment showed strong binding-competition between patient sera and antibodies b12, 4.8d (which targets the co-receptor binding site) and 447-52D (which targets the V3 loop). There was no competition seen against 2G12. The experiment also revealed a clear decline in competition for binding in samples taken during the peak of viremia, which recovered as pVL declined, in good accordance with the data obtained from the neutralization experiments, but using an independent system of evaluation (Figure 3.4B).

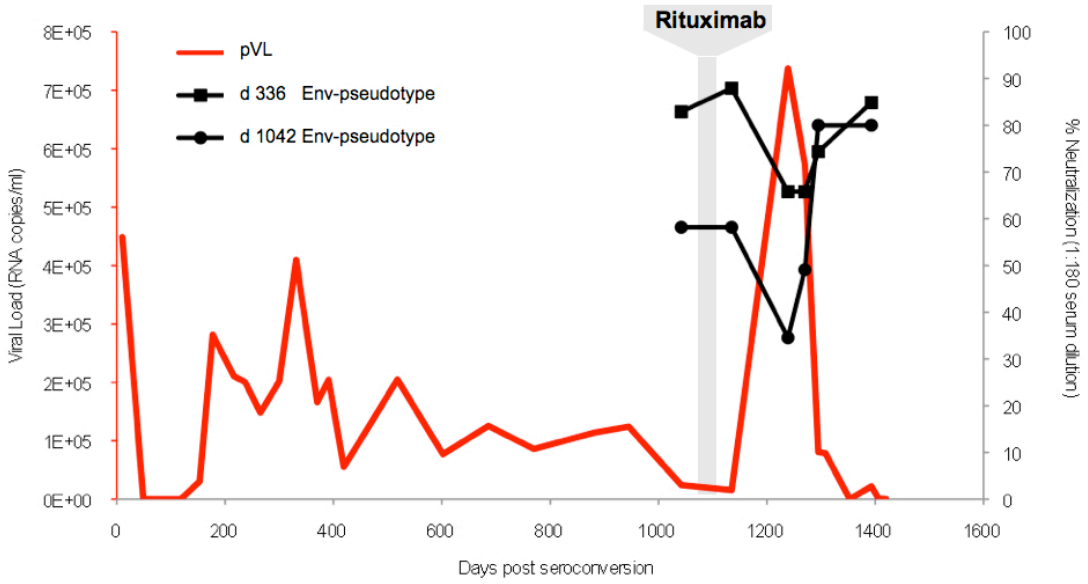


Figure 3.3: Autologous neutralising antibody titres over time

Pseudotyped virus particles were constructed using full-length envelope clones that represent majority quasispecies at 336 or 1042 days post-seroconversion. Plasma VL (red line; left axis) is shown over time (x-axis) alongside the neutralising effects (%; right y-axis) of patient sera against autologous viruses constructed with cloned *env* genes derived from 336 d (■) and 1042 d (●) post-seroconversion. Serum was sampled sequentially before and after Rituximab therapy (1075-1106 d, shaded grey), and assayed at concentrations ranging from 1:20 – 1:4370; the neutralising effects of sera diluted 1:180 are shown here.

Table 3.1: IC50 neutralisation titres against autologous and heterogous virus

Days post sero-conversion	Autologous virus			Heterologous virus: Clade B reference strains (NIAID, NIH)									VSV-G
	d336	d1042	Mutant** 339E 363R	HXB2	6535	QH0692	PVO	TRO	AC10.0	pREJO 4541	pTRJO 4551	pRHPA 4259	
1042	>775	150	*n/t*	850	220	220	90	120	90	140	50	50	<20
1135	>775	225	300	1000	*n/t*	*n/t*	*n/t*	*n/t*	*n/t*	*n/t*	*n/t*	*n/t*	*n/t*
1239	220	<120*	*n/t*	875	*n/t*	*n/t*	*n/t*	*n/t*	*n/t*	*n/t*	*n/t*	*n/t*	20
1295	250	150	*n/t*	325	*n/t*	*n/t*	*n/t*	*n/t*	*n/t*	*n/t*	*n/t*	*n/t*	20
1308	360	375	*n/t*	1000	140	140	90	200	70	140	140	200	<20
1353	600	425	600	3950	300	170	90	130	160	220	140	120	*n/t*
negative serum	<20	<20	20***	<20	20***	<20	<20	<20	<20	20***	20***	<20	*n/t*

The half maximal inhibitory concentrations of sera (IC50, reciprocal dilutions) are given for virus pseudotyped with representative Env-clones derived from d336 and d1042 plasma. (*) 1:120 was the highest concentration of d 1239 serum that could be tested against d 1042 virus due to limited serum availability. (**)Mutant 339E/363R was constructed to represent envelope changes identified following Rituximab therapy. Patient sera potently neutralized the HXB2-Env pseudotype and exhibited moderate neutralizing potency and breadth against heterologous pseudoviruses constructed from a standardized reference panel of heterogeneous envelopes (NIAID, NIH). Low-level neutralizing effects of pooled HIV-negative sera (***) were attributed to cellular toxicity which was not observed at concentrations below 1:20. Furthermore, serum concentrations below 1:20 did not inhibit VSV-G pseudotyped virus indicating that neutralizing responses were predominantly HIV-Env specific. "*n/t"* denotes virus-serum combinations that were not tested due to serum availability.

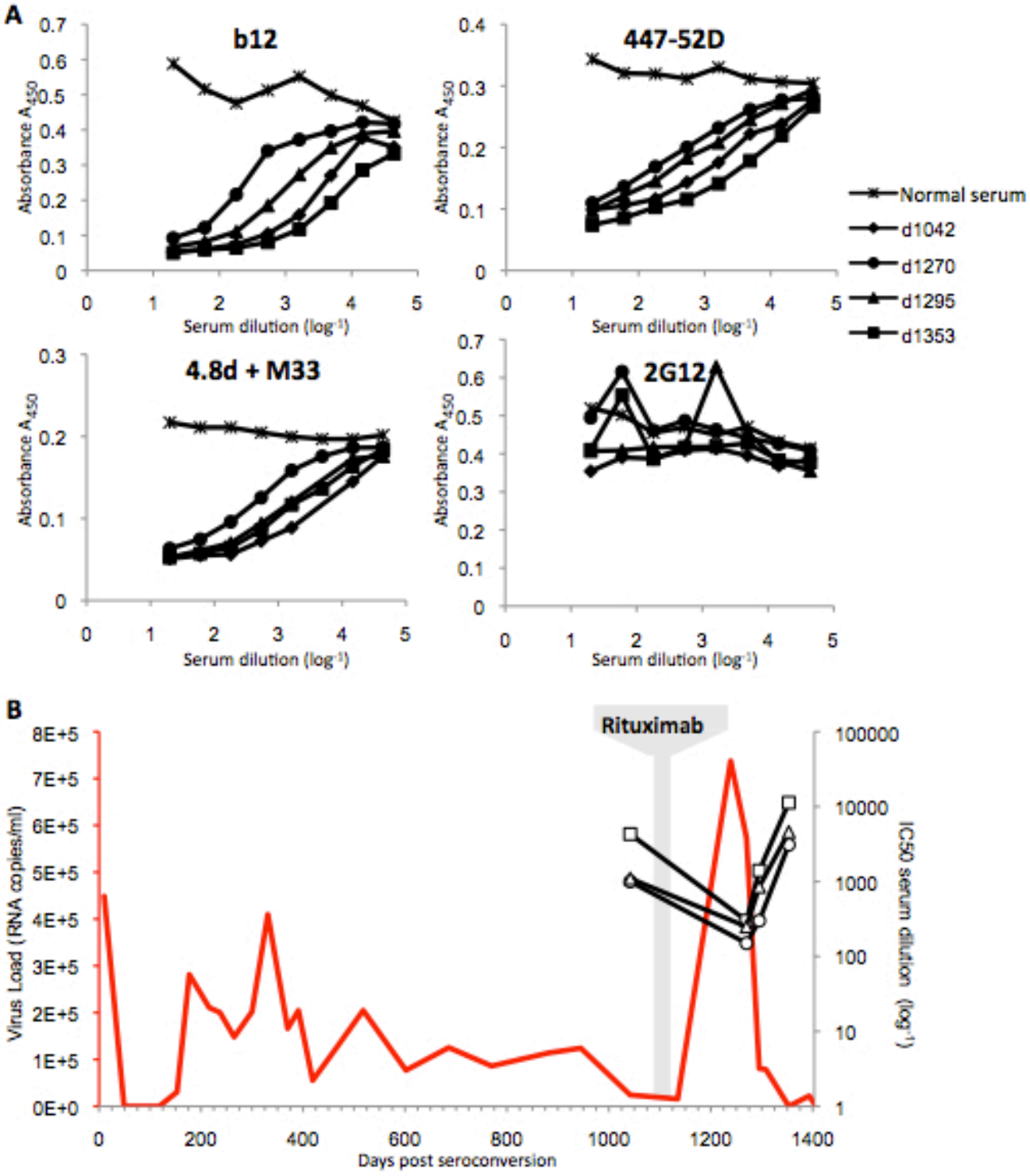


Figure 3.4: Characterisation of patient antibodies by competition-ELISA

Panel A: Serial dilutions of patient serum (x-axes) collected on 1042 d (◆), 1270 d (●), 1295 d (▲) and 1353 d (■) post-seroconversion reduced gp120 binding with NAbs: b12, 447-52D, and 4.8d, thereby reducing colourometric change and light

absorption (A_{450nm}, y-axes). Pooled serum collected from HIV-uninfected subjects (✳) did not compete with any NAb for gp120-binding. All sera failed to compete with 2G12 for gp120-binding. The results shown in the line plot is the mean result of the triplicate experiments.

Panel B: shows the reciprocal dilutions of sera, sampled before and after Rituximab therapy, that reduce gp120-NAb binding by 50% (IC50, right y-axis). One hundred percent binding was taken as that seen in microtitre-wells containing either IgGb12 (□), 4.8d (△) or 447D (○) in the absence of patient sera. Corresponding VLs are shown by the red line (left y-axis).

3.4.4 Immune selection and neutralisation susceptibility of HIV-1 Env over time

Since disturbance of the B cell compartment may impact specifically on sites under prior antibody selection we next analysed immune selection. The apparent *in vivo* impact of the NAbs in this individual prompted a search for the targets of binding. To identify potential NAb targets we first analysed immune selection using CODEML. We observed two sites under selection (Figure 3.5 and 3.6) - position 339 (dN/dS=7.6; $p<0.0001$) and position 363 (dN/dS=6.977, $p=0.03$), which have been previously implicated in the binding of CD4bs specific NAb b12 and the carbohydrate specific 2G12 (Scanlan, et al. 2002).

To examine the effect of sequence changes on neutralization susceptibility, pseudoviruses mimicking Q363R and R339N/E were constructed by site-directed mutagenesis. These mutations were shown to modify neutralization by antibody b12, which targets the CD4 binding site (CD4bs) (Figure 3.7) (Scheid, et al. 2009). In addition, 339N completes an NDT glycosylated motif that induced neutralization-susceptibility to the carbohydrate specific antibody 2G12. The emergence of 339N and 363R virus occurred prior to Rituximab therapy. Following Rituximab we saw rapid re-emergence of b12-susceptible quasispecies possessing 363Q (ie. the original residue), which switched to 363R as NAbs returned. This was followed by the acquisition of a 339E mutation which further reduced susceptibility to b12, and also reduced IC50 titres of d1042 serum by half (Table 3.1). Importantly, declining NAb titres and peak pVL were temporally associated with these key mutations.

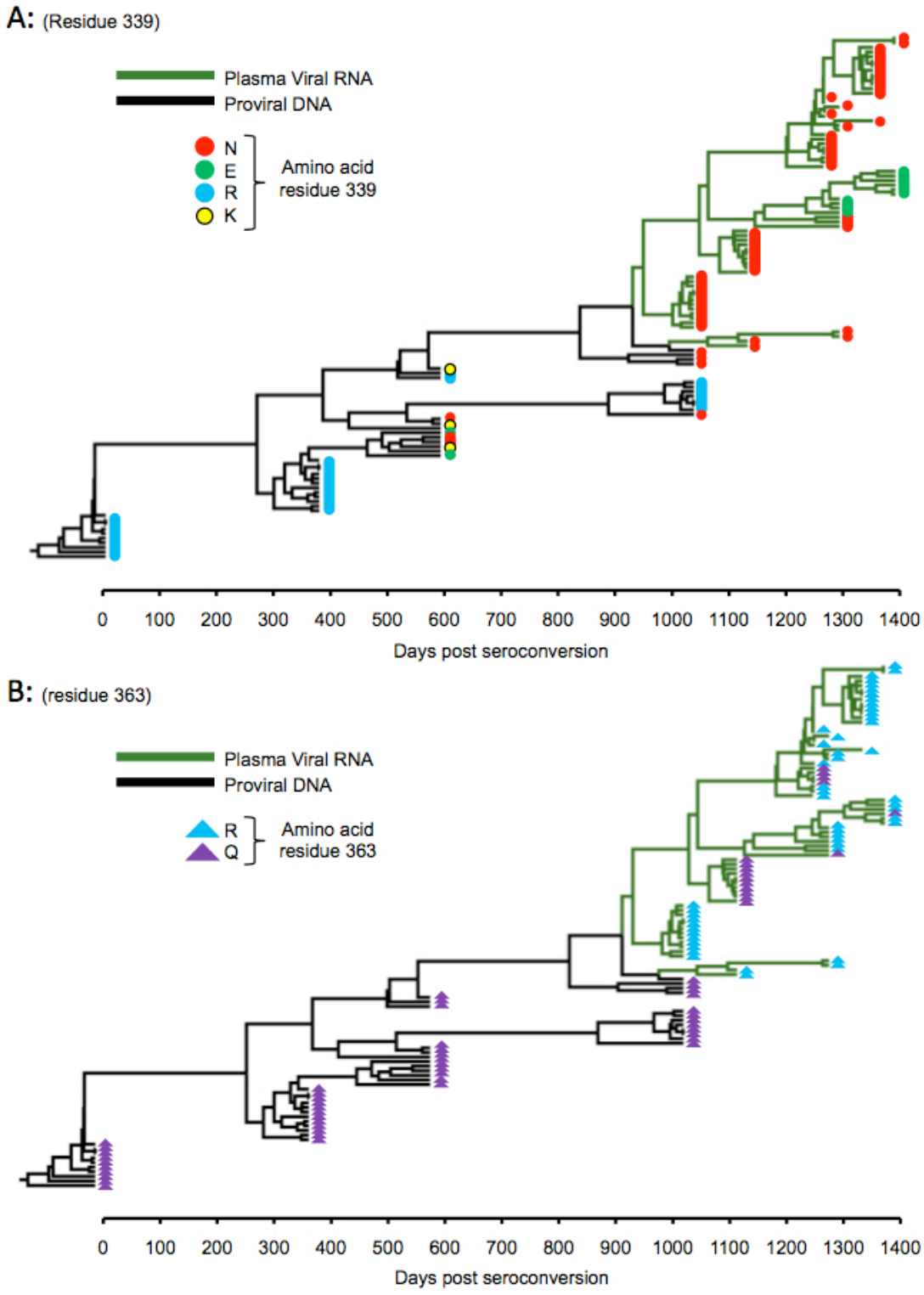


Figure 3.5: Ancestral analysis of emerging strains

Panel A: The amino acid status change in viral-Env protein (position 339) over time. A time-structured highest likelihood tree, showing reconstruction of the probable

ancestral amino acid changes at Env position 339 of the plasma viral samples and proviral samples over time. The amino acid status is reconstructed using parsimony described in Figure 3.2.

Panel B: The amino acid status change in viral Env protein (position 363) over time. Data were calculated and presented as for panel A.

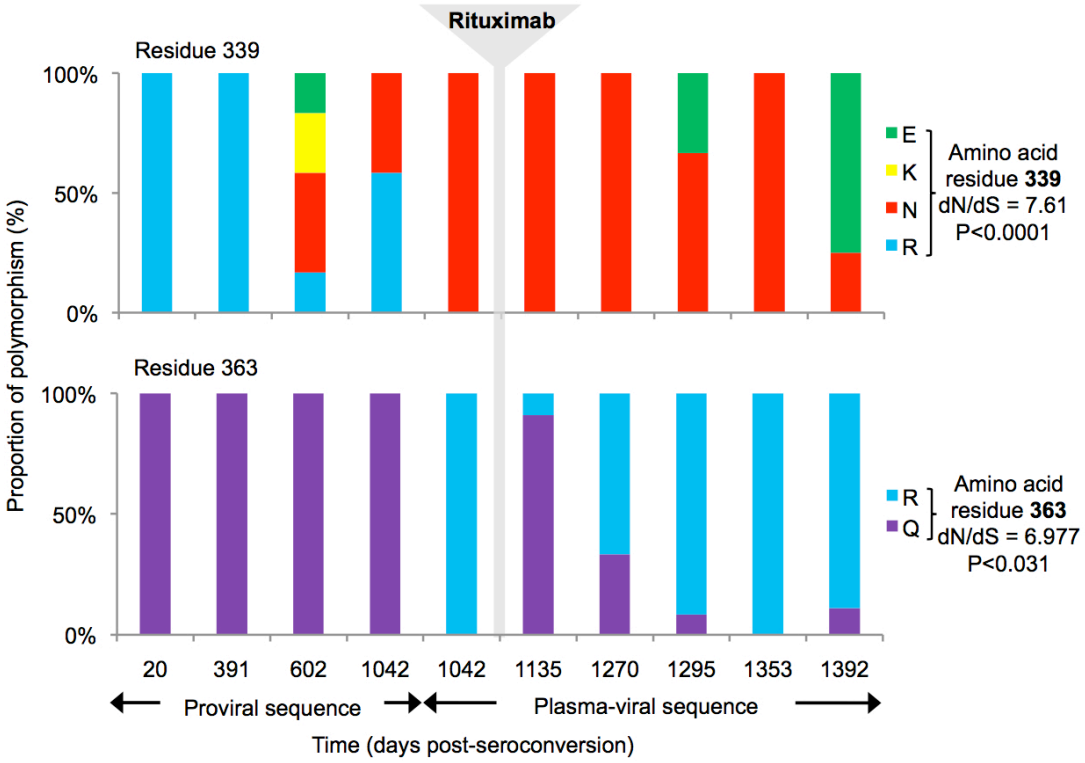


Figure 3.6: Tracking of selected enveloped mutants

This figure displays the relative frequencies of polymorphisms at sites 339 and 363 from historical samples and plasma samples obtained at the time of Rituximab therapy. Each bar shows the frequency of mutants at these positions.

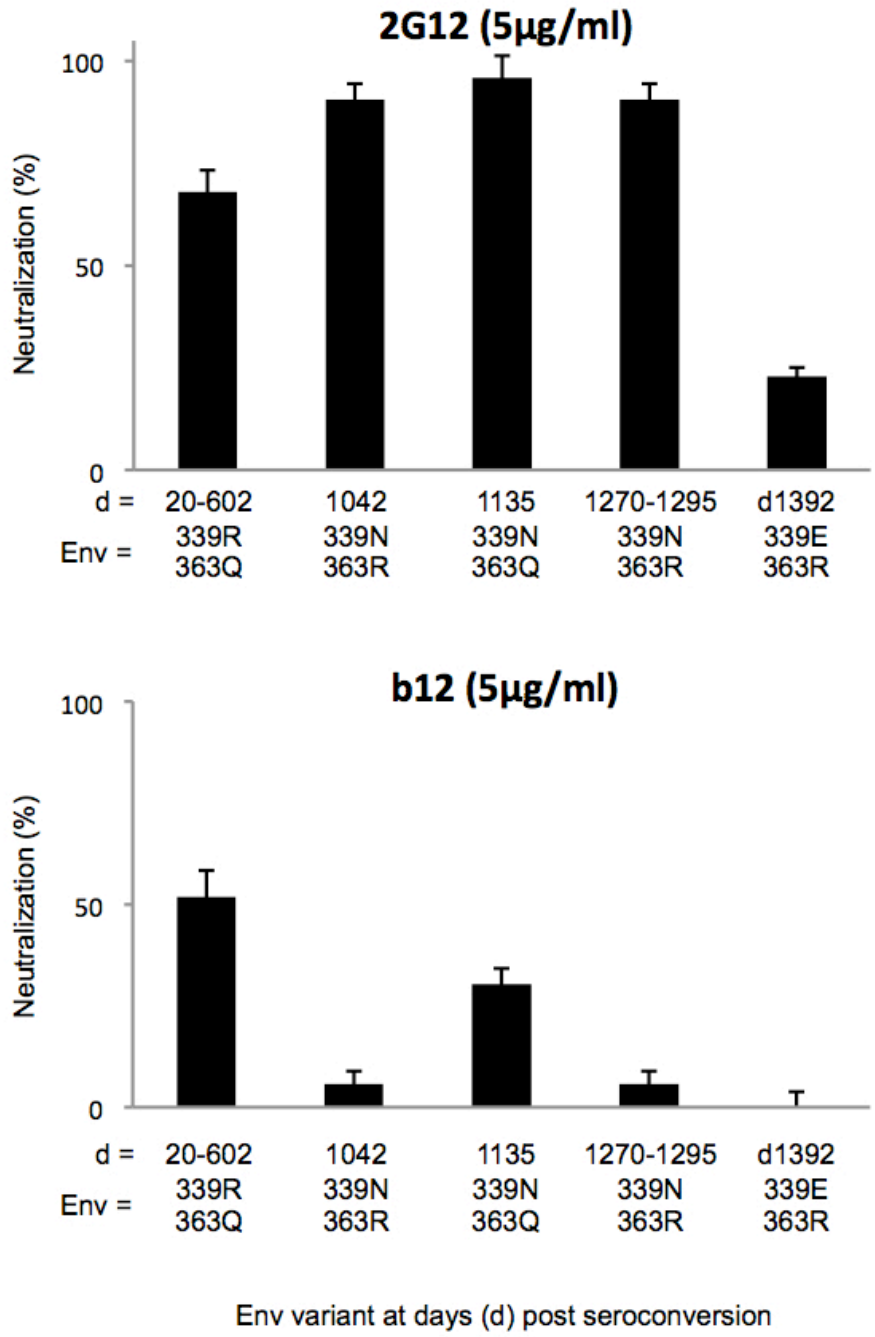


Figure 3.7: Impact of selected mutations on neutralisation by NAbs

Mutant psudoviruses were constructed by site directed mutagenesis of 336 d envelope and used in neutralisation assays to assess relative sensitivities to NAbs 2G12 (top) and b12 (bottom). Viruses are listed on the x-axes in the order they occurred *in vivo* and the percentage of each pseudovirus neutralised by 5µg/ml of NAb is given by the

y-axes. Psudoviruses mimicking the predominant variant after Rituximab therapy and immediately before peak viraemia (d 1135; 339N 363Q) were transiently more susceptible to b12 than variants present before Rituximab therapy (d1042; 339N 363R). Pseudovirus carrying an N-linked glycan at position 339 were more susceptible to neutralization by the carbohydrated-specific NAb 2G12. The error bars depict standard deviation from triplicate measurements.

3.4.5 Analysis of other factors potentially linked to the impact of Rituximab

We set out to investigate other possible causes for the observations made, particularly the possibility of an indirect impact on cellular immunity (Liossis and Sfikakis 2008). The patient was identified to possess HLA class I tissue type HLA-A*0101, HLA-A*1101, HLA-B*0801 (homozygous), and HLA-Cw*0701 (homozygous). The longitudinal sequencing of HIV-I Gag, Pol and Nef genes were analysed in the context of HLA-mediated epitope restriction. Mutations were detected in four epitopes in Gag and Nef at time points unrelated to rituximab therapy (Table 3.2). Peptide matching was done using optimised autologous peptide (from baseline sample), and wild type (from population consensus). Epitopes were used to assess autologous CTL responses by IFN γ -ELISPOT analyses. Peripheral blood mononuclear cells were not available post-rituximab, therefore cells collected after 1 year (d 391) and 3 years (d 1042) post seroconversion were used. Responses against four peptides were detected 1 year post-seroconversion, and two of these had been lost by year three, shortly before rituximab therapy. Of the two remaining responses, neither of the corresponding epitopes exhibited sequence changes during the period when control of VL was transiently lost and subsequently regained (d 1042 to d 1393 post-seroconversion). Although some mutations in pre-Rituximab sequences were seen, the majority of the epitopes remained stable. Importantly, control over pVL was regained without further escape/reversion at these sites.

Secondly, immune activation through recrudescence of other infectious agents may play a role in this case. Concurrent with the pVL rise, the patient also developed elevated liver transaminases, associated with reactivation of previously controlled hepatitis B (HBV) infection (Garcia-Rodriguez, et al. 2008). However, control of

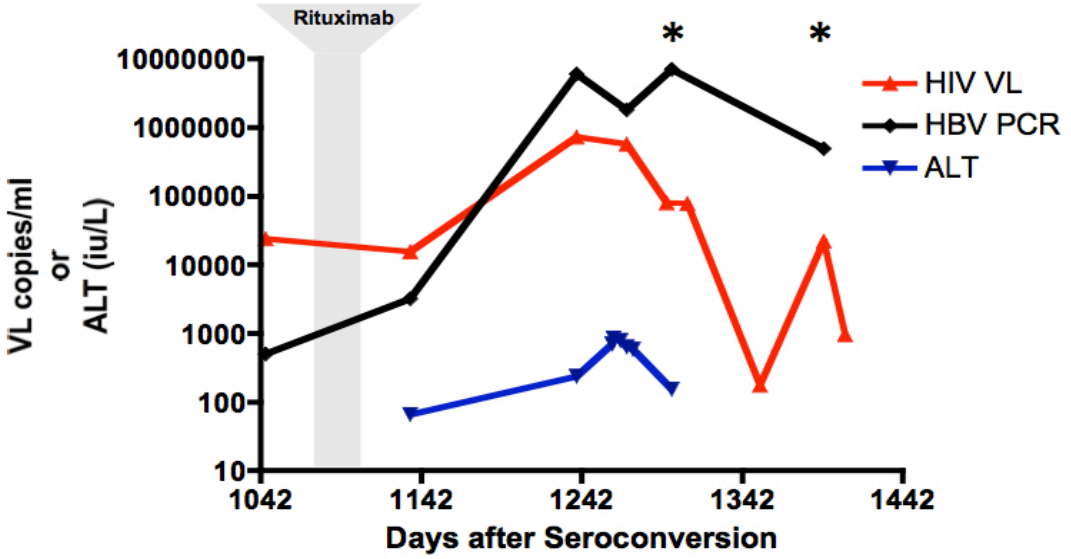
HIV was observed at a time when HBV load was still high, suggesting that HBV replication was not driving HIV replication (Figure 3.8) in this patient. The patient had otherwise no other known concurrent viral infection.

Table 3.2: Analysis of CD8+ T cell responses and epitope evolution over time

Epitope	Autologous sequence	Response pre-rituximab Year 1	Response pre-rituximab Year 3	Mutation pre-rituximab	Mutation post-rituximab
A11 gag TK8	TLYCVHQK	1,200	–	K8N	Q7A
A11 pol AK11	ACQGVGGPGHK	300	–	–	G9A
A11 nef QK10	QVPLRPMTYK	–	–	–	–
A11 pol AK9	AIFQCSMTK	–	–	–	–
B8 gag p17 GK9	GGSKKYQLK	–	–	K3R/K7Q	K3R/K7Q
B8 gag p17 EV9	ELKSLYNTV	160	200	–	–
B8 gag p24 DI8	DIYKRWII	640	600	–	–
B8 gag p24 DL9	DCKTILKAL	–	–	–	–
B8 nef WM9	WPKVRERM	–	–	K3G/E6D	K3R
B8 nef FL8	FLKEKGGL	–	–	K5Q	K5Q
B57 gag TW10*	TSNLQEQIGW*	–	–	N3T (242)*	N3T (242)*

CD8+ T cell responses were analysed using *ex vivo* IFN-γ-ELISPOT analysis against a panel of HLA-restricted autologous peptides (spot forming units/million PBMCs are shown). The HLA Class IA and Class IB genotypes are A0101, A1101 and B0801, B0801, respectively. Sequencing of the corresponding peptides was carried out at the timepoints shown, and mutating residues have been underlined. * This epitope is a B*57 escape mutant, transmitted to, but not recognised by the recipient, and which shows reversion pre-Rituximab therapy. The results reveal that pre-Rituximab there were two main CD8 T cell responses detectable, neither of which showed evidence of immune-escape throughout the Rituximab-induced viral flare and recovery period.

a.



b.

Days post-seroconversion:	882	1239	1295	1308	1700
HBsAb level (IU/ml):	12.43	7.8	<1	2.86	100.76

c.

Day	IgG (g/l)	IgA (g/l)	Paraprotein g/l (IgM kappa)	Clinical stage
1028	16.1	0.7	38	Pre rituximab
1094	16	0.8	37	Pre rituximab
1144	15.6	0.7	29	Post rituximab
1204	12	0.5	24	
1254	10.8	0.6	25	
1298	11.6	0.7	26	During acute illness
1449	17.3	0.8	23	On ART post Convalescence

Figure 3.8: Clinical course of HIV-1, HBV, hepatitis (ALT) and immunoglobulin levels in relation to Rituximab therapy

Panel A: The plasma viral load of HIV-1, HBV and liver function marker ALT are displayed over time (days since seroconversion). The time of Rituximab therapy is indicated. Asterisks indicate HBeAg detection (tested negative at d677).

Panel B: below x-axis of Panel A indicate the change of the serum HBsAb over time.

Panel C: immunoglobulin level over time.

3.5 Discussion

In this study, B cell depletion led to a decline in total IgG levels and NAb titre concurrent with rising pVL. Recovery of NAb titres coincided with control over viral load and viral diversity, implying the major role for CD20+ cells in the control of chronic HIV infection. Multiple analytical techniques were employed to longitudinally characterise the distinct virus population that emerged as NAbs returned. The common ancestor of this new virus population was defined by polymorphisms at two residues exposed on the surface of the virus envelope, both of which showed evidence of positive selection. These data informed a strategy of mutagenesis which confirmed these sites altered susceptibility to known NAbs including the CD4 binding site specific antibody, IgG1b12. Subsequent antibody-envelope glycoprotein binding studies demonstrated this patient possessed b12-like antibodies. Although this is a study of a single patient, the fact that rituximab was given in the absence of antiviral therapy coupled with the ability to analyse sequences from the time of seroconversion onwards make this case especially important.

A number of potential explanations for the impact of B cell depletion exist. Firstly, there may be a profound loss of antiviral Neutralising antibodies (NAb). However, while CD20 is present on all mature B cell populations it is absent on plasma cells, and CD20 depletion will also not impact on pre-existing antibody, which has a half-life of 3-4 weeks. Consistent with this we did not observe any change in titres of neutralising antibodies against laboratory strains (JR-CSF and NL4-3; data not shown), and we did not see re-emergence of historical strains from early in infection, including those under prior selection.

However, it is likely that other defects in production of Nab or non-Nab may be important. In particular depletion of naïve B cells may have delayed the production of antibodies against newly emerging strains. A critical role for naïve B cells in containment of diverse influenza strains has been recently shown (Rangel-Moreno, et al. 2008).

It is possible that depletion of B cells had an indirect impact on viral load through modulation of CD4 or CD8+ T cell subsets in one of follow ways: 1) depletion of CD20+ cells may have promoted proliferation and activation of CD4 T cells in lymphoid complexes, generating increased targets for HIV infection, or, 2) as has been postulated in auto-immunity, the B cells may act as both regulatory and antigen presenting cells and a reduction in T cell activation and an increase in T_{reg} has been observed (Cross, et al. 2006, Liossis and Sfikakis 2008, Vallerskog, et al. 2007). We did not have material to examine this here but prospective studies of the impact of rituximab on antiviral T cells may shed light on this issue. Furthermore, while we cannot exclude this, the mechanism for such an indirect effect is not clear and, importantly, no impact on CD8+ T cell responses has been observed in the SIV models quoted above (Miller, et al. 2007). The delay of several weeks that is seen between Rituximab treatment and the change in viral load is much more consistent with the observed decline in antibody titres than a direct B cell/T cell interaction, which would be expected to take effect immediately. Additionally the changes in the prevailing sequence (reversion and subsequent re-selection of a mutated NAb epitope; Figures 3.2, 3.4 and 3.5) are consistent with relaxation and reapplication of NAb-mediated selection pressure.

The impact of other reactivating viruses on HIV load is complex. As in this case, Rituximab treatment is known to have the potential to reactivate hepatitis B virus (HBV) (Aksoy, et al. 2007), even in those with previously controlled infection (HBsAg-ve, HBcoreAb +ve). This points to an important role for B cells in long-term control of HBV – a virus where CD8+ T cells play an important role in acute disease. The relationship between HBV pVL and HIV pVL is not fully understood, but overall these appear to be independent (Rockstroh 2006). In this case HBV and HIV pVL rose within a similar timeframe, one may postulate that concurrent viral infection may lead to pro-inflammatory responses and hence favouring HIV replication. The anti-HIV Nab may have decreased secondarily to combined effect of substrate binding exhaustion (during peak viraemia) and Rituximab activity. However, control over HIV occurred spontaneously while HBV load remained high, indicating that they were differentially regulated in this patient (Figure 3.8), and since we do not observe similar flares in HIV VL during acute hepatitis C superinfection (Thomson et al, unpublished data), the role of HBV co-infection appears less likely to be dominant. In addition, the hypothesis of "antibody depletion sequestrated by the increasing viraemia" is not supported by the non-specific decreasing in all types of immunoglobulin (overall IgG level, anti-HBV, and anti-HIV) seen in this study.

Overall, our findings support previous suggestions of an evolutionary balance between virus and NAbs (Burton, et al. 2005, Wei, et al. 2003). At first sight these findings appear at odds with previous studies that fail to confirm an association between long-term virus control and neutralizing antibody response, especially in elite controllers (Bailey, et al. 2006). However, as shown here, the dynamics of immune responses are critical to their interpretation, so complex relationships could easily be

missed in cross-sectional studies, especially analyses of those with favourable CD8+ T cell responses (Goulder and Watkins 2008).

In conclusion, this unique study suggests that B cells, and their secreted NAbs can impact upon HIV viral load in chronic infection. By comparison the impact of the most protective HLA Class I allele, B*57, is approximately 1 $\log_{10}$, and CD8+ depletion studies in macaques showed equivalent changes to those seen here in 2/3 animals (Kiepiela, et al. 2004, Schmitz, et al. 1999). This evidence, derived directly from observations in man, may inform the rational design of future immunotherapies and HIV-vaccines.

CHAPTER 4:

Molecular epidemiology and drug-associated mutations in the Free State, South Africa

---

This work is based on a manuscript published in Antiviral Therapy (Huang, et al. 2009).

4.1 Background

In South Africa (SA), 5.5 million people are infected with HIV(ASSA 2007, UNAIDS/WHO 2007, Venter 2005). Since 2004, the SA government has made antiretroviral therapy (ARV) available to HIV positive individuals through public sector ARV programs - an initiative with significant clinical, public health and economic impact (Department of Health 2003, Egger, et al. 1997, Frater, et al. 2002, Palella, et al. 1998, Sow, et al. 2007, Venter 2005).

Until recently, only a few small to medium-sized South African cohorts have reported the molecular characterisation and baseline pre-therapy resistance profiles of HIV infection, predominantly amongst the KwaZulu Natal, Gauteng and Limpopo provinces (Bessong, et al. 2006, Gordon, et al. 2003, Pillay, et al. 2002, Shekelle, et al. 2007). There is currently no similar information available from other less urbanised regions of SA, such as Free State Province with a population of approximately 3 million, an antenatal clinic HIV prevalence of 33.9%, and the life expectancy estimate of 46.5 years (ASSA 2007). Under the Free State Comprehensive Care, Management and Treatment of HIV and AIDS program, highly active antiretroviral therapy

(HAART) has been available since 2004 as a combination stavudine and lamivudine with either nevirapine or Kaletra, and has significantly reduced the number of HIV-related deaths (hazard ratio 0.14) (Fairall, et al. 2008).

Guidelines exist for the surveillance of drug resistance in patients receiving HAART and in drug-naïve individuals – the World Health Organization (WHO) and Stanford Drug Resistance Database publish a list of mutations for use in surveillance of transmitted drug resistance mutations (TDRM) and the International AIDS Society (IAS)–USA produce a list for use in the monitoring of treated populations. The WHO, Stanford and IAS–USA committees regularly update the definitions and clinical significance (phenotypical resistance, clinical association and outcome of meta-analyses) of drug resistance mutations and drug-associated polymorphisms (Johnson, et al. 2008, Stanford University 2007). Recently, the updated TDRM list from the Stanford database has become unified with the WHO guidelines (Bennett, et al. 2009, Shafer, et al. 2007). The mutations in the TDRM surveillance list are selected on the basis that they are non-polymorphic, clinically significant and applicable across all subtypes. The implementation of such a standard should facilitate a specific method to assess changes in drug resistance prevalence in populations over time (Bennett, et al. 2009, Gilks, et al. 2006, Team 2003).

In patients who report that they are drug-naive, resistance mutations might either reveal transmitted variants or might indicate that patients are not, in fact, drug-naive or are unaware of previous ARV exposure (Garcia-Diaz, et al. 2008). The extent to which unrecognized access contributes to the level of resistance in patients enrolling onto public sector ARV is unknown in South Africa, and very difficult to measure.

Prior to the introduction of the public sector ARV treatment programme, ARV drugs were available for the prevention of mother-to-child transmission (PMTCT) through private clinics and other unregulated routes, although previous studies suggest a low prevalence of baseline drug resistance in South Africa (Bessong, et al. 2006, Gordon, et al. 2003, Pillay, et al. 2002, Shekelle, et al. 2007).

Here, we studied 425 HIV type-1 (HIV-1)-positive patients newly recruited to the public sector ARV treatment programme in the Free State province, South Africa and reporting to be drug-naive. We characterized the molecular epidemiology of HIV-1 infection within this region and measured the prevalence of drug resistance and other polymorphisms associated with drug exposure in these individuals. We found a correlation between low CD4 T cell counts and drug-selected polymorphisms, suggesting that many mutations in drug-naive individuals are not transmitted, but are the result of acquired resistance through unrecognised ARV access.

4.2 Aims

(i) To describe the viral molecular epidemiology of Free State Province, South Africa.

(ii) To investigate the prevalence of drug-associated mutations, and resistance to HAART in the pre-therapy cohort.

(iii) Model the impact of ART availability on baseline antiretroviral resistance.

4.3 Results

4.3.1 Characterisation of Patients Infected with HIV in Free State, South Africa

The 884 patients were stratified according to CD4 T cell count, and viral sequences were analysed from 425 patients to form the low (<100 cells/µl; n=195), intermediate (200–349 cells/µl; n=120) and high (>500 cells/µl; n=110) CD4 T cell counts stratification groups. The mean CD4 T cell counts for all patients within the low, intermediate and high groups were 45.4, 270.0 and 651.8 cells/µl, respectively. The mean age of the cohort was 36.1 years (Table 4.1). From the 425 patients representing the three CD4 T cell counts stratifications, 44 were initially sequenced locally in South Africa as part of a preliminary analysis in previous study (Seebregts, et al. 5-8 June 2007), and 381 were then sent to the Peter Medawar Building for Pathogen Research (Oxford, UK). There was no difference in the sampling strategy for the two groups. Complete *protease* gene sequences (99 amino acids) and amino acids 1–530 of the *RT* gene were successfully amplified from 390 and 397 patients, respectively. The REGA subtyping tool revealed all isolates to be subtype C. All sequences were combined with reference sequences from the BioAfrica database and their phylogeny analysed using maximal likelihood trees.

Table 4.1: Details of HIV-1 positive patients recruited to three subgroups according to CD4 T cell count

Category	Complete cohort	CD4+ T-cell count <100 cells/μl	100–199 cells/μl	200–349 cells/μl	350–499 cells/μl	≥500 cells/μl	*P*-value (χ^2 test for trend)
Patients in cohort, *n* (%)	884	195 (22.1)	212 (24.0)	244 (27.6)	123 (13.9)	110 (12.4)	–
Patients sampled, *n*	425	195	–	120	–	110	–
Patients with both protease and RT gene sequences, *n*	390	192	–	112	–	86	–
Mean age, years (±SD)	36.1 (9.1)	38 (8.9)	–	35.4 (9.0)	–	34.2 (9.2)	–
Mean CD4+ T-cell count, cells/μl (±SD)	271.3 (200.82)	45.36 (28.66)	–	269.96 (40.86)	–	651.80 (178.91)	–
Patients with drug resistance mutations, *n* (%)[a]	9 (2.3)	7 (3.6)	–	1 (0.9)	–	1 (1.2)	0.134 (0.055)
Total drug resistance mutations, *n*	11	9	–	1	–	1	–
Patients with non-accessory drug-associated mutations, *n* (%)[b]	16 (4.1)	13 (6.8)	–	2 (1.8)	–	1 (1.2)	0.015 (0.004)
Total non-accessory drug-associated mutations, *n*	19	16	–	2	–	1	–
Patients with any polymorphism at drug resistance-associated sites, *n* (%)[b]	64 (16.4)	37 (19.3)	–	15 (13.4)	–	12 (14.0)	0.193 (0.099)
Total naturally occurring drug-associated mutations, *n*	73	44	–	15	–	–	–

All patients listed in this table were infected with HIV type-1 (HIV-1) subtype C. Patients were stratified into three discrete groups according to baseline CD4+ T-cell count. Patients not falling into these groups were not analysed. The total number of mutations are also shown because some sequences contained more than one mutation. [a]Defined by the International AIDS Society–USA Drug Resistance Mutation list. [b]Defined by Stanford Drug Resistance Database.

4.3.2 Free State HIV sequences are representative of those from the rest of SA, suggesting multiple subtype C introductions into the population

In Figure 4.1, the maximum likelihood tree shows 390 HIV-1 *pol* sequences from the Free State in the context of local and global HIV-1 subtype C viruses from the BioAfrica database. The Free State strains (n=390) are in red, the South African strains (n=428) are in green, and remaining global subtype C strains (n=551) are in black. In Figure 4.2, the tree shows the phylogeny of the Free State patients (n=390) in the context of local HIV-1 subtype C virus strains from other provinces within SA (n=428). The Free State strains are highlighted in red, the drug resistant strains are highlighted in blue, whilst the SA reference strains (mainly from Gauteng, Kwa-Zulu Natal and Western Cape Province) are highlighted in green.

The Free State sequences appear to be distributed randomly amongst both Southern African (Figure 4.1) and South African (Figure 4.2) reference sequences. Quantitative analysis of clustering showed that 52.5% of sequences occurred in 61 clusters of between 2 and 18 sequences (mean 3.4 patients per cluster: 27 clusters of two, 10 clusters of three, 8 clusters of four, 3 clusters of five, 1 clusters of six, 2 clusters of seven, and 3 individual clusters of eight, nine and eighteen sequences). The analysis also shows that 70.6% of the Free State sequences (in isolation or in clusters) have a South African reference sequence as the nearest neighbour, but there is also clustering of sequences sampled from the Free State compared with other major South African centres (Slatkins and Maddison test, $P<0.001$). In addition, 96.1% of Free State strains are most closely related to a Southern African sequence and 97.2% of Free State strains are most closely related to an African sequence. Only 2.8% of the Free State sequences are closely related to a non-African sequence.

Sequences that contained any drug-associated polymorphism or a mutation from the IAS–USA surveillance list are indicated and are randomly distributed. The absence of linkage of these variant sequences does not support significant transmission of drug resistance within this cohort (Slatkins and Maddison test, P=0.85).

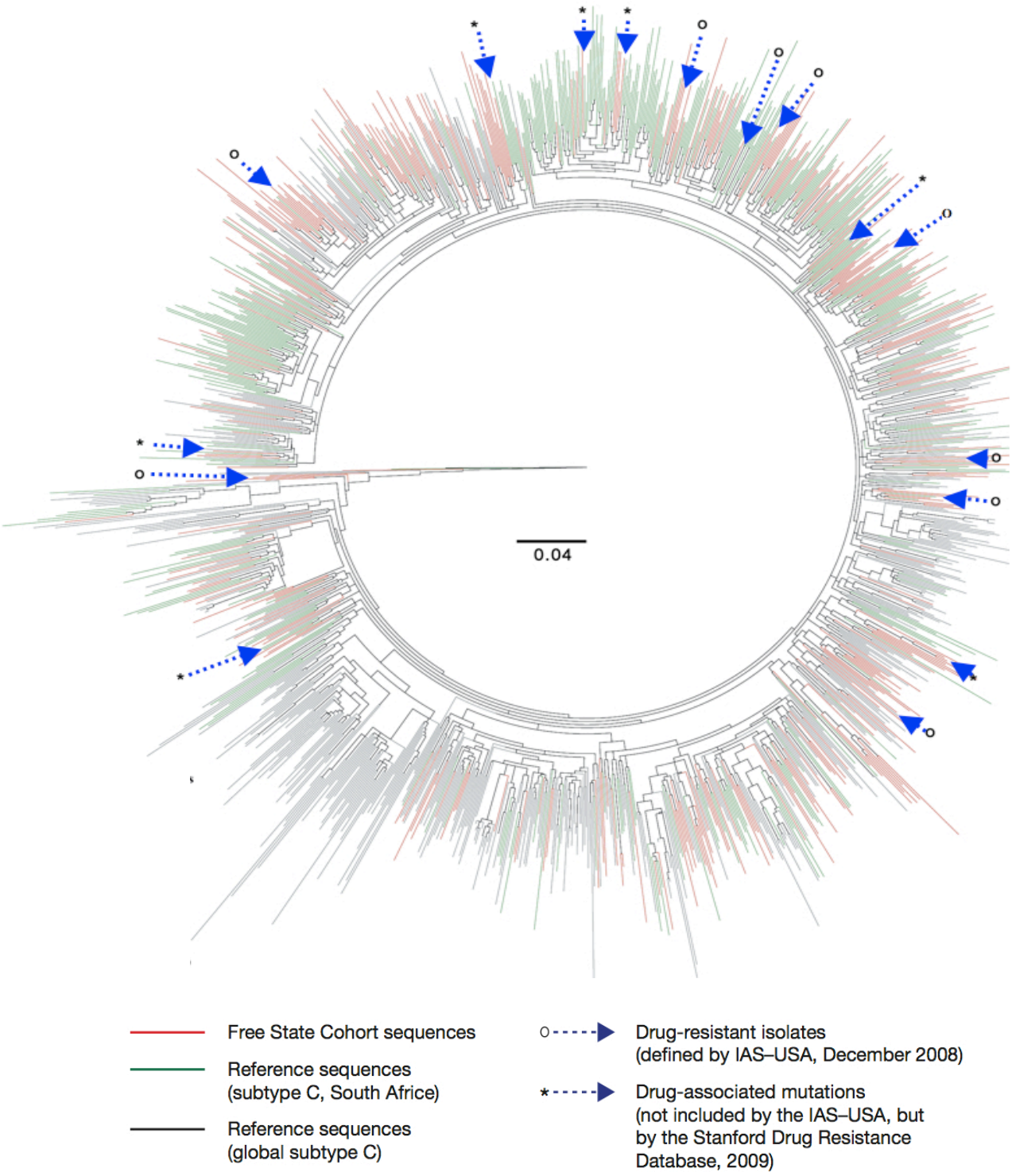


Figure 4.1: Phylogenetic tree of the HIV-1 molecular epidemiology in the Free State, South Africa in the context of global HIV-1 subtype C *pol* reference sequences

The *pol* gene of HIV-1 isolates sampled from Free State patients (red) are shown in the context of published South African (green) and non-South African (black) HIV-1 subtype C reference sequences. Resistant isolates defined by International AIDS Society (IAS)–USA and other drug-associated mutations defined by the Stanford Drug Resistance Database are shown (blue dotted arrows, ending in "o" and "*", respectively). This maximal likelihood tree was constructed using the general time reversible substitution model with optimised proportions of invariable sites and gamma distribution (GTR+G+I) using PAUP* (version 4.0 beta) and PhyML (version 2.4.4) software.

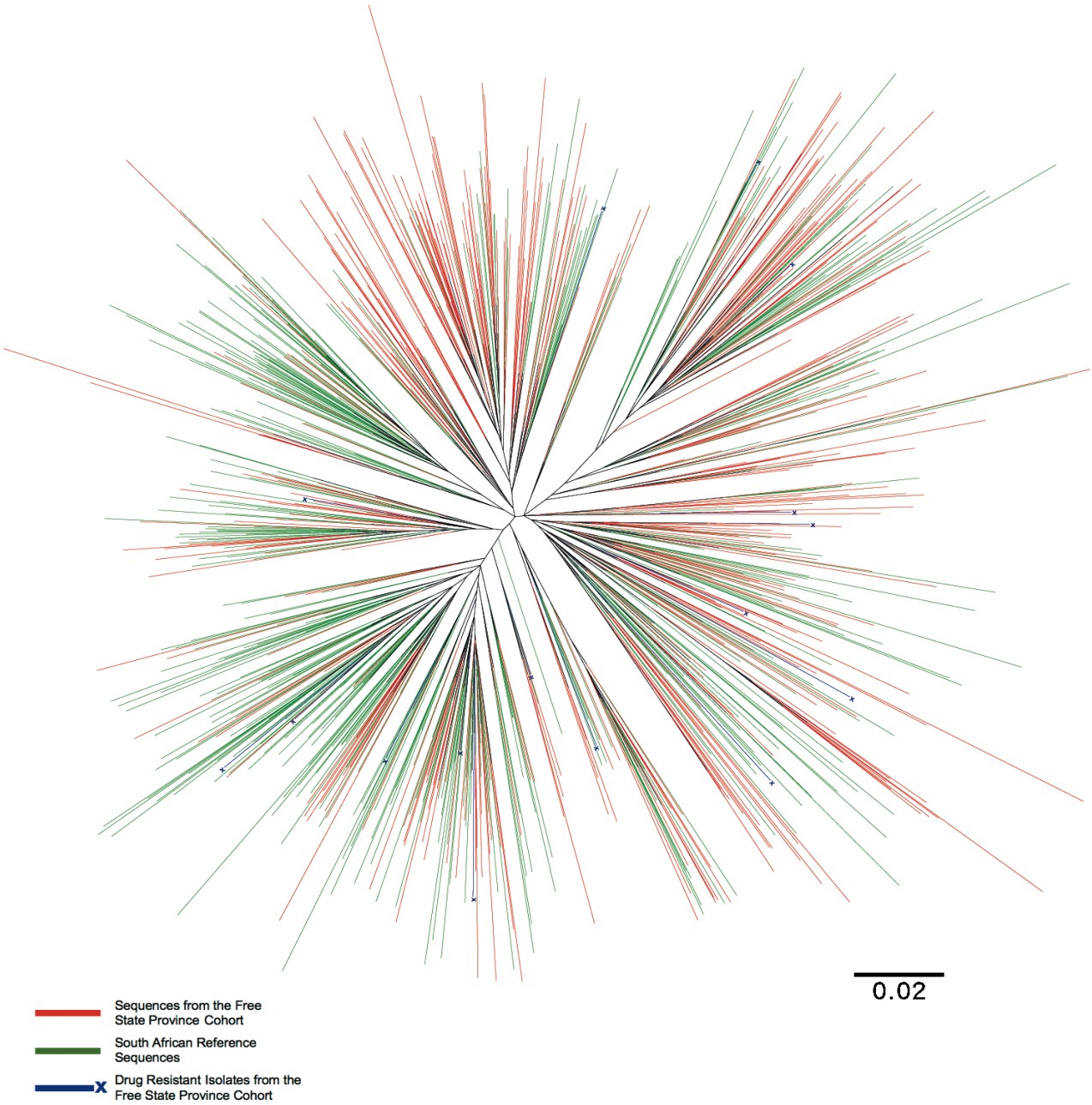


Figure 4.2: Phylogenetic tree describing HIV molecular epidemiology of the Free State Province and its drug resistant subjects in the context of South African HIV-1 subtype C reference sequences

The *pol* gene of HIV isolates sampled from the Free State patients (in red) are shown in the context of published South African HIV-1 Subtype C database (in green). The two HIV population sequences showed remarkable homology, sharing common internal nodes. Only 1 group of 18 Free State isolates clustered with more than 10 sequences. The drug resistant isolates from Free State patients (blue stars) distribute sporadically across the tree. This maximal likelihood tree was constructed using the general time reversible substitution model with optimised proportions of invariable sites and gamma distribution (GTR+G+I) using PAUP* (version 4.0 beta) and PhyML (version 2.4.4) software.

4.3.3 Low prevalence of pre-therapy drug resistance in the Free State

All available pre-therapy *pol* gene sequences (n=390 for PR; n=397 for *RT* amino acids 1-530) from the Free State cohort were submitted to the Stanford HIV Resistance Database for identification of polymorphic sites associated with drug resistance. Although all potential resistance mutations were identified, specific attention was paid to the mutations from three drug classes currently being provided as part of the Free State public HAART regimen, namely PI, NRTI, and NNRTI.

According to the Stanford Drug Resistance Database list, the prevalence of non-accessory mutations was 4.1%, comprising 16 patients carrying a total of 19 non-accessory mutations (Table 4.1, Accessory mutations are defined as "atypical mutations or subtype-associated polymorphisms possibly related to secondary drug resistance" (Stanford University 2007)). The overall prevalence of clinically significant drug resistance mutations (according to the IAS–USA classification) was low (2.3%), comprising 9 patients carrying 11 mutations (Table 4.1). A total of 64 (16.4%) patients sampled had a mutation at any of the drug resistance amino acid sites, including those that have been found to be naturally occurring polymorphisms (Table 1). Of these, the most common accessory mutations for the *protease* gene were L10I/V (n=7, 1.8%), L11IV (n=4, 1.0%) and L89I (n=7, 1.8%). For the *RT* gene, the most common mutations were V118I (n=17, 4.4%) and G333E (n=4, 1.0%).

The 16 patients with non-accessory mutations – according to the Stanford definition – are detailed in Table 4.2. Of these, nine had clinically significant mutations according to the IAS–USA list. These were Y181C (n=2), K103N (n=3), Y188L (n=1), V106M (n=1), V108I (n=1) and K219E (n=1) in the *RT* gene, and M46L (n=1) and N88S

(n=1) in the *protease* gene. The referral centres and local clinics where these patients had been recruited were either visited or contacted; however, in all but five cases the patients had been lost to follow-up. Of the four who had maintained undetectable viral loads on therapy, none had clinically significant IAS–USA defined mutations at baseline (V179D, M46L, A98G and T69N). The one patient identified with a detectable viral load on ARVs (13,000 RNA copies/ml after 5 months) had V179D at baseline, although no sequence data was available from subsequent samples to determine the development of resistance (Table 4.2). Of the 16 patients with significant resistance, 6 were male. Of the 10 females, the median age was 34.5 years (range 25–56).

The non-accessory mutations identified in the cohort are explored in more detail in Table 4.3. For each mutation, the number in each CD4 T cell counts stratification is shown and compared with the reported prevalence in databases of drug-naive subtype B and C populations and in drug-experienced subtype C populations. To determine the clinical implications of these identified mutations, the Stanford database MARVEL report for each is summarized and the significance of the mutations in the surveillance lists of the WHO (updated 2009) and IAS–USA (updated December 2008) are compared (Table 4.3). For all the mutations, the prevalence in databases of drug-experienced subtype C cohorts was higher than in drug-naive patients with implications that these mutations might act as markers of drug exposure, even if not conferring clinically relevant resistance.

Table 4.2: Patients with drug-associated mutations to antiretroviral drugs during pre-therapy assessment

Patient identification	Gender	Age, years	CD4+ T-cell count, cells/μl	PI major mutation (grade of resistance)[a]	NRTI mutation (grade of resistance)[a]	NNRTI mutation (grade of resistance)[a]	Response to therapy
OX2032	F	35	1	–	–	E138K (potentially low) and Y181C (high)	Lost to follow-up
OX428	M	37	2	–	–	A98G (potentially low)	Lost to follow-up
OX927	F	25	3	–	K219E (low)	Y181C (high)	Lost to follow-up
OX195	M	41	5	–	T69N (potentially low)	–	Lost to follow-up
OX693	M	31	14	–	–	V179D (potentially low)	VL=13,000 copies/ml 5 months after starting therapy
OX1	F	42	23	–	–	K103N (high)	Lost to follow-up
OX613	F	24	28	–	–	V179D (potentially low)	VL<25 copies/ml 28 months after starting 3TC/d4T/EFV
OX1082	M	37	30	–	–	V106M and Y188I (both high)	Lost to follow-up
OX2233	F	28	42	–	–	K103N (high)	Lost to follow-up
OX677	F	31	60	M46L (low)	–	–	VL<25 copies/ml 23 months after starting 3TC/d4T/EFV
OX2011	M	37	63	–	–	A98G (potentially low)	VL<25 copies/ml 33 months after starting 3TC/d4T/EFV
OX5	F	56	64	–	T69N (potentially low)	–	VL<25 copies/ml 36 months after starting 3TC/d4T/EFV
OX1006	M	43	64	–	–	V108I (potentially low)	Lost to follow-up
OX2519	F	32	205	–	–	V179D (potentially low)	Lost to follow-up
OX312	F	38	215	–	–	K103N (high)	Lost to follow-up
OX626	F	39	550	N88S (intermediate)	–	–	Lost to follow-up

The table shows observed pre-therapy drug-associated mutations (class of antiretroviral agent, mutation position and grade of resistance), arranged in ascending order of CD4+ T-cell count and virological response (if known). [a]Grades of resistance are defined using the Stanford Drug Resistance Database. d4T, stavudine; EFV, efavirenz; F, female; M, male; VL, viral load; 3TC, lamivudine.

Table 4.3: Drug-associated mutations in the Free State Cohort

	PI		NRTI		NNRTI							
Category	M46L	N88S	T69N	K219E	A98G	K103N	V106M	V108I	E138K	179D	Y181C	Y188L
Free State Province cohort												
Patients with CD4⁺ T-cell count <100 cells/µl (*n*=192), *n*[a]	1	0	2	1	2	2	1	1	1	2	2	1
Patients with CD4⁺ T-cell count 200–350 cells/µl (*n*=114), *n*[b]	0	0	0	0	0	1	0	0	0	1	0	0
Patients with CD4⁺ T-cell count >500 cells/µl (*n*=91), *n*[c]	0	1	0	0	0	0	0	0	0	0	0	0
Stanford drug resistance database												
Mutation prevalence in drug-naive subtype B patients, (*n*=7,404 for PR, *n*=5,539 for RT), %	0.3	0	0.4	0	0.2	0.3	0	**0.6**	0.2	**2**	0	0
Mutation prevalence in drug-naive subtype C patients (*n*=2,145 for PR, *n*=1,979 for RT), %	0.1	0	0.1	0	**0.5**	0.3	0	0.3	0.2	**0.5**	0.1	0.1
Other HIV-1 subtypes (drug-naive) with mutation prevalence ≥0.5% (subtypes A, D, F, G, AE and AG)	Subtype G (**0.5%**, *n*=619)	No	Subtype AE (**0.5%**, *n*=762)	No	No	No	No	Subtype AG (**1.2%**, *n*=1,025)	No	Subtype D (**4%**, *n*=320), subtype AE (**1.5%**, *n*=762)	No	No
Mutation prevalence in drug-experienced subtype C patients (*n*=282 for PR, *n*=1063 for RT), %	2.8	3.2	4.2	7	8.2	28	13	4	1.3	4.5	11	4.1
Genotype–treatment correlation?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Genotype–phenotype correlation?	1.8–8.2	0.1–8.9	NA	NA	NA	22–51	11–356	NA	NA	NA	1.3–148	4.3–400
Genotype–virological correlation?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	Yes	Yes	NA
Mutation grading (Stanford Drug Resistance Database)	Low	Intermediate	Potentially low	Low	Potentially low	High	High	Potentially low	Low	Potentially low	High	High

Table 3 shows the non-accessory drug-associated mutations in the Free State cohort and in reference cohorts (values ≥0.5% are in bold). The genotype–treatment correlation row shows whether the mutation is associated with therapy. The genotype–phenotype correlation row shows *in vitro* evidence (fold resistance) for associated phenotypic resistance. The genotype–virological correlation row shows whether mutations have an effect on viral load on therapy. All results are summarized from MARVEL and the Stanford Drug Resistance Database. The significance of each resistance mutation is compared with the drug mutation lists of the International AIDS Society–USA and the World Health Organization. *P*-values were calculated using the χ^2 test for trend, comparing the resistance statistics across the three stratified patient populations. The level of significance for clustering of mutations among low CD4⁺ T-cell counts (χ^2 test for trends) was *P*=0.004 for highly active antiretroviral therapy and *P*=0.005 for non-nucleoside reverse transcriptase inhibitors (NNRTIs). [a]Total mutations =16. [b]Total mutations =2. [c]Total mutations =1. HIV-1, HIV type-1; NA, not applicable; NRTI, nucleoside reverse transcriptase inhibitor; PI, protease inhibitor; PR, protease; RT, reverse transcriptase.

4.3.4 Clustering of drug-associated mutations among patients with low CD4 T cell counts

Drug-associated polymorphisms (based on the Stanford database) were concentrated among patients with low CD4 T cell counts – 6.8% of patients with CD4 T cell counts <100 cells/µl carried non-accessory mutations compared with 1.8% and 1.2% of patients with intermediate and high CD4 T cell counts, respectively (P=0.015; Table 4.1 and Figure 4.3). The prevalence of resistance according to the IAS–USA definition was 3.6%, 0.9% and 1.2% for low, intermediate and high CD4 T cell counts groups, respectively (Table 4.1). Although not statistically significant, there were more accessory mutations among patients with low CD4 T cell counts (19.3%) compared with the intermediate (13.4%) and the high (14.0%) CD4 T cell counts groups. When using the more relaxed Stanford definition, mutations clustered among patients with low CD4 T cell counts for all mutations (Figure 4.3; P=0.004). Although not statistically significant, when only considering clinically relevant mutations, there was a trend for clustering of resistance among low CD4 T cell counts for the IAS–USA (P=0.055) and the WHO TDRM lists (P=0.086; Table 4.3).

The NNRTI class, in particularly Nevirapine, was associated with significant baseline drug-associated mutations amongst the low CD4 T cell counts patient strata. The genotypes and phylogeny of the NNRTI resistance strains were heterogenous (Figure 4.3 and Table 4.2) and were not derived from any common founder (Figure 4.1 and 4.2). Of the patients with low CD4 T cell counts (<100 cells/µl), 5.2% possessed non-nucleoside reverse transcriptase inhibitor (NNRTI)-associated mutations compared with 1.8% of patients with intermediate and 0.0% of patients with high CD4 T cell counts (P=0.005; Figure 4.3 and Table 4.3). This association remained significant

when restricted to NNRTI mutations on the IAS–USA list (P=0.031; Table 4.3). The distribution of NNRTI mutations within patients with low CD4 T cell counts was concentrated further among the individuals with lowest CD4 T cell counts (Table 4.2), who also tended to have higher grade resistant mutations.

The major mutation genotypes observed for NNRTI resistance were K103N (grade 5; 3 isolates), V106M (grade 5; 1 isolate), Y181C (grade 5; 2 isolates), and Y188L (grade 5; 1 isolate). Other NNRTI mutations included K103R (grade 1; 3 isolates), V108I (grade 2; 1 isolate), A98G (grade 2; 2 isolates), V179D (grade 2; 3 isolates), E138K (grade 3; 1 isolate), (Table 4.2).

In contrast, the prevalence of resistance to non-NNRTI classes of HAART was much lower (Table 4.2 and 4.3). Only two major PI resistance mutations in *protease* were observed (M46L (grade 3) and N88S (grade 4)) and two NRTI mutations (T69N (grade 2) and K219E (grade 3)).

Certain individuals were identified with clusters of mutations, indicative of potential multi-class resistance (Table 2-3). One patient (OX927) had significant resistance to both NRTI (K219E) and NNRTI (Y181C). One patient (OX1082) possessed two significant NNRTI resistance mutations, V106M and Y188L. Another (OX2032) carried E138K and Y181C.

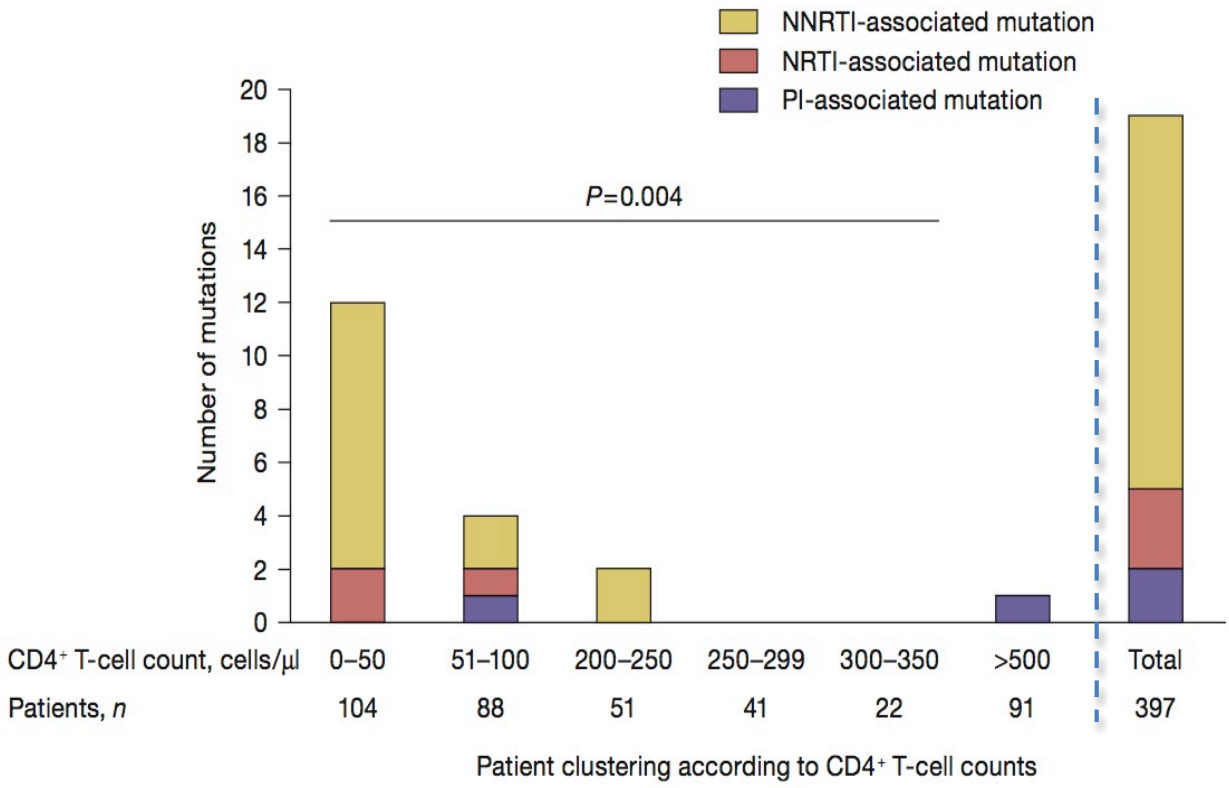


Figure 4.3: Distribution of drug-associated mutations in pre-therapy patients according to CD4 T cell counts

The plot depicts all pre-therapy patients carrying drug-associated mutations to any of the highly active antiretroviral therapy constituents, including protease inhibitors (PIs), nucleoside reverse transcriptase inhibitors (NRTIs) and non-nucleoside reverse transcriptase inhibitors (NNRTIs), stratified according to CD4 T cell count. The P-value was calculated using the $\chi 2$ test, comparing the mutations across CD4 T cell counts.

4.3.5 Mathematical modelling of acquisition of drug resistance

(In this section the mathematical modelling was carried out in collaboration with Dr Helen Fryer and Professor Angela McLean)

The presence of inducible drug-resistance mutations in a population prior to recruitment to the ARV program indicates either access to antiretroviral therapy from alternative routes or transmission of drug-resistant strains. A mathematical model was developed to address two questions. Firstly, how long would it take for the observed level of inducible resistance mutation prevalence to develop in the absence of non-governmental access to drugs? Secondly, if alternative routes of drug access are maintained at the current rates, what is the likely prevalence of drug resistance in the next ten years?

Regarding the first question on the attributable impact on baseline inducible drug resistance of non-governmental access to ARVs, the model estimates that it should take between 20 and 35 years for a treatment naïve population exposed to NNRTI as part of HAART to achieve the observed level of resistance in the Free State cohort (Figure 4.4a). This supports the existence of additional sources of NNRTI exposure and that the Free State population is not entirely treatment naïve before the government program.

In answer to the second question, the model predicts that in patients with low CD4 T cell counts the prevalence of NNRTI resistance will increase to 7.1% (confidence interval (CI): 5.7 to 8.2%) over the next 10 years, if the additional sources of NNRTI exposure are maintained (Fig 4.4b). The resistance prevalence can however be stabilised to 5.2% (CI = 3.6 to 6.2%) over next 10 years, if additional exposure to

NNRTI causing resistance can be reduced to 1% (currently 4.2%) (Figure 4.4c). If all additional NNRTI exposure could be restricted, the present resistance prevalence could not sustain itself and would be expect to fall.

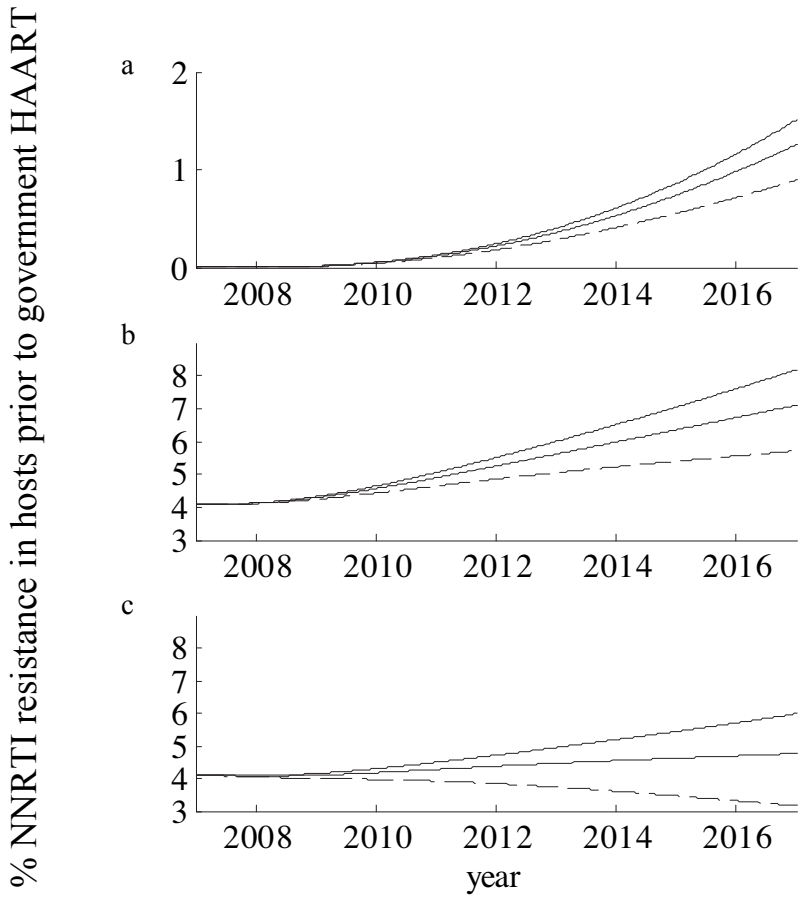


Figure 4.4: Model predictions of how the prevalence of NNRTI resistance in hosts, prior to enrolment onto governmental ARTs, is expected to develop over time in the Free-State population

Panel A: A projection of how resistance would develop in the absence of alternative, non-governmental sources of ARTs, if there was no initial resistance to NNRTI.

Panel B: A projection of how NNRTI resistance will increase over the next 10 years if the percentage of infected hosts acquiring resistance from alternative access to therapy continues at the current level of 4% (estimated using the model).

Panel C: A projection of resistance emergence over the next 10 years, if the percentage of infected hosts acquiring resistance from alternative access to therapy reduces to 1%.

Limited data were available to estimate the rate at which drug resistant virus reverts to drug sensitive virus in untreated hosts (ψ), hence all model predictions were made for three different reversion rates: $\psi = 0.13$ years^{-1} (dashed line), $\psi = 0.05$ years^{-1} (solid line) and $\psi = 0.01$ years^{-1} (dotted line).

4.4 Discussion

Here, we present a cohort analysis of the molecular epidemiology of the HIV-1 epidemic in the Free State Province of South Africa, and of the prevalence of circulating pre-therapy drug resistance in this major site for the government ARV treatment program. There were three key findings. Firstly, there was evidence of multiple introductions of HIV-1 into the Free State, but random distribution of drug resistance-associated polymorphisms. Secondly, the overall prevalence of pre-therapy drug resistance was low, but drug-selected polymorphisms were concentrated among patients with low CD4 T cell counts. Thirdly, baseline drug resistance may be driven by exposure to ARVs available through non-governmental routes and that, unless drug availability is controlled, resistance prevalence is likely to rise.

Patients attending antiretroviral clinics in Free State province during 2006 demonstrated a mean CD4 T cell counts of 271 cells/μl, with 44.8% of patients possess a CD4 cell counts of less than 200 cells/μl, significantly lower than other published chronic HIV cohorts (Kiepiela, et al. 2004). In an earlier study of 2777 chronically HIV infected individuals in the Pelonomi Hospital, Bloemfontein recruited between 1991 and 1997 (van der Ryst, et al. 1998), the mean CD4 T cell counts was 421 cells/μl. The low mean CD4 T cell counts seen in the Free State Cohort is likely a result of the cohort we studied – ARV clinics tend to be enriched with patients presented during symptomatic phase of late HIV. There was no public ARV program available during 1990s. The availability of HAART from the government is expected to improve quality of life and reverse this population-wide fall in CD4 cell counts (Venter 2005).

The Free State is located in central South Africa and is relatively isolated from other South African centres. Our phylogenetic analysis confirmed that Free State province is dominated by HIV-1 subtype C infection, much like the rest of South Africa (Table 4.1). There were multiple small but distinct clusters within the Free State, which suggests the mixing of HIV-1 strains from across South Africa, with migration into and through the Free State, possibly a result of the mining industry, transport routes and other economic reasons (Figure 4.1 and 4.2). However, despite these clusters, viruses with drug-associated mutations are distributed randomly, suggesting that they have evolved recently as a result of individual drug exposure rather than being transmitted within the Free State or introduction into the province.

The prevalence of mutations considered to be clinically important was low (2.3%), in concordence with the findings of previous smaller studies from South Africa (Bessong, et al. 2006, Gordon, et al. 2003, Pillay, et al. 2002, Shekelle, et al. 2007). However, in a supposedly drug-naive cohort, mutations predominated among patients with low CD4 T cell counts. The clustering of mutations with low CD4 T cell counts suggests that, rather than being transmitted, these polymorphisms had been selected by drug exposure. Transmitted mutations should be more common in recently infected individuals as they may revert to wild type over time in the absence of therapy. Chronically infected patients with low CD4 T cell counts would be expected to have fewer mutations if transmission was the only source of drug resistance. Our cohort does not comply with WHO guidelines for the identification of transmitted resistance (i.e. patients with recent infection, aged <25 years and no history of pregnancy (Bennett, et al. 2008, Gilks, et al. 2006, Team 2003)), and although transmission of mutations is documented in other cohorts (Gifford, et al. 2007, Gilks,

et al. 2006, Pillay 2007, Richman, et al. 2004), a combination of phylogeny and the association with lower CD4 T cell counts in our data suggested that transmission was unlikely to be a significant cause of resistance in the Free State.

The Stanford drug resistance database classifies mutations according to drug inducibility, viral phenotype and clinical outcome (Stanford University 2007). In our cohort, the non-accessory mutations not conferring major resistance were T69N, A98G, V179D (all of which were potentially low grade) and E138K (low grade). The WHO TDRM list also excludes these four mutations and one additional mutation (V108I for NNRTI) because they are considered polymorphic, with a pretreatment prevalence of >0.5% in at least one HIV-1 subtype. However, certain mutations excluded by the WHO list are not polymorphic in clade C populations: the prevalence of T69N, V108I and E138K is <0.5% and prevalence of A98G and V179D equals 0.5%. As many of these low-grade resistance mutations are excluded from the IAS–USA and WHO surveillance lists, we have been careful to use the term 'drug-associated mutations' when these mutations were described. We included an analysis with these low grade and potentially low grade mutations, although they are not currently clinically important, because they are selected by therapy and might have a role as surrogate markers of treatment exposure. It is clearly ideal to have a single mutation list to screen all populations; however, the possibility exists that for South Africa, where the dynamics of therapy provision and the more recent introduction of universal access to HAART makes for a unique situation, a subtype C- specific list might become applicable.

Patients were counselled by trained local HIV-1 nurses regarding drug history whether through the prevention of PMTCT, ARV programmes in different provinces, countries and private clinics or other non-governmental sources. A key route of ARV exposure is the nevirapine-only regimen initially used for PMTCT from 2000 (Department of Health 2003, Jourdain, et al. 2004, Shekelle, et al. 2007). Those patients with major NNRTI mutations were predominantly women of child-bearing age. Unfortunately, despite contacting or visiting the referring clinics, the demographic details to determine the extent of this effect were not available for this analysis because of loss of follow-up for most of the affected patients. Nevirapine is an inexpensive and effective agent and a key constituent of HAART regimens in resource-limited settings. In our cohort, nevirapine had the highest prevalence of drug-associated mutations. The drug has a long half-life, a simple mutational pathway and is prone to rapid resistance even with the single doses used in PMTCT (Grossman, et al. 2004, Jourdain, et al. 2004, Shekelle, et al. 2007).

According to the mathematical model, the prevalence of NNRTI resistance is likely to increase, if additional non-governmental ART exposure is not controlled. The implications of this would include reduced drug efficacy and a requirement for clinical drug resistance surveillance, with its inherent financial costs. Newer antiretroviral classes, such as integrase and CCR5 inhibitors are currently unavailable through SA government programs re-enforcing the need to preserve the NNRTI class. In December 2007, South Africa revised its PMTCT regimen to a combination of zidovudine and nevirapine. Although this measure should improve PMTCT efficacy and reduce the emergence of NNRTI resistance, the logistics of introducing this

change might delay its implementation in all areas of South Africa and a proportion of mothers might still receive nevirapine monotherapy.

Nonetheless, PMTCT cannot be responsible for all the observed resistance in the cohort – 6 of the 16 patients were men and the age of the females ranged from 25 to 56 years. These patients predominantly had lower grade mutations, suggestive of exposure but not clinical resistance. It is possible that some individuals had previously received treatment in other clinics, for example, contract workers travelling between provinces. In addition, patients with lower CD4 T cell counts are more likely to be symptomatic, to use drugs prescribed to family or friends and might have previously sought medication through other non-government sources (Novak, et al. 2005, Uy, et al. Jul 2007). At the time of sampling, drugs such as nevirapine, lamivudine, stavudine and didanosine were also available through private practitioners in the Free State, although the duration for which drugs would be prescribed is dependent on the patient's ability to pay (Dahab, et al. 2008, Garcia-Diaz, et al. 2008).

In conclusion, this large cohort from South Africa reveals new molecular epidemiological and drug resistance surveillance data. The prevalence of drug-associated mutations among patients is reassuringly low, but the association with low CD4 T cell counts is previously undescribed and warrants close monitoring of the virological response when these patients start therapy.

CHAPTER 5:

Study of Cytotoxic T cell Mediated Immune Selection in Advanced HIV Infection

---

The data presented in Chapter 5 and 6 are based on a manuscript that is currently submitted to the Journal of Experimental Medicine.

5.1 Background

5.1.1 Correlates of disease progression in HIV infection

In the majority of untreated individuals Human Immunodeficiency Virus Type 1 (HIV-1) infection leads to Acquired Immunodeficiency Syndrome (AIDS), associated with opportunistic infections and malignancies. Some patients progress to AIDS quickly, whilst others maintain undetectable plasma viral loads without therapy and do not become unwell for many years. Different HLA Class I alleles are associated with different rates of progression (O'Brien, et al. 2001). However, the mechanism underlying this effect is unclear. Deciphering the correlates of this heterogeneous protection is clearly of importance, as there are implications for the design of vaccines and other interventions.

It is evident that the pace of progression is multifactorial - a mixture of host and pathogen genetics combined with factors such as immune activation, enteric bacterial and product translocation, immune selection, and viral adaptation. Genome-wide

association studies reveal a limited number of SNPs and genes that correlate with progression (Fellay, et al. 2007, Herbeck, et al.), but HLA Class I and the human MHC associate reproducibly (Fellay, et al. 2009). The role of the cell-mediated immune system in HIV-associated disease has received much scrutiny, especially the effect of different HLA Class I alleles. Well-documented examples include the protection conferred by HLA B*57 and B*27 (Carrington and O'Brien 2003, Gao, et al. 2005) in Caucasian individuals and HLA B*58 and B*81 in patients from South Africa (Kiepiela, et al. 2004).

In determining the mechanism by which HLA Class I influences outcome, there are likely to be both host and viral factors. For example, selection of different numbers of self-peptides in the thymus may facilitate broader cross-reactivity of certain T cell clones on exposure to HIV viral variants (Kosmrlj, et al. 2010). Viral adaptation to the selection pressures invoked by both antiretroviral drugs (ARVs) and the immune system are well documented in the forms of drug resistance (Shafer, et al. 2008, Shafer, et al. 2007) and immune escape mutations (Goulder and Watkins 2004, Phillips, et al. 1991), respectively. The latter are widespread across the HIV-1 genome (Bhattacharya, et al. 2007, Brumme, et al. 2007, Brumme, et al. 2008), and may influence outcome in individual patients (Feeney, et al. 2004) and across different populations (Kawashima, et al. 2009).

The prevalence of escape from an effective immune response is determined by the strength of the imposed selection pressure and may explain why the prevalence of escape is greater in HLA Class I alleles associated with protection (Frater, et al. 2007). Although the adapted virus maintains a fitness advantage in the presence of the

selection pressure conferred by cytotoxic T cells (CTL), there may be a significant drop in replicative capacity compared to a wild-type virus in a selection-free environment (Martinez-Picado, et al. 2006, Miura, et al. 2009). We have previously hypothesised that immune escape may therefore result in the maintenance of relatively lower viral loads and clinical advantage (Frater, et al. 2007). This is supported by high reversion rates of escape mutations selected by HLA Class I alleles associated with the greatest clinical advantage following transmission to HLA-mismatched recipients (Brumme, et al. 2008, Duda, et al. 2009).

5.1.2 HLA-mediated CTL selection in progression to AIDS

Studies in acute infection have demonstrated strong CTL mediated immune selection and evidence of viral escape mutation in cognate HLA epitopes (Borrow, et al. 1994, Jin, et al. 1999, Koup, et al. 1994, Matano, et al. 1998, Schmitz, et al. 1999). In contrast, the role of ongoing CTL immune pressure on epitope evolution during chronic and late-stage infection remains unclear (Shankarappa, et al. 1999). Cytotoxic T-lymphocytes (CTL) are thought to be crucial for the control of HIV but ultimately fail to control viraemia in most infected patients (Frater, et al. 2007, Goulder, et al. 1997, Goulder, et al. 1997, Klenerman and Hill 2005). Most of the existing observations came from studies on elite controllers of HIV, who are infected individuals that remain asymptomatic with very low VL for prolonged periods of time. Strong CTL and HLA mediated selection appear to be key forces controlling viraemia.

The immune system ultimately collapses in most untreated HIV infections. Whether cytotoxic T cells maintain selection pressure on HIV in advanced disease at low CD4 T cells counts has not been well established. There are a few small studies and case

reports on immune responses in advanced disease, but as it is now standard of care to prescribe HAART to halt disease progression, the study of HIV and immune response during AIDS is difficult (Allen, et al. 2005, Draenert, et al. 2004, Goulder, et al. 2001, Hay, et al. 1999, Kalams, et al. 1999, Koibuchi, et al. 2005, Lee, et al. 2002, Shankarappa, et al. 1999).

During chronic HIV infection, the CD4 T helper cell population deteriorates qualitatively and quantitatively, potentially impacting on HIV specific CTL and their targeting of both circulating or newly evolved virus (Altfeld and Rosenberg 2000, Grossman, et al. 2006, Kalams, et al. 1999, Matloubian, et al. 1994, Rosenberg, et al. 1997). It is likely that CD8+ve T cells may still be functional in AIDS, although with varied avidity, less polyfunctionality and less differentiation (Addo, et al. 2007), and targeting Env rather than Gag (Zhuang, et al. 2008). It has also been noted that HIV specific CTL precursors and interleukin-2 expandible CTL also reduce as disease progresses (Carmichael, et al. 1993, Jin, et al. 1998). Some studies suggest memory T cells may be preserved with strong, broadly directed CTL of high avidity persisting in chronic infection, but they appear to be poorly activated and do not drive viral escape (Draenert, et al. 2004, Hay, et al. 1999, Koibuchi, et al. 2005). Simultaneously, viral escape impacts on the nature of CTL response. The 'immunodominant' epitopes predominantly targeted by CTL change as disease advances, and the TCR repertoir diversifies structurally and functionally with increasing VL (Goulder, et al. 2001, Kalams, et al. 1999, Lee, et al. 2002).

Other analyses of disease progression came from cohorts of patients on HAART. Studies comparing patients who responded differently to HAART have observed that

broader and stronger CD4 and CTL responses are seen more frequently in patients with intermittent and persistent viraemia (whilst on therapy), and are associated with subsequent viral failure and drug resistance (Karlsson, et al. 2004, Meyer-Olson, et al. 2006).

Evidence of altered CTL function has also been suggested by viral evolution data. Firstly, the limited studies to date have demonstrated that very few viral mutations emerge as disease advances (Draenert, et al. 2004, Frater, et al. 2007, Hay, et al. 1999, Kelleher, et al. 2001, Koibuchi, et al. 2005, Shankarappa, et al. 1999). Secondly, studies of viral divergence and diversity have shown a reduction in diversity in early disease, but static viral divergence in AIDS (Grenfell, et al. 2004, Shankarappa, et al. 1999). Taken together, these results suggest CTL selection becomes 'static' as disease progresses, so becoming unable to select further mutants. However, most of the above evidence is based on small cohorts or case studies and some of the cases studied involve patients already on ARV programs.

If CD8 T cell imposed selection pressure is 'relaxed' due to HIV-induced immune exhaustion this might facilitate reversion of costly escape mutations, a restoration of viral fitness and a subsequent rise in viraemia. Reversion of costly drug resistance mutations has been associated with a rise in viral load and clinical progression (Gandhi, et al. 2003), and therefore a precedent exists to potentially explain the rise in viral load associated with the onset of AIDS. Alternatively, CTL pressure may be maintained, but the virus might develop secondary compensatory mutations which restore the replicative cost of the initial escape mutation. Compensatory mutations

have been reported in chronic infection (Brockman, et al. 2007, Kelleher, et al. 2001), but whether they explain the AIDS-associated rise in viraemia is not known.

In this chapter we aim to explore whether in patients with very low CD4 counts (CD4 cell counts <100 cells/μl) - in the advanced stages of immune deficiency – there is a restoration of HIV viral fitness associated with either the development of compensatory viral mutations or the reversion of mutations due to 'relaxation' of CTL-imposed selection. We hypothesise that the presence of a fitter virus in AIDS might contribute to the rise in viraemia and increased pathogenicity. We investigated 935 untreated HIV-1 infected South African individuals from Bloemfontein and Durban. This cohort was enriched for patients with advanced AIDS and low CD4 counts allowing us to explore the immune responses and viral sequences associated with late stages of disease. We focussed on the selection in the HIV-1 Gag protein, as immune responses to these protein have been associated with lower plasma viral loads (Kiepiela, et al. 2007) and significant viral fitness costs (Martinez-Picado, et al. 2006, Miura, et al. 2009). The results are compared to that of chronic non-progressors (CD4 count >500 cells/μl), and other HIV proteins (Pol and Nef). We found evidence for both the accrual of mutations (as a consequence of on-going HLA-mediated escape or potentially compensatory mutations) and for reversion, the latter possibly due to 'immune relaxation'. Despite the complex interplay between host and pathogen, these data support a key role for the restoration of viral fitness as a determinant of AIDS progression and help explain the rise in viraemia in terminal stages when there are few residual CD4 positive target cells.

5.2 Hypothesis

Depletion of CD4 T cells in advanced AIDS weakens CTL selection pressure, limiting new epitope mutations and permitting the reversion of sites conferring high fitness costs. This change in CTL selection pressure may vary according to HLA class I genotype and HIV protein.

5.3 Aim

To study whether HLA class I mediated CTL responses continue to impose selection pressure on different HIV proteins in patients with advanced AIDS. To identify immune selection in the HIV genes I studied:

(i) Viral evolution and immune escape in the cognate CTL epitopes

(ii) *Ex vivo* interferon gamma (IFNγ) CD8 T cell ELISPOT responses against limited cognate epitopes

5.4 Results

5.4.1 Structure and characteristics of the two South African cohorts

Patients were recruited from two South African cohorts, one in Bloemfontein (BFN), Free State and one in Durban (DBN), KwaZulu-Natal. In summary, each of the cohorts consisted of adult patients recruited from neighbouring provinces in the central east region of South Africa, who are chronically infected with HIV-1 subtype C and have not received ARV intervention previously. Both cohorts have been

described in detail elsewhere (Huang, et al. 2009, Kiepiela, et al. 2004). For HIV *gag* sequences, 824 patients have both HIV *gag* sequence and high-resolution HLA class I genotype information available for phylogenetic dependency network analysis, although 21 of the patients have no CD4 count information (Table 5.1). Similarly, 687 and 676 patients had HIV *pol* or *nef* sequences with HLA class I genotype information available, respectively. For 914 patients the CD4 T cell counts were known, for 843 viral load data were available. For 822 patients both CD4 and VL were known. In addition to sequencing of the *pol* gene described in Chapter 4, HIV genes, *gag* (covering genes encoding p17 and p24 proteins, 1080 nucleotides) and *nef* (entire gene, 618 nucleotides) were also sequenced using primers and methods described in Chapter 2.

The Bloemfontein cohort comprises 1491 drug-naïve patients at all stages of HIV infection recruited through the South African government ARV program. HLA class I genotype information was available from 260 participants: 96 in the high (CD4 count > 500 cells/μl) and 164 in the low (CD4 count <100 cells/μl) strata (Table 5.1). The samples were used for the HLA-association viral sequence study and to make recombinant viral strains for the analysis of fitness. To increase the power of the HLA-association sequencing analysis we incorporated HIV viral sequences from the Durban cohort, which comprised 656 adults, naive to antiretroviral therapy. Of these, 219 were asymptomatic women identified through antenatal clinics. The remaining 437 subjects were recruited from outpatient clinics. Clinically, the mean CD4 cell counts and median viral load of the Durban cohort differs significantly from that of Bloemfontein cohort (Figure 5.1a and 5.1b). The Bloemfontein cohort recruited patients from ARV dispensing clinics, who often present at more advanced or

symptomatic stages of the disease in contrast to antenatal and outpatient clinics.

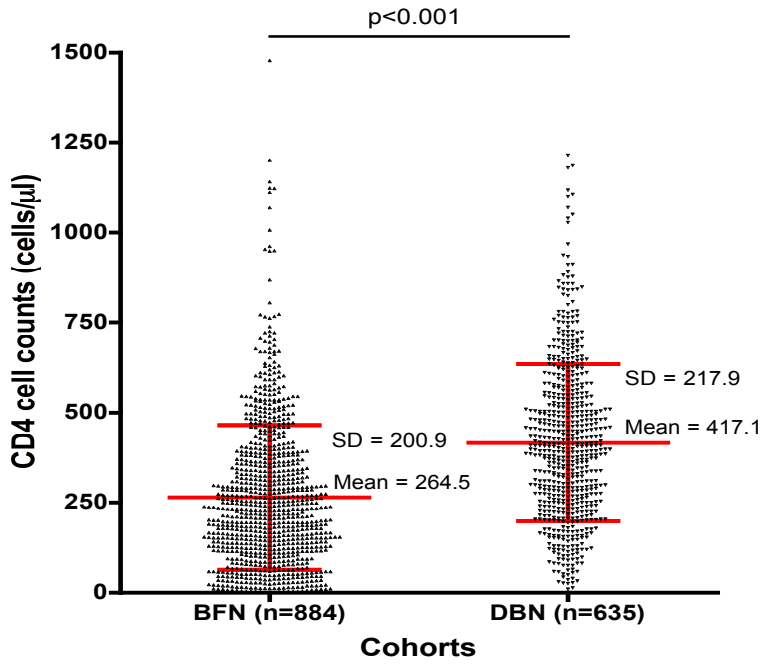
Despite differences in gender constitution and recruiting clinics, the phylogenetic analysis conducted on HIV *pol* (in Chapter 4.3.2, Figure 4.1 and 4.2) and HIV *gag* and *nef* (data not shown) confirmed comparable molecular epidemiology of the two cohorts, justify the merging of the two datasets. In collaboration with Microsoft (Seattle), have conducted an independent assessment of the phylogeny of the two cohorts (using methods described, (Carlson, et al. 2008)), and have confirmed the cohorts are suitable for a combined analysis using phylogenetic dependency networks. There are no distinct local epidemics in either of the cohorts, both cohorts are infected by HIV-1 subtype C, and cohort samples provided are representative of the molecular epidemiology seen elsewhere in the national and regional databases (Chapter 4.3.2). Furthermore, the geographical approximation between the two provinces and the population ethnic constitution (and therefore also HLA class I allele frequencies) are suitable for data combination, as discussed in Chapter 2.1.2 and Figure 2.1.

Table 5.1: Breakdown of cohorts structure and stratification

Cohorts	Stratifications	CD4 cell counts (cells/µl)					Subtotal of patients with CD4 information	CD4 count unavailable	Total Patients
		<100	101-200	201-400	401-500	>500			
BFN	Genotyped for HLA	164	not done	not done	not done	96	260	607	1491
	HLA & gag known	128	not done	not done	not done	78	206		
	HLA & pol known	162	not done	not done	not done	86	248		
	HLA & nef known	160	not done	not done	not done	90	250		
	Viral replication assayed	88	not done	18	not done	42	148		
	Total (with plasma sample)	183	213	308	71	109	884	607	1491
DBN	Genotyped for HLA	32	66	220	114	203	635	21	656
	HLA & gag known	32	63	204	108	190	597	21	618
	HLA & pol known	24	51	156	80	115	426	13	439
	HLA & nef known	21	48	149	80	124	422	4	426
	with ELISpot conducted	38	63	282	143	249	775	not included	775
Combined cohorts (BFN+DBN)	Genotyped for HLA	196	66	220	114	299	895	21	916
	HLA & gag known	160	63	204	108	268	803	21	824
	HLA & pol known	186	51	156	80	201	674	13	687
	HLA & nef known	181	48	149	80	214	672	4	676

(BFN = Bloemfontein cohort, DBN = Durban cohort)

a)



b)

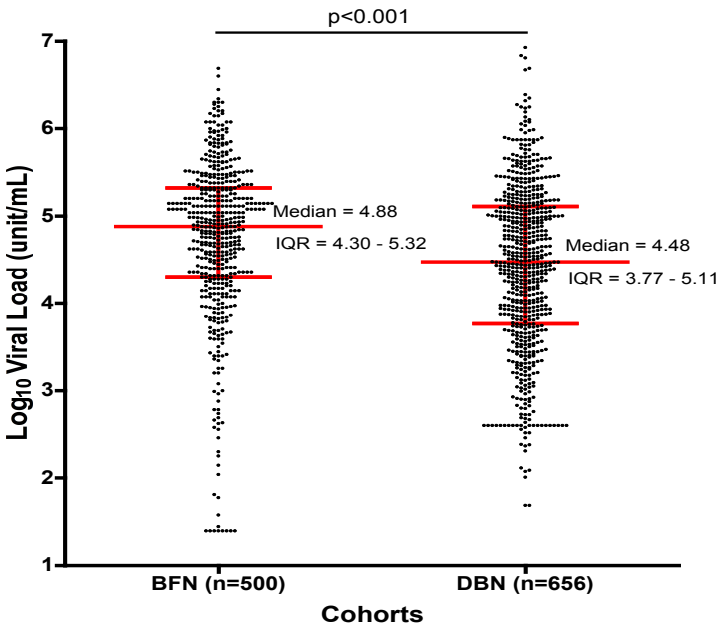


Figure 5.1: CD4 cell counts and Log_{10} viral load distribution of the two cohorts

a) The distribution (including mean and standard deviation, SD) of the CD4 cell counts of Bloemfontein (BFN) and Durban (DBN) cohorts are displayed. The p-value was calculated using unpaired t-test, two tailed. b) The distribution (including median and interquartile range, IQR) of the Log_{10} viral load of the two cohorts are displayed. The p-value was calculated using Mann-Whitney test, two tailed.

5.4.2 HLA class I allele frequencies distribute unevenly at the extremes of CD4 count strata

The cross-sectional distribution of HLA-A, B and C alleles of the combined cohorts is analysed in the context of the CD4 T cell counts (Table 5.2). Significant differences are noted in the host distribution of four HLA-A, six HLA-B and four HLA-C alleles across patients at either extreme of CD4 T cell counts. In particular, one of the HLA-A alleles (HLA-A*7401) and two of the HLA-B alleles (HLA-B*5703 and B*8101) appear to be enriched within the more favourable HIV immunological stages (higher CD4 T cell counts, with odds ratio >1). The opposite is also observed, with three HLA-A alleles (HLA-A*3001, A*6601 and A*8001), four HLA-B alleles (HLA-B*1801, B*4202, B*4701, and B*5802) and four HLA-C alleles (HLA-C*0210, C*0602, C*1601 and C*1701) associating with more advanced AIDS prevalence. It is worth noting that only the HLA class I molecules distributing significantly unevenly between the extremes of CD4 cell counts strata are shown in Table 5.2. In total, there are 40 HLA-A alleles, 57 HLA-B alleles and 34 HLA-C alleles that are present at ≥0.5% prevalence within the two cohorts.

Table 5.2: List of HLA class I alleles that distribute differentially across extremes of CD4 cell counts strata

		CD4 cell count ≤100 cells/µl (n=196)	CD4 cell count ≥500 cells/µl (n=300)	Total patients in two cohorts (n=916)	Odds ratio (of protection), and 95% confidence interval		P VALUE: (Fisher's exact test)
HLA-A	3001	23.47%	15.33%	17.58%	0.59	(0.37 - 0.93)	0.025
	6601	9.69%	3.67%	6.33%	0.35	(0.16 - 0.76)	0.007
	7401	5.10%	11.67%	9.17%	2.46	(1.19 - 5.09)	0.016
	8001	2.04%	0.00%	1.31%	0.00	(0 - 1.33)	0.024
HLA-B	1801	5.10%	1.67%	2.73%	0.32	(0.11 - 0.94)	0.034
	4202	4.59%	0.00%	1.97%	0.00	(0 - 0.57)	<0.001
	4701	2.55%	0.00%	0.66%	0.00	(0 - 1.05)	0.009
	5703	1.53%	10.00%	5.46%	7.15	(2.15 - 23.77)	<0.001
	5802	31.63%	16.67%	23.14%	0.43	(0.28 - 0.66)	<0.001
	8101	7.65%	15.33%	10.81%	2.19	(1.18 - 4.04)	0.012
HLA-C	0210	15.82%	5.67%	5.24%	0.32	(0.17 - 0.6)	<0.001
	0602	39.29%	25.33%	30.46%	0.52	(0.36 - 0.77)	0.001
	1601	11.22%	6.00%	9.39%	0.50	(0.26 - 0.97)	0.043
	1701	28.06%	19.00%	20.74%	0.60	(0.39 - 0.92)	0.021

5.4.3 Identification of HLA class I linked polymorphisms in HIV-1 gag in different CD4 cell counts strata

To identify evidence of differences in selection pressure according to HLA Class I in patients with high and low CD4 cell counts, we first carried out an HLA- association study in HIV *gag*. We accounted for potential confounding effects such as phylogeny, founder effect and multiple comparisons by using phylogenetic dependency networks (Carlson, et al. 2008) - an approach which reports associations between specific viral polymorphisms and HLA Class I alleles according to p value, q value (an estimate of false positive error) and a phylogenetically-adjusted odds ratio.

Initially, we analysed 824 HIV-1 gag sequences from the 'total' cohort with known HLA information (Table 5.1) and then we carried out sub-group analyses on patients with either high (>500 cells/μl, n=268) or low (<100 cells/μl, n=160) CD4 cell counts. A total of 114 HLA class I linked polymorphisms were observed (list not shown), 40 of which showed significant associations in either or both extreme of CD4 cell counts strata (Table 5.3). Within the 40 HLA-associated polymorphisms related to CD4 cell counts stratification, 31 amino acid sites (of 363 amino acids included for HIV Gag analysis, 17 sites in gag p17 and 14 sites in gag p24) were polymorphic in association with 8 HLA-A, 13 HLA-B, and 5 HLA-C Class I alleles ($q<0.20$, Table 5.3). Polymorphisms in 5 of the sites in gag p24, and 4 in gag p17 were associated with more than one HLA class I locus or allele. The analysis also identified 14 associations where polymorphic sites revert to wild type amino acid in the HLA mismatched hosts ($q<0.20$, column titled "Is reversion seen in HLA mismatched host? (R)") (Matthews, et al. 2008). A total of 13 reverting sites were observed in 31 amino acid sites across gag p24 and p17, relating to 2 HLA-A, 7 HLA-B and 3 HLA-C alleles.

HLA class I linked polymorphisms were compared with previously reported CTL epitopes (column "Is the amino acid part of a known cognate epitope (E), or flanking (F) region (within 5 amino acid of epitope)", Table 5.3), defined through inclusion in the optimal epitope 'A-list' of the Los Alamos Database or characterisation of gamma interferon ELISPOT responses on the Durban cohort (Kiepiela, et al. 2007). Fourteen HLA class I linked polymorphisms were in epitopes or flanking regions restricted by 1 HLA-A, 9 HLA-B and 1 HLA-C alleles across 11 amino acid sites.

Next, the q values for the statistically significant associations for the whole cohort, the 'high' CD4 count subgroup (>500 cells/μl) and 'low' CD4 count subgroup (<100 cells/μl) are shown. Interestingly, phylogenetic dependency network analysis of cohorts according to CD4 count (entire cohort, high and low CD4 cell count, table 5.3) identified different lists of significant HLA class I linked associations. Although some of the missing associations between different strata may be a result of limitation in size and therefore power of statistical analysis, certain associations becomes more apparent within specific strata (but not the entire population).

Lastly, the $\log_2$-adjusted odds ratios for the 'high' and 'low' groups are reported as a measure of strength of the association. In the final column, the p value is shown, when there is a significant difference between the two odds ratios. This will be explain in further detail in the next section (5.4.4).

Table 5.3: HLA class I-associated polymorphisms in HIV-1 clade C Gag protein in different CD4 count strata

HIV-1 protein	HLA Class I	Population majority consensus amino acid at associated site	HXB2 site and direction of HLA associated polymorphism	Is the amino acid part of known epitope (E), or within flanking (F) region (within 5 amino acid)	Is there evidence for reversion in HLA mismatched host (R)?	q-value of HLA-related association in each strata			Phylogentically corrected Log_2 Odds Ratio (pLOR) towards "escape"		
						Entire population	High CD4 cell count strata (>500 cells/µl)	Low CD4 cell count strata (<100 cells/µl)	High CD4 cell count strata (>500 cells/µl)	Low CD4 cell count strata (<100 cells/µl)	p-value (likelihood ratio test between pLOR of high and low CD4 strata)
Gag p17	A29	I	I7V				0.189				
Gag p17	A74	K	K12N/Q		R	0.000	0.000		18.83	- (infinity)	0.0070
Gag p17	B5801	K	K12Q		R		0.147				
Gag p17	A3303	K	K15S			0.026	0.062				
Gag p17	A6801	K	T15	F		0.131	0.086				
Gag p17	B45	H	H28R					0.058	- (infinity)	15.05	0.0005
Gag p17	C17	H	H28Q			0.000	0.183	0.017			
Gag p17	B4201	M	M30	E		0.004	0.122		14.14	9.44	0.0443
Gag p17	B0801	A	A45S		R			0.128	- (infinity)	5.29	0.0051
Gag p17	A6802	E	E55D					0.015	-6.35	18.13	0.0009
Gag p17	A3402	K	Q62				0.149		(infinity)	-10.24	0.0130
Gag p17	A3001	K	K62S				0.183		10.89	- (infinity)	0.0151
Gag p17	A30	Q	Q65H		R		0.086		9.27	-11.37	0.0086
Gag p17	B8101	Q	K69				0.163		(infinity)	-0.21	0.0027
Gag p17	A26	K	K76R				0.149				
Gag p17	B15	Y	79Y		R		0.188		14.22	-1.66	0.0193
Gag p17	C04	K	K91N		R			0.016	15.78	4.21	0.0004
Gag p17	B08	Q	Q112K				0.159				
Gag p17	B1402	K	K114				0.069		7.73	- (infinity)	0.0498
Gag p17	B1801	T	T115				0.146				
Gag p17	A2902	Y	Y132F				0.188				
Gag p24	B57	A	A146P	E		0.000	0.000	0.048			
Gag p24	C17	A	P146		R		0.102				
Gag p24	B1510	A	A146	E		0.008	0.153				
Gag p24	C0304	I	147I	E			0.102		(infinity)	1.41	0.0209
Gag p24	B57	I	I147L/M	E	R	0.000	0.000				
Gag p24	C17	Q	Q182					0.017	0.79	4.18	0.0019
Gag p24	B8101	Q	Q182S	E		0.000	0.000	0.017	13.97	16.79	0.0349
Gag p24	B8101	T	T186S	E	R	0.000		0.000	18.38	7.25	0.0397
Gag p24	C0210	V	V218		R			0.056			
Gag p24	B5801	T	T242N	E		0.000	0.000	0.000			
Gag p24	B57	T	T242N	E	R	0.000	0.000	0.000			
Gag p24	A74	I	I247		R		0.087				
Gag p24	C04	S	N252				0.189				
Gag p24	B35	D	D260E	E		0.000	0.000				
Gag p24	B1401	K	K302R	E	R	0.000	0.015				
Gag p24	B44	D	D312E	E	R	0.000	0.183				
Gag p24	B0702	D	D319E				0.086				
Gag p24	C08	T	T332A			0.131		0.015			
Gag p24	B0702	G	S357G	E		0.000	0.001	0.000			

5.4.4 Variation in the strength of associations across HLA class I linked polymorphisms in Gag at different CD4 stratifications is suggestive of changing selection pressure

Table 5.3, therefore, allows us to see which associations between HLA class I and mutations in Gag are stronger or weaker in patients with advanced disease compared with patients with early infections. In the 40 significant associations between a viral polymorphism and an HLA Class I allele in either, or both of, the 'high' or 'low' CD4 count subgroups for the HIV-1 *gag* genes, 16 had significantly different $\log_2$-adjusted odds ratios in the two subgroups (Figure 5.2a). Here, it is worth noting that the positive and negative signs in these columns describe the direction of associations, whether it is towards or away from the HLA-specific wild type status (which is not always identical to the population consensus, see "population majority consensus amino acid status at associated HXB2 site" column of Table 5.3).

We identified 11 associations at 10 amino acid sites in which the strength of the association was weaker in patients with low CD4 cell counts compared to the high CD4 count group (Figure 5.2a). There were 4 associations at 5 sites in which the association was stronger in the low CD4 count group compared with the high group. The weakening of certain associations suggests an enrichment of wild-type sequences in patients with low CD4 counts. Why might HLA associations become less strong as disease progresses? One possible explanation is that as CTL pressure weakens due to immune exhaustion and loss of breadth, escape mutations with replicative fitness costs revert to fitter wild type variants. This reversion back to wild-type, resulting in a reduction in the observed $\log_2$-adjusted odds ratio has been dubbed as 'immune relaxation'. Thus, low-cost mutations accumulate over the course of infection, leading

to higher prevalence in Nef and Pol, whereas high-cost mutations begin reverting, leading to a lowered prevalence in Gag. An alternative explanation is that only patients with wild-type virus progress to AIDS and therefore they are enriched in the low CD4 count strata. It is, of course, also possible that both processes are occurring.

One association demonstrating significantly weakening at low CD4 cell counts was the T186S mutation in the TPQDLNTML (TL9) epitope, restricted by HLA-B*8101, with a decrease in $\log_2$-adjusted odds ratio from 18.38 (CD4 count >500 cells/μl) to 7.25 (CD4 count <100 cells/μl). The p-value for likelihood ratio test is 0.040. The fitness impact of HLA-B*8101 related T186S mutation is further suggested by its tendency to revert upon transmission into non-HLA-B*8101 hosts. Another two HLA-B*8101 related polymorphisms are also listed in Table 5.3 and Figure 5.2a (at position K69 and Q182S). However, neither of these reverts in HLA mismatched hosts (Table 5.3). The Q182S mutation is located in the same epitope (TL9) as T186S but displays a remarkably preserved odds ratio as the CD4 cell counts decreases – the $\log_2$-adjusted pLOR increases from 13.97 to 16.79. There are no significant linkage disequilibrium or clear pattern of co-evolution noticed between Q182S and T186S. We hypothesis the two mutations are involved in different pathways that mediate escapes from the CD8 T cell responses, and investigate their effects against CD8 T cell ELISPOT and viral replication fitness in later section, accordingly.

In contrast, another HLA class I linked mutation under significant and strong selection pressure is the T242N mutation within the B*57 and B*5801 restricted epitope TSTLQEQIGW (TW10). Selection was consistently significant across all strata of CD4 cell counts and there was evidence for reversion in HLA negative hosts

(Table 5.3). The log_2-adjusted odds ratio for the association with the T242N mutation was 25.36 and 23.20 in high and low CD4 cell counts strata respectively (p=0.121, not significantly different). In particular, none of the six HLA-B*57 hosts harboured virus with wild type variant (T242) in low CD4 cell counts strata.

These data are compatible with different strengths of selection in late disease for different HLA Class I alleles. One might expect that changes in viral fitness with disease progression, might be specific to certain HLA Class I alleles or associated key mutations. We therefore constructed recombinant viruses from patients with low and high CD4 cell counts to explore further the impact of sequence variation on viral fitness (see Chapter 6).

5.4.5 Variation in the strength of associations across HLA class I linked polymorphism in Pol and Nef according to CD4 stratification

In order to gain a wider understanding of the changes in CTL selection against HIV across different CD4 cell counts strata, we have extended the phylogenetic dependency networks and pLOR analyses to cover two other important HIV proteins – Pol (Table 5.4 and Figure 5.2b) and Nef (Table 5.5 and Figure 5.2c).

We report 91 and 20 associations between a viral polymorphism and an HLA Class I allele in either the 'high' or 'low' CD4 count subgroups for the HIV-1 *pol* and *nef* genes, respectively. Of these, 28 and 7 associations in HIV-1 *pol* and *nef*, respectively, had significantly different log_2-adjusted odds ratios in the two subgroups (Figure 5.2b and Figure 5.2c, respectively). Figure 5.2 shows that in *gag* there were five (31.3%) associations that were more prevalent in the 'low CD4' group, compared with 11 that

were less prevalent. For *pol* there were 24 (85.7%) associations that were stronger and only 4 that were less common in advanced disease. Finally, for *nef*, all 7 (100%) associations had greater $\log_2$-adjusted odds ratios in the low CD4 count patients compared with the 'high CD4' group. These data show that, where we could detect a difference, most associations became increasingly prevalent at lower CD4 counts (36/51; 70.6%), suggestive of continued selection pressure, but that certain associations – particularly in HIV-1 *gag* – became weaker.

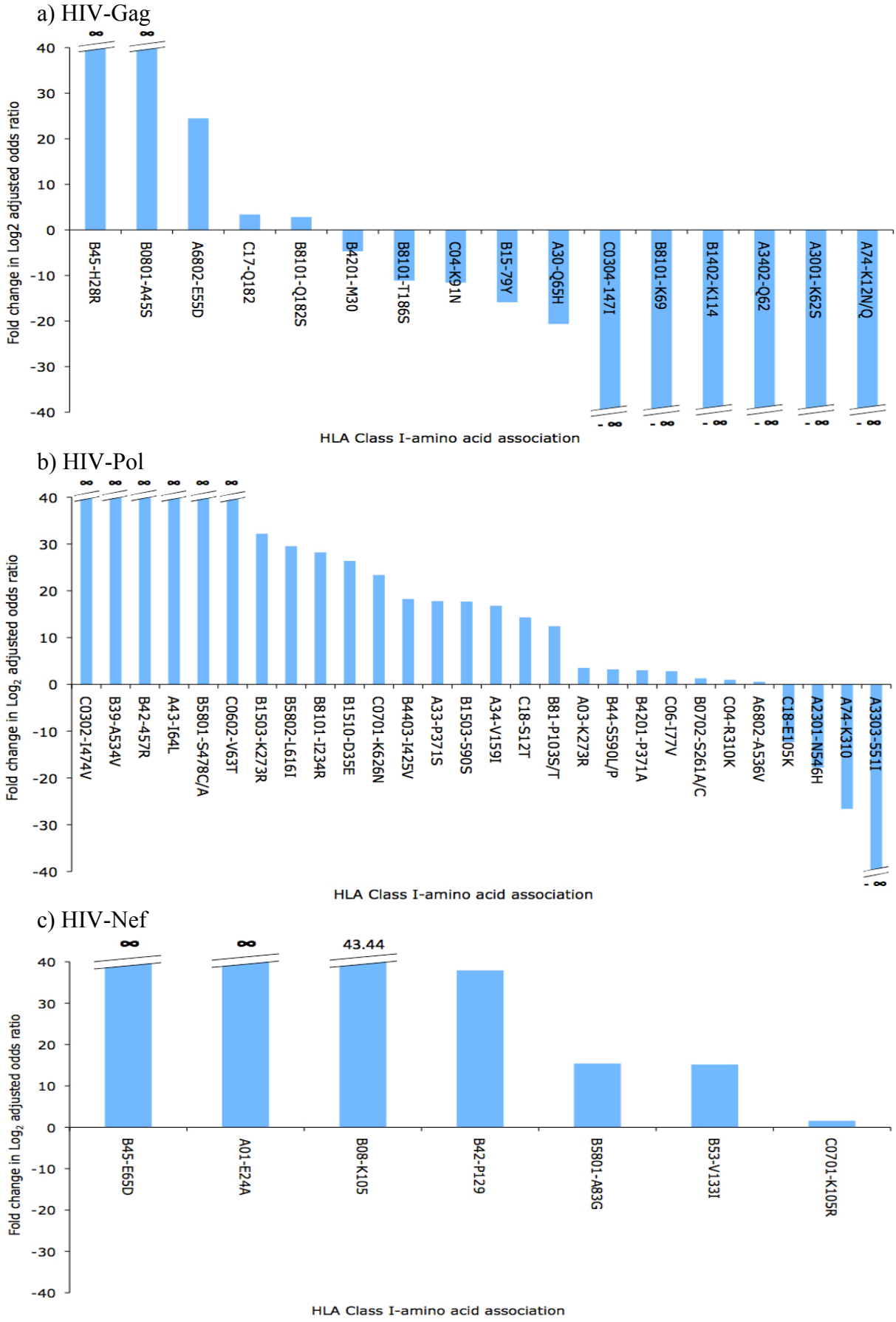


Figure 5.2: Changes in associations between HLA Class I alleles and HIV amino acid polymorphisms for patiehts with CD4 cell counts <100 cells/µl compared with patients with CD4 counts >500 cells/µl

The change in the strength of associations between HLA Class I alleles and specific HIV amino acid polymorphisms in patients with high and low CD4 cell counts is shown for HIV-1 (a.) Gag, (b.) Pol and (c.) Nef. Change in strength of the association is reported on the Y-axis as a log(2)-adjusted value. Associations are recorded on the X-axis according to the restricting HLA Class I allele and the specific viral polymorphism – e.g. B45-H28R in Gag means that in the presence of HLA B*45, Gag amino acid 28 mutates from a Histidine (H) to an Arginine (R). In this case there is at least a 40-fold change in the log(2)-adjusted odds ratio for this association in patients with CD4 counts <100 cells/μl compared with less progressed patients with CD4 cell counts >500 cells/μl. (All fold changes are capped at a value of 40).

Table 5.4: HLA class I-associated polymorphisms in HIV Pol in different CD4 count strata

HIV-1 protein	HLA Class I	Population majority consensus amino acid at associated site	HXB2 site and direction of HLA associated polymorphism	Is the amino acid part of known epitope (E), or within flanking (F) region (within 5 amino acid)	Is there evidence for reversion in HLA mismatched host (R)?	q-value of HLA-related association in each strata			Phylogentically corrected Log_2 Odds Ratio (pLOR) towards "escape"		
						Entire population	High CD4 cell count strata (>500 cells/µl)	Low CD4 cell count strata (<100 cells/µl)	High CD4 cell count strata (>500 cells/µl)	Low CD4 cell count strata (<100 cells/µl)	p-value (likelihood ratio test between pLOR of high and low CD4 strata)
Protease	C0401	L	L10		R		0.178				
Protease	C18	S	S12T					0.010	2.13	16.45	0.0005
Protease	B5802	S	T12S			0.000		0.007			
Protease	C06	S	12S					0.007			
Protease	B1801	I	I19T					0.162			
Protease	B1510	E	D35E				0.155		-5.82	20.55	0.0015
Protease	B44	E	E35D	E		0.000		0.000			
Protease	A6601	N	N37			0.167		0.121			
Protease	C0602	L	V63T	E		0.192		0.190	13.30	(infinity)	0.0044
Protease	C16	L	P63L		R			0.013			
Protease	C18	L	P63S		R	0.000		0.064			
Protease	C08	L	63L					0.100			
Protease	A43	I	I64L		R	0.167		0.097	- (infinity)	20.66	0.0477
Protease	C06	V	I77V			0.004		0.013	3.60	6.41	0.0129
Protease	A66	V	V77I					0.196			
RT	B81	P	P103S/T	E		0.000	0.155	0.000	10.43	22.85	0.0092
RT	C18	E	E105K			0.001	0.178		10.97	-1.39	0.0207
RT	B81	E	E105	E		0.000	0.178				
RT	B18	K	K119R		R			0.073			
RT	B15	E	E152D					0.035			
RT	B4101	E	E152D			0.141		0.080			
RT	A34	V	V159I					0.121	0.54	17.36	0.0155
RT	B35	E	E221K	E		0.003	0.178				
RT	C02	G	S222			0.002		0.015			
RT	B8101	I	I234R					0.131	-5.66	22.55	0.0379
RT	B4202	I	I234V					0.036			
RT	B0702	S	S261A/C	E		0.000		0.041	15.26	16.55	0.0348
RT	C0702	S	S261					0.029			
RT	B07	T	T264I	E		0.000		0.000			
RT	A03	K	K265R	E	R	0.005		0.121			
RT	C0702	E	E268					0.010			
RT	B1503	K	K273R	E		0.002		0.003	-14.09	18.11	0.0302
RT	A03	K	K273R	F		0.000		0.035	15.22	18.75	0.0444
RT	A80	E	E276D			0.171		0.092			
RT	B4403	E	E306K/N/R	E	R	0.000		0.001			
RT	B1801	E	E306			0.167		0.131			
RT	C06	E	D306					0.143			
RT	C04	K	R310K		R			0.072	16.98	17.95	0.0001
RT	A74	K	K310		R		0.178		16.39	-10.22	0.0059
RT	A80	D	D349					0.195			
RT	A30	I	I356		R			0.121			
RT	A33	P	P371S					0.100	-12.13	5.66	0.0196
RT	B4201	P	P371A	E				0.000	12.06	15.09	0.0040
RT	C1701	P	P371A/S			0.000	0.054				
RT	B42	I	I373V	E	R	0.000		0.037			
RT	A3001	K	K374	F				0.064			
RT	A03	R	K376R	E		0.000		0.006			
RT	A3001	K	K380R			0.000		0.005			
RT	B57	A	A385V				0.178				
RT	C0210	A	A385T				0.198				
RT	B15	A	A385V		R			0.099			
RT	B45	E	E396K					0.153			
RT	C04	D	E423D		R	0.057		0.143			
RT	B4403	I	I425V					0.003	2.39	20.64	0.0259
RT	C03	M	R456				0.178				
RT	B42	R	457R					0.110	-20.83	(infinity)	0.0039
RT	C06	R	R457K		R	0.000		0.019			
RT	A6601	R	R457K		R			0.097			
RT	B5802	T	458T		R			0.103			
RT	C03	A	A470		R			0.150			
RT	C0302	I	I474V		R	0.165		0.000	- (infinity)	(infinity)	0.0015
RT	B15	L	L476I		R			0.128			
RT	B5801	S	S478C/A	E	R	0.004		0.001	- (infinity)	16.17	0.0001
RT	A32	E	E498	E				0.080			
RT	A3201	T	T499I	E		0.001		0.010			
RT	A29	T	499T					0.081			
RT	B13	D	D503N			0.007	0.051				
RT	B39	A	A534V					0.033	- (infinity)	(infinity)	0.0018
RT	B5802	A	I534					0.196			
RT	A6802	A	A536V	E	R	0.007		0.053	15.20	15.75	0.0177
RT	A6601	A	A536V	F		0.001	0.178	0.038			
RT	A2301	N	N546H		R	0.073	0.165		12.30	-5.48	0.0030
RT	B4403	E	E548D					0.025			
RT	A3303	I	551I				0.198		(infinity)	- (infinity)	0.0000
RT	B5802	I	I551A		R			0.072			
RT	B4202	V	V566I					0.143			
RT	B08	V	V566I					0.145			
RT	B08	S	S567T			0.000		0.003			
RT	B08	L	L568					0.190			
RT	A3402	T	T569A		R	0.082	0.155				
RT	B18	Q	Q582H	E		0.001		0.073			
RT	B1503	S	590S	F	R			0.150	-12.70	5.01	0.0196
RT	B44	S	S590L/P	E	R	0.000		0.000	0.74	3.95	0.0408
RT	B1401	S	S590L	F			0.178				
RT	B4403	E	E591	E		0.018		0.011			
RT	B3910	I	I594V		R	0.000		0.005			
RT	B4202	K	K611N		R	0.071		0.120			
RT	B5802	L	L616I					0.009	-11.95	17.59	0.0084
RT	C0701	K	K626N					0.188	-5.73	17.64	0.0045
Integrase	B4403	E	E669A	E	R	0.000		0.011			
Integrase	B4403	E	E670D	E	R	0.000		0.000			

Table 5.5: HLA class I-associated polymorphisms in HIV-1 clade C Nef protein in different CD4 count strata

HIV-1 protein	HLA Class I	Population majority consensus amino acid at associated site	HXB2 site and direction of HLA associated polymorphism	Is the amino acid part of known epitope (E), or within flanking (F) region (within 5 amino acid)	Is there evidence for reversion in HLA mismatched host (R)?	q-value of HLA-related association in each strata			Phylogentically corrected Log_2 Odds Ratio (pLOR) towards "escape"		
						Entire population	High CD4 cell count strata (>500 cells/µl)	Low CD4 cell count strata (<100 cells/µl)	High CD4 cell count strata (>500 cells/µl)	Low CD4 cell count strata (<100 cells/µl)	p-value (likelihood ratio test between pLOR of high and low CD4 strata)
Nef	B07	I	I10R	F				0.178			
Nef	A01	E	E24A					0.172	- (infinity)	15.19	0.0024
Nef	B42	A	A26P		R	0.013		0.053			
Nef	A6802	A	A53E					0.134			
Nef	A0201	C	C55	F				0.033			
Nef	B45	E	E65D	E				0.150	- (infinity)	20.98	0.0498
Nef	B45	E	E66D	E		0.000		0.178			
Nef	C0702	R	R71K		R	0.000	0.110	0.000			
Nef	B8101	L	L76T/V	E	R	0.000		0.018			
Nef	B5801	A	A83G	E		0.159		0.003	11.33	26.74	0.0008
Nef	B4403	Y	Y102H	F		0.000		0.000			
Nef	B08	K	K105	E				0.178	-23.73	19.71	0.0001
Nef	C0701	K	K105R	E	R	0.000	0.124	0.000	21.77	23.34	0.0290
Nef	B4403	E	E108D	E		0.000		0.011			
Nef	B18	E	E108D	E		0.000		0.134			
Nef	C0602	Q	Q125H	E	R	0.149		0.030			
Nef	B42	P	P129	E				0.150	-15.22	22.70	0.0036
Nef	B53	V	V133I	F				0.078	6.21	21.38	0.0240
Nef	A23	F	F143Y	F		0.000		0.000			
Nef	C06	S	S187		R			0.104			

5.4.6 Changes in the absolute breadth of CTL mediated INFγ ELISPOT responses in advanced AIDS patients

This work was carried out in collaboration with Professor Philip Goulder as cryopreserved T cells were not established from the Bloemfontein cohort at the time of this study.

In order to measure the T cell-imposed selection pressure acting on HIV at different stages of the infection, ELISPOT assays using overlapping peptides (OLPs) covering the entire HIV genome were analysed from 775 patients from the Durban cohort. Results were stratified according to CD4 cell counts and are presented according to the number of OLPs seen per patient for each gene (Figure 5.3a), and the percentage contribution of each protein to all responses in each strata (Figure 5.3b). The figure shows that over the course of infection, there is a general narrowing of the breadth of responses (Figure 5.3a). However, the relative contribution of Gag to the total responses decreases ($p<0.0001$, $r^2=0.055$) in contrast to the proportion of responses to Env ($p<0.0001$, $r^2=0.024$) and Vif ($p=0.003$, $r^2=0.011$), which increased as the CD4 count fell. We hypothesized that the reduction in breadth of CD4 T cell responses might result in, and therefore correlate with, a decline in selection pressure, particularly in Gag in light of the relative shift away from Gag targeting.

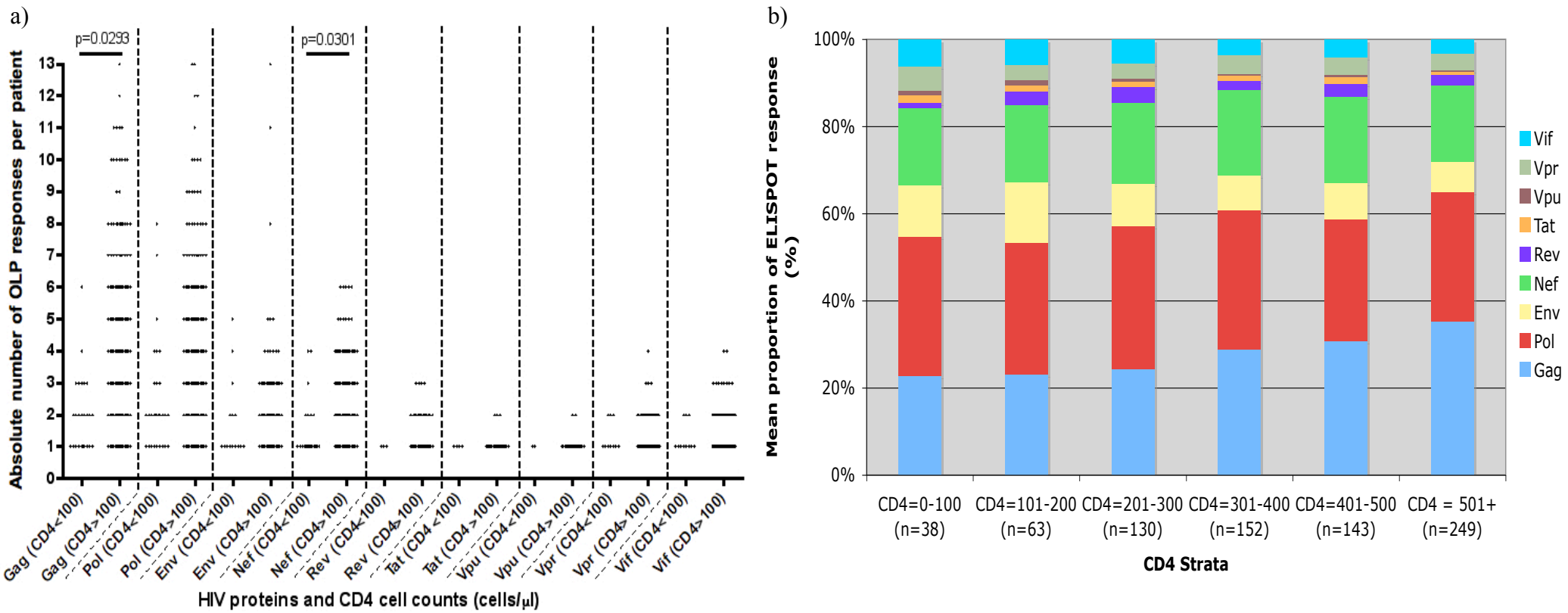


Figure 5.3 IFNγ ELISPOT responses in patients with CD4 counts above and below 100 cells/µl

Recognition of overlapping peptides covering the full HIV genome is shown in patients with CD4 cell counts above and below 100 cells/µl. ELISPOT responses are displayed as (a.) the absolute breadth of responses, reflecting the mean number of overlapping peptides recognised by each patient for each HIV protein and, (b.) the relative breadth, reflecting the proportion of ELISPOT responses targeted against different HIV proteins in patients within different CD4 count strata.

We next investigated whether differences existed between 'protective' HLA-B alleles (such as HLA-B*8101 (n=86, Figure 5.4a, and 5.4b), HLA-B*57 and HLA-B*5801 (n=136, Figure 5.4c and 5.4d)) and 'detrimental' alleles (such as HLA-B*5802, n=144, Figure 5.4e and 5.4f). Both the absolute and relative CTL breadth are shown for each of the HLA class I alleles. HLA-B*8101 demonstrated only significant narrowing of the relative breadth against HIV-Gag as CD4 cell counts decreased ($p=0.0141$, $r^2=0.070$), whereas for HLA-B*57 and B*5801 there was significant broadening of the relative breadth against the Env and Vif proteins as the CD4 cell counts decreased ($p=0.0232$, $r^2=0.038$ and $p=0.0444$, $r^2=0.030$, respectively). In contrast, HLA-B*5802 displayed both significant narrowing of relative breadth against HIV-Gag ($p=0.0002$, $r^2=0.092$), and broadening of relative breadth against HIV-Env ($p=0.0000$, $r^2=0.075$), as CD4 cell counts decreased.

Lastly, based on results in Table 5.3 and Figure 5.2, we asked: whether targeting of certain HIV-Gag epitopes by their cognate HLA class I alleles may have differential associations with clinical outcome, according to cross sectional CD4 cell counts or plasma viral load. Here, we chose two well-characterised HLA restricted epitopes: HLA-B*8101 restricted TL9 (in OLP 25, Figure 5.5a and b), and HLA-B*57 and B*5801 restricted TW10 (in OLP33, Figure 5.5c and d).

Amongst the 86 HLA-B*8101 patients with CD4 cell counts available, 60 patients made responses to TL9 epitope, 26 did not. In Figure 5.5a and b, the median viral load and mean CD4 cell counts of HLA-B*8101 patients who make responses to TL9 was 9230 units/mL (IQR = 744.5-78500 units/mL) and 536.7 cells/μl (SD = 266.4 cells/μl), respectively. In contrast, the median viral load and mean CD4 cell counts of

HLA-B*8101 patients who did not make responses to TL9 was 15245 units/mL (IQR = 2945-100900 units/mL, p=0.32, Mann-Whitney test) and 425.6 cells/μl (SD = 202.3 cells/μl, p=0.06, student t-test), respectively. There was no significant difference in the magnitude of the TL9 specific ELISPOT responses in the reactive HLA-B*8101 patients that progressed to CD4 cell counts <200 cells/μl).

In the case of HLA-B*57 and B*5801 positive patients with available CD4 cell counts data (n=136), 34 patients made responses to the TL9 epitope, and 102 did not. In Figure 5.5c and d, the median viral load and mean CD4 cell counts of HLA-B*57 and B*5801 patients who make responses to the TW10 was 3655 units/mL (IQR = 477 – 27946 units/mL) and 467.3 cells/μl (SD = 200.0 cells/μl), respectively. In contrast, the median viral load and mean CD4 cell counts of HLA-B*57 and B*5801 patients who did not make responses to TW10 was 14986 units/mL (IQR = 1580-65800 units/mL, p=0.06, Mann-Whitney test) and 492.6 cells/μl (SD = 236.0 cells/μl, p=0.58, student t-test), respectively. There were not enough HLA-B*57 and B*5801 patients with CD4 cell counts <200 cells/μl to allow an evaluation of the ELISPOT responses amongst TW10 responders that progressed to AIDS.

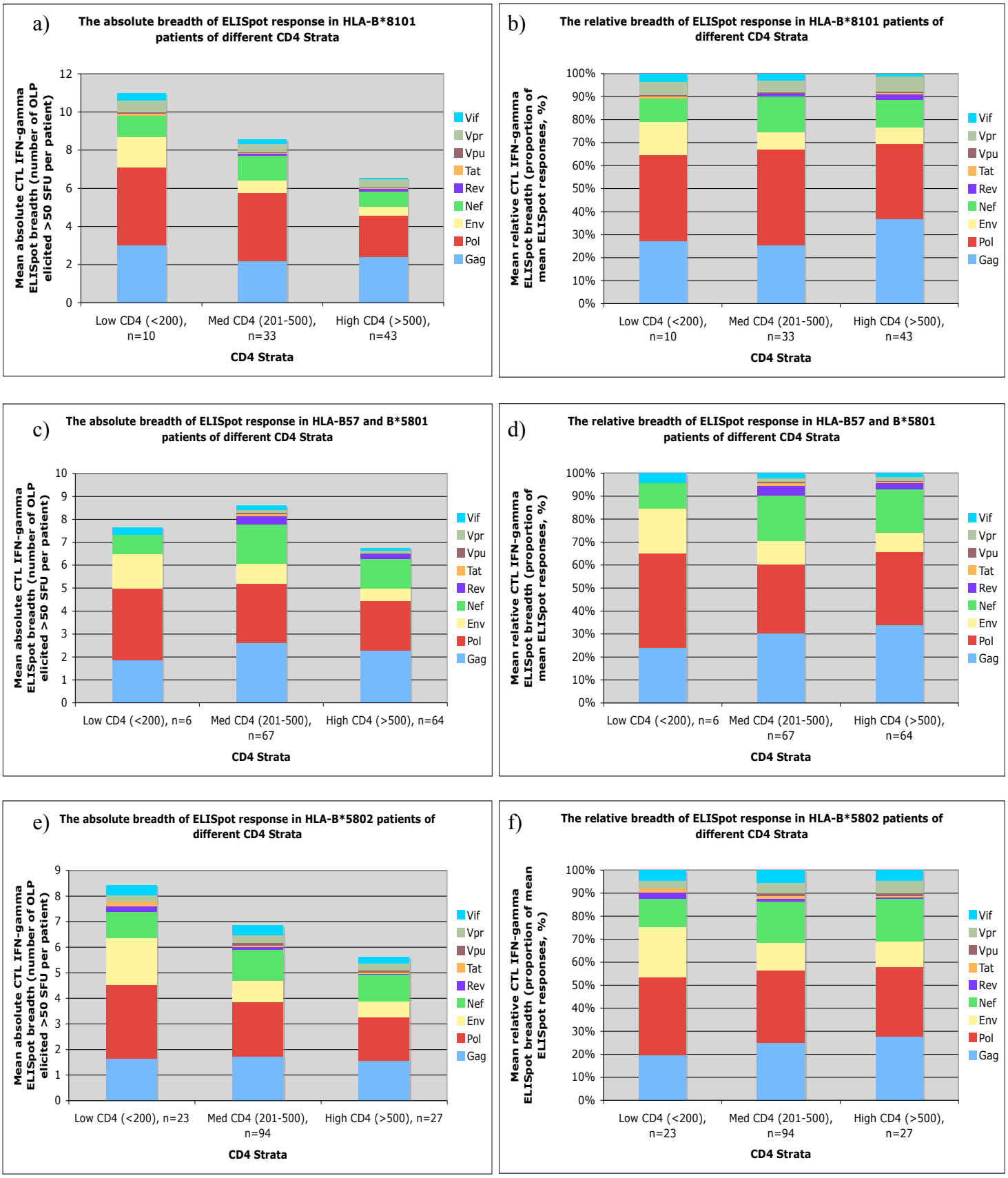
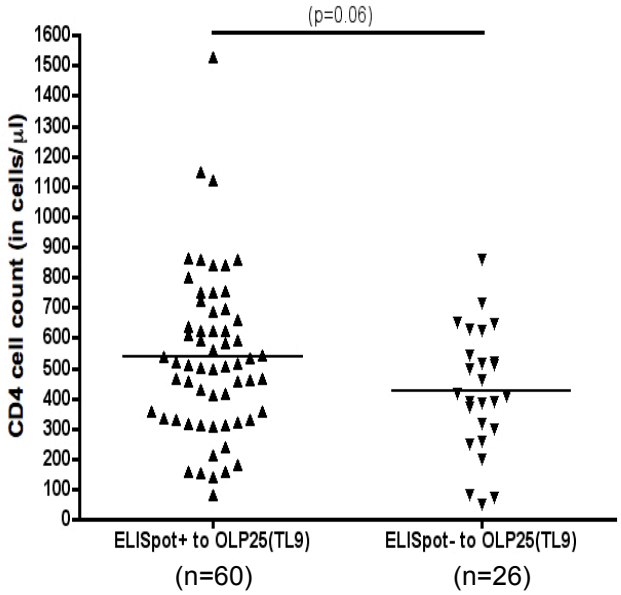


Figure 5.4 The changes in the CTL IFNγ ELISPOT responses against overlapping peptides (OLP) over different CD4 strata in specific HLA class I alleles

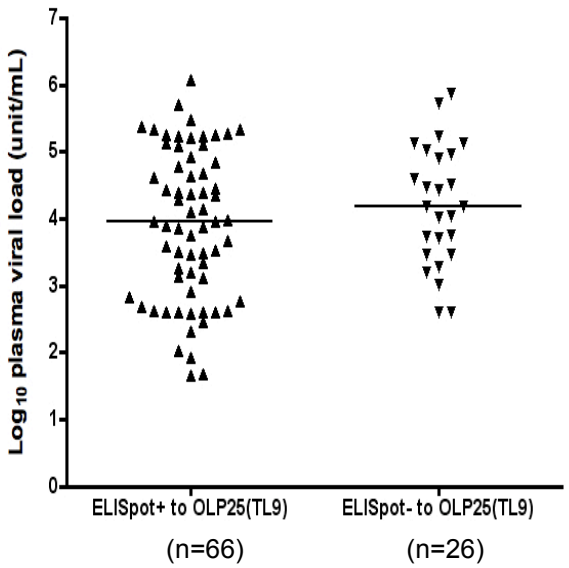
The changes in IFNγ produced in response to overlapping peptides (OLP) from different HIV proteins stratified according to CD4 T cell counts (cells/μl). The 'absolute breadth' of ELISPOT (panel a, c and e) refers to the mean number of OLPs recognised by patients with specific HLA alleles. The 'relative breadth' of ELISPOT (panel b, d, and f) refers to mean proportion of the responses targeted at different HIV proteins. The HIV protein covered by the OLP design includes: Gag, Pol, Env, Nef, Rev, Tat, Vpu, Vpr and Vif.

---

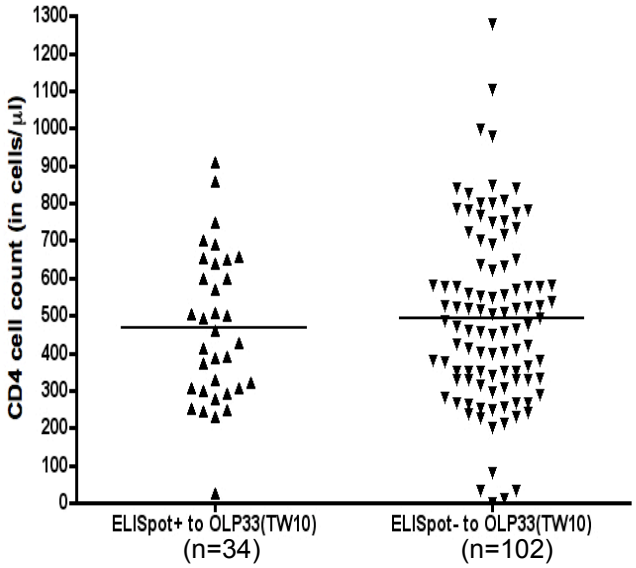
a) Comparison of CD4 cell count between the HLA-B*8101 patients that made INFγ ELISPOT responses against OLP25 (containing Gag TL9)



b) Comparison of plasma viral load between the HLA-B*8101 patients that made INFγ ELISPOT responses against OLP25 (containing Gag TL9)



a) Comparison of CD4 cell count between the HLA-B*57 and B*5801 patients that made INFγ ELISPOT responses against OLP33 (containing Gag TW10)



b) Comparison of plasma viral load between the HLA-B*57 and B*5801 patients that made INFγ ELISPOT responses against OLP33 (containing Gag TW10)

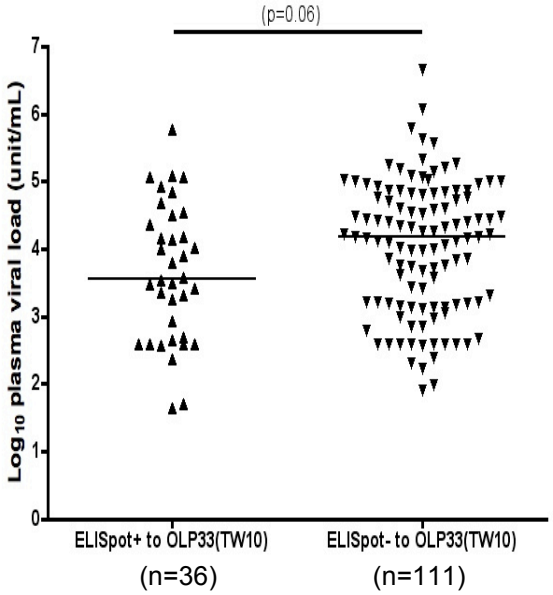


Figure 5.5 Comparing the clinical parameters of specific ELISPOT responses against HIV-Gag epitopes (TL9 and TW10) by their respective cognate HLA class I alleles (HLA-B*8101, and HLA-B*57 and HLA-B*5801)

Panel a and b: The clnical parameters of CD4 cell counts (a) and plasma viral load (b) are compared between HLA-B*8101 patients who respond to wild type OLP 25 (containing the TL9 epitope) in the ELISPOT assay, and HLA-B*8101 patients who do not respond to OLP 25.

Panel c and d: The clnical parameters of CD4 cell counts (c) and plasma viral load (d) are compared between HLA-B*57 and B*5801 patients who respond to the wild type OLP 33 (containing the TW10 epitope) in the ELISPOT assay, and HLA-B*57 and B*5801 patients who do not respond to OLP 33.

The CD4 cell counts columns are displayed with mean bars and compared via student t-test, whereas VL was displayed with median bars and compared via Mann Whitney test.

---

*5.4.7 CD8 ELISPOT assay suggests T186S, but not Q182S, is immune escape mutation of HLA-B*8101 restricted TL9 epitope*

The T186S and Q182S mutations within the TL9 epitope of Gag have shown different patterns of CTL selection mediated by HLA-B*8101 allele (see section 5.4.4). Using the results available from the ELISPOT assay, we investigated the relationship between mutations and strength of ELISPOT responses. Interestingly, within the HLA-B*8101 patients that recognises OLP 25 (containing TL9 epitope), the presence of T186S is associated with reduced ELISPOT magnitude (measured by spot forming units, Figure 5.6, p=0.0425, Mann-Whitney test). In contrast, the mutational status at Q182 did not correlate significantly with the changes in ELISPOT assay. This data suggest that T186S is an escape mutation against CTL selection that recognises OLP25 (containing the TL9 epitope).

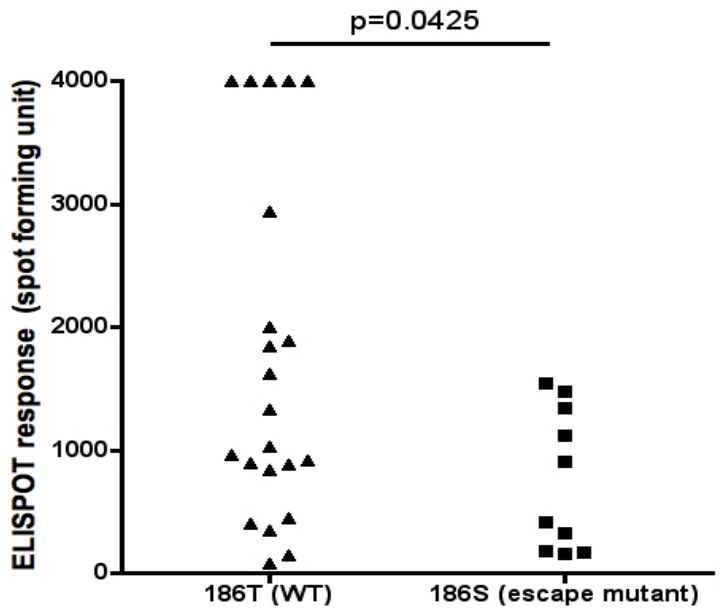


Figure 5.6 Comparison of ELISPOT responses between T186 (WT) and T186S amongst HLA-B*8101 patients

5.5 Discussion

In this study, we have explored the hypothesis that progression to AIDS is associated with altered HLA mediated selection pressure and the raised viral load in AIDS could be associated with a rise in viral replicative capacity. We initially looked for evidence of differential selection in patients with high and low CD4 cell counts using a statistical HLA Class I association study. The analysis was initially conducted only in Bloemfontein (BFN) cohort, and later extended to include the Durban cohort (DBN) to improve statistical power. By combining sequence and HLA data from the Bloemfontein cohort and the Durban cohort, we were able to define two groups of patients with CD4 counts less than 100 cells/μl (n=196) and greater than 500 cells/μl (n=299). These CD4 count strata were defined to represent extremes of HIV-related disease without reducing the power of the analysis. The WHO have defined that antiretroviral therapy should be commenced at a CD4 count of 350 cells/μl. Below 200 cells/μl, opportunistic infections become more common. A count of 100 cells/μl or below, therefore, represents advanced HIV disease and progression to AIDS. Above 500 CD4 cells/μl it would be highly unusual for there to be opportunistic infections and the most recent NIH guidelines do not recommend treating above this level (DoHaH 2009).

To identify mutations in the HIV proteome that were associated with individual HLA Class I alleles we used an established technique – phylogenetic dependency networks (Carlson, et al. 2008) - which has been widely implemented in the HIV literature (Bhattacharya, et al. 2007, Brumme, et al. 2008, Matthews, et al. 2008, Rousseau, et al. 2008). In this approach, the risk of error due to founder effect and misidentification of co-variant sites is corrected for by integrating the phylogeny of the sequences into

the analysis. Identified associations are hypothesised to be the result of cytotoxic T cell imposed selection pressure and in other analyses such associations have been proven experimentally. Although there have been a number of HLA-association studies reported, none have stratified patients according to disease progression and compared the strength of selection in different strata according to $\log_2$-adjusted odds ratios.

In our cohort of 895 patients we identified 151 HLA-associated mutations in HIV-1 gag, pol or nef that were significant in patients with either high or low CD4 cell counts. By comparing $\log_2$-adjusted odds ratios in the CD4 count strata, we found that there was evidence of both immune relaxation and intensification in patients with low CD4 counts compared with high CD4 counts. This suggests that for some HLA Class I–restricted responses the selection pressure imposed by CTL is maintained at very low CD4 cell counts but that for others there may be reversion of escape mutations compatible with a reduction in selection pressure. Unlike Gag and Pol, which are structural and enzymatic functional proteins of HIV, Nef (an accessory protein of HIV) confers no sites of relaxing selection association. This is suggestive of fitness implication of Gag and Pol proteins, at least at certain domains, when compared to Nef protein.

Although we identified a change in the pattern of HLA class I mediated polymorphisms in AIDS, the sequencing bioinformatics data alone are not sufficient to infer changes in selection pressure. We have also tested the hypothesis by looking for evidence of changes in T cell-imposed selection pressure measured by identification of viral sequence variation and gamma interferon ELISPOT assays. We

found that the breadth of the response narrowed as the CD4 count declined and that, specifically, the contribution of Gag responses to the overall response decreased. It has previously been reported that acute HIV infection is associated with narrow, high affinity T cell responses which broaden out as patients progress into chronic infection (Streeck, et al. 2009). Here, we see some evidence that is in keeping with previous studies, where the higher CD4 cell counts group (often associated with more early stages of infection) seem to have a narrower absolute ELISPOT breadth that broadens in the lower CD4 cell counts group, with the exception of lowest CD4 cell counts strata (<100 cells/μl).

However, the reverse occurs in the lowest CD4 cell counts strata, with a return to a significantly narrower T cell ELISPOT responses as CD4 cell counts become less than 100 cells/μl ($p=0.029$, Mann Whitney test). The absolute ELISPOT breadths against HIV Gag and Nef decreased most significantly ($p=0.003$ and 0.030, respectively). In terms of relative breadth, Gag narrows but Env and Vif broadens in their proportion within all ELISPOT responses as patient progress to AIDS. It would be interesting to record whether there was a loss of high affinity T cells in the advanced stage patients in this cohort as this might be compatible with a weakening in selection pressure. This is work that will be conducted in the next phase of the Bloemfontein study.

Two key sites potentially reflecting both immune 'persistence' and 'relaxation' were the HLA B*57/B*5801 associated T242N mutation and the B*8101 associated T186S mutation, respectively. For the former there was evidence of maintained (and possibly, on-going) mutation associated with the compensatory mutations, whereas for the

latter we found evidence for reversion. An alternative explanation for immune relaxation is that the patients who progress are the ones who – despite carrying the restricting HLA Class I allele – do not make a response and therefore do not select mutants, therefore enriching patients with low CD4 counts with wild-type viruses. This would be compatible with the findings that targeting certain key epitopes may be associated with different speeds of clinical progression (Dinges, et al. 2010), but does not explain the eventual outcome of those patients enriched for immune escape mutations with high CD4 cell counts.

To determine which occurs would require a longitudinal study of progression to AIDS which would be unethical, and we are therefore limited to cross-sectional analyses. It is informative that in HLA B*8101 positive patients from DBN cohort with data on ELISPOT responses, sequence information and CD4 cell count, of those with CD4 cell counts between 200 – 500 cells/μl, 50% (3/6) of patients who did not recognise TL9 carried the T186S mutant, suggesting a loss of responses. For patients with CD4 counts < 100 cells/μl there were no ELISPOT negative patients with T186S (0/2), suggesting that these may had reverted rather than having been consistently wild type, although in this study numbers are too small to be conclusive.

Taken together, we have provided evidence suggestive of changes within HLA class I-mediated CTL selection on the HIV as patients progress to advanced AIDS. Changes in HLA class I-mediated sites include those sites conferring potential fitness costs, suggested by ELISPOT responses, necessity to revert within HLA class I mismatched hosts following transmission, and the enrichment of wild type state in advanced disease. Whilst we do not have direct evidence to distinguish the two

potential causes of the enrichment of wild type amino acids in advanced AIDS ('immune-relaxation' versus a biased progression of patients with wild type sequences), the observation suggest that fitness may play a key role in disease progression. To confirm the exact impact of mutations at these crucial HLA class I-mediated sites, we have performed experiments to determine viral replication capacity (VRC), in Chapter 6.

CHAPTER 6:

Compensatory mutations and immune relaxation are associated with increased viral fitness in advanced AIDS

6.1 Background

In Chapter 5, we have observed evidence suggestive of changing HLA class I mediated CTL selection across different HIV proteomes over decreasing CD4 count, especially the progression into advanced AIDS. We have noted the narrowing of absolute breadth in ELISPOT responses against the HIV proteome (especially those targeting HIV Gag and Nef proteins) in advanced AIDS patients. In terms of proportion within ELISPOT responses, Gag narrows but Env and Vif broaden in their relative breadth as CD4 count decreases. Furthermore, the reduced HLA class I associated polymorphisms in advanced AIDS, whether due to "immune-relaxation" reversion or failure of escape-inducing CTL responses, infers fitness implications of escape mutations at crucial certain Gag sites selected by certain HLA class I (including T186S in TL9 epitope restricted by HLA-B*8101 and T242N in TW10 epitope restricted by HLA-B*57/B*5801).

In this chapter we question whether progression to AIDS is associated with an increase in viral fitness, and whether this is related to changes in T cell imposed selection pressure at certain sites in HIV-Gag. Based on the results from Chapter 5, we measure the viral fitness of these variants to correlate the replicative cost of

adaptation with outcome. We investigated 148 untreated HIV-1 infected patients from Bloemfontein that host specific HLA class I alleles ("protective"-associating alleles including HLA-B*8101 and HLA-B*57/B*5801, "adverse"-associating alleles including HLA-B*5802, and a mixture of "neutral"-associated alleles that has no significant associations with clinical outcomes). Viral fitness was measured using a recombinant assay in which a patient-derived PCR amplicon of HIV-1 gag-protease was inserted into a gag-protease-deleted pNL4-3 plasmid. We found that viral fitness was greater in patients with advanced disease with examples of both viral compensatory mutations and of reduction in escape mutations, still possibly due to 'immune relaxation'. Despite the complex interplay between host and pathogen, these data support a key role for the restoration of viral fitness in association with AIDS progression and help explain the terminal rise in viraemia.

6.2 Hypothesis

In chronic, untreated HIV infection:

(i) HLA class I mediated CTL on specific structurally or functionally important sites (such as HIV Gag positions T186S in TL9 epitope restricted by HLA-B*8101 and T242N in TW10 epitope restricted by HLA-B*57/B*5801) induce immune-escape that compromises viral replication capacity (VRC).

(ii) The increased viraemia in AIDS is associated with increased VRC, which either occurred as a result of adaptation through accumulating compensatory mutations, or as a result of reversion of immune-escape mutations to wild type at sites of high fitness cost ('Immune-relaxation').

6.3 Aim

To compare the VRC of chimeric HIV (which recombines autologous *gag-protease* in the presence of laboratory NL4-3 viral backbone cassette) isolated from specific HLA class I alleles between the high (>500 cells/μl) and low (<100 cells/μl) CD4 cell counts strata. The results are also compared to other clinical measurements such as plasma viral load.

6.4 Results

6.4.1 Increased viral replication capacity (VRC) of chimeric HIV-1 NL4-3 at low CD4 counts and high plasma viral loads

To investigate the impact of the gag-protease mutations selected by different HLA class I alleles at different clinical stages, we constructed 148 chimeric HIV-1 NL4-3 viruses containing autologous gag-protease from patients of the Bloemfontein cohort (Table 5.1 and Table 6.1). Patients with specific HLA class I alleles previously associated control of HIV in this population (HLA-B*57/5801/8101) or lack of control (B*5802) (Kiepiela, et al. 2004) were selected for chimeric virus construction and stratified by CD4 count. 88 viruses were constructed from patients with low CD4 cell counts (<100 cells/μl), comprising HLA-B*57 (n=2), HLA-B*5801 (n=15), HLA-B*5802 (n=54) and HLA-B*8101 (n=10) plus 14 'neutral' (i.e. not protective) HLA Class I alleles. From patients with high CD4 cell counts (>500 cells/μl), 42 viruses were made: HLA-B*57 (n=7), HLA-B*5801 (n=8), HLA-B*5802 (n=15), HLA-B*8101 (n=10) plus 10 'neutral' HLA Class I alleles. Sixteen patients co-

expressed two of these HLA class I alleles. A further 18 viruses were made from patients with intermediate CD4 T cell counts, all randomly selected without knowing HLA class I alleles. The VRC of the chimeric virus was expressed as percentage growth rate compared to the pNL4-3Δgag-pro plasmid recombined with the wild type HIV-1 NL4-3 gag-pro insert.

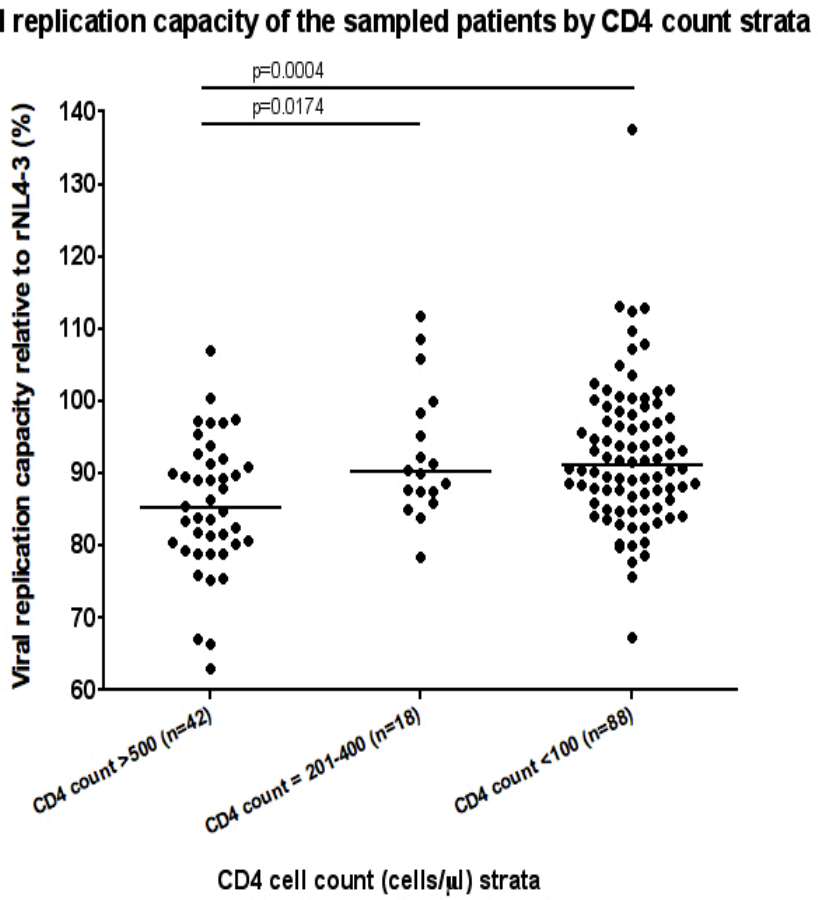
Initially we determined whether CD4 cell counts alone was associated with VRC. Figure 6.1a shows that the VRC of the chimeric viruses from patients from different CD4 T cell counts strata were significantly different (Figure 6.1a). The VRC from viral isolates from patients with the lowest CD4 T cell counts (<100 cells/μl, median 91.15%; IQR: 86.1-98.0%) was significantly higher than from patients with high CD4 T cell counts (>500 cells/μl, median 85.19%; IQR: 79.8% - 91.8%), p=0.0004 (Mann-Whitney test).

There was also a weak but statistically significant positive correlation between VRC and $\log_{10}$ plasma viral load (Figure 6.1b, p=0.0003, r^2=0.088, n=142). The slope of the best-fit values (0.029±0.0079) indicated that a 34.4% increase in VRC accounted for an increase of one $\log_{10}$ in plasma viral load.

Table 6.1: Number of patients selected for construction chimeric viruses in terms of specific HLA-B alleles and CD4 count strata

	HLA-B alleles				
	B57	B*5801	B*5802	B*8101	Other "Neutral" HLA-B alleles
High CD4 counts (>500 cells/ul)	7	8	15	10	42 (Continuous CD4 cell counts Spread)
Low CD4 counts (<100 cells/ul)	2	15	54	10	

(a)



(b) The relationship between viral replication capacity and viral load

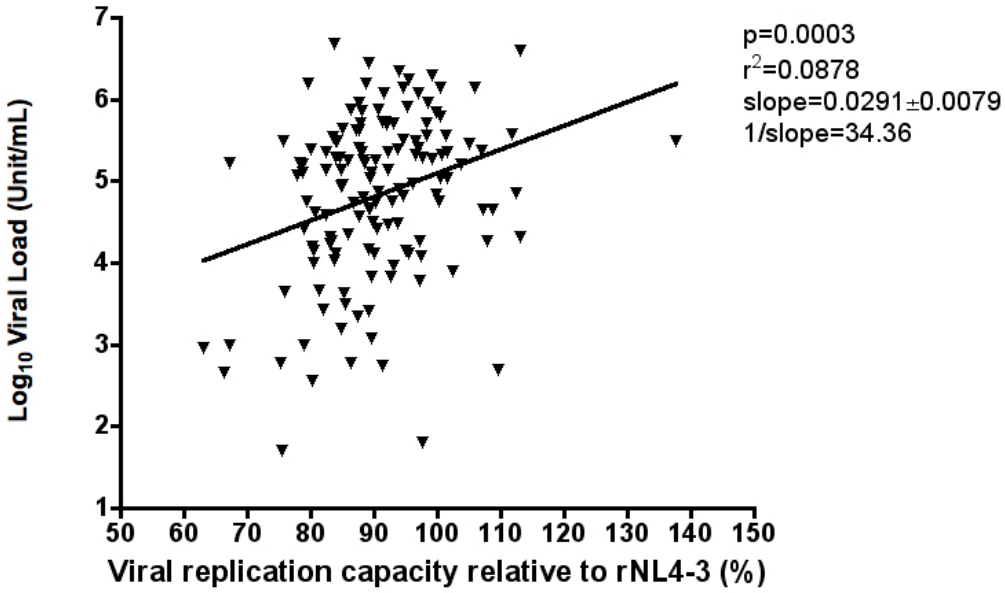


Figure 6.1: Viral replication capacity (VRC) according to CD4 count and plasma viral load

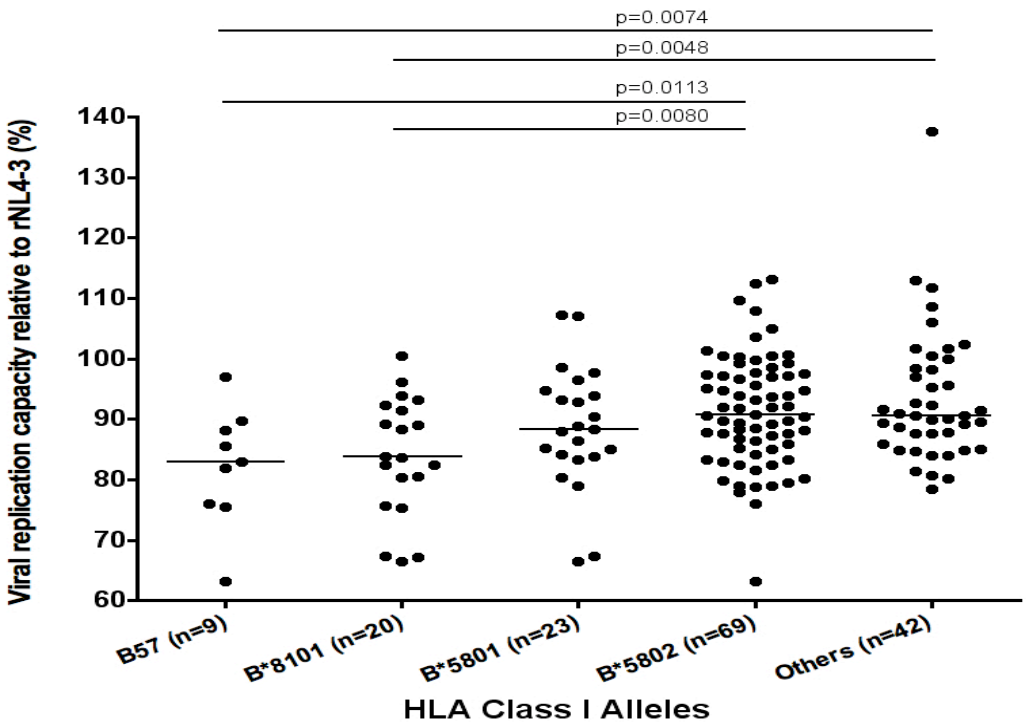
The viral replicative capacity (VRC) of chimeric NL4-3 viruses recombined with patient autologous gag-protease genes is compared with a laboratory recombinant HIV-1 NL4-3 strain. In (a.), The VRC of isolates from the Bloemfontein cohort is presented stratified by CD4 T cell count. The p-values are calculated using the Mann-Whitney test. The α-value for the study is equal to 0.05, however, for a significant test result, the p-value should be less than 0.025 after Bonferroni correction for the multiple comparison. In (b.), plasma viral load ($\log_{10}$ copies/ml) is correlated against the VRC from samples from the Bloemfontein cohort. In both panels VRC is presented as a percentage relative to the fitness of the control NL4-3 virus.

6.4.2 VRC of chimeric **gag-protease** *NL4-3 viruses varies according to HLA class I*

As higher VRCs were associated with higher plasma VL and lower CD4 cell count, we looked to see how this effect varied according to HLA Class I (Figure 6.2a). The VRC of chimeric viruses derived from patients with the 'protective' alleles HLA B*57 (median=83.0%, IQR: 75.8-88.9%; n=9) and HLA B*8101 (median=83.7%, IQR: 78.1-91.9%; n=20) were lower than from patients with 'neutral/other' HLA class I alleles (median=90.6%, IQR: 81.1-91.8%; n=26); (p=0.007 and p=0.005, respectively). The VRC of chimeric viruses derived from patients with HLA B*57 and HLA B*8101 were also lower than those derived from patients with HLA B*5802, which is associated with more rapid clinical progression (an allele with potentially 'adverse' effect, median=90.8%, IQR: 85.1-97.5%, n=68) (p=0.010 and p=0.008, respectively, Mann-Whitney test).

To see if the association between HLA Class I and viral fitness was additionally impacted by disease stage we stratified the VRCs according to high and low CD4 cell counts (Figure 6.2b). This sub-group analysis shows that the median VRC values are higher for each HLA Class I allele in the low CD4 count group suggestive of an increase in viral fitness in advanced disease, although in most cases these values are not statistically significant. This is likely to be an effect of being underpowered, as patient numbers become smaller with each sub-categorisation. Interestingly, there is a significant difference in the VRCs of viruses in the non-protective 'neutral/other' HLA Class I alleles (p=0.0092, Mann-Whitney test), showing that although there is an HLA-specific effect on viral fitness there is an additional association with CD4 cell count, independent of the more beneficial HLA Class I alleles.

(a) Viral replication capacity of the sampled patients by HLA alleles



(b) Viral replication capacity of the sampled patients by HLA alleles and CD4 counts

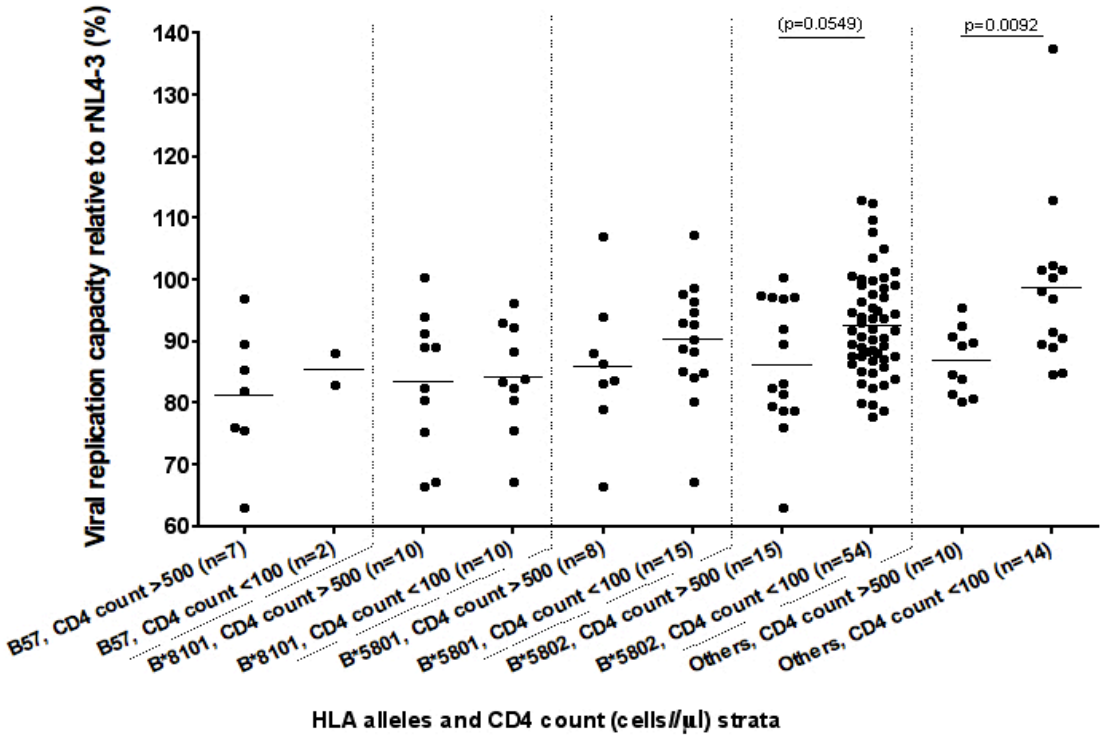


Figure 6.2: Viral replication capacity (VRC) according to HLA Class I and CD4 cell counts

(a) Viral replication capacity (VRC) of chimeric NL4-3 viruses containing patient autologous gag-protease stratified by HLA Class I. The VRC is expressed as percentage growth rate compared to control strains. The p-values are calculated using the Mann-Whitney test. The α-value for the study is equal to 0.05, however, for a significant test result, the p-value should be less than 0.025 after Bonferroni correction for the multiple comparison. Panel (b) shows further stratification according to CD4 cell count. The y-axis depicts VRC, expressed as percentage growth and the x-axis depicts HLA class I, with each HLA class I category divided further into "low" CD4 T cell count strata (<100 cells/mm3) and "high" CD4 T cell count (> 500 cells/mm3). The p-value is calculated using Mann-Whitney test.

---

*6.4.3 The impact of different mutations on VRC in the context of HLA-B*8101 and HLA-B*B*57/B*5801*

These analyses have shown that viral fitness is greater in viruses from patients with advanced disease, and that certain protective HLA Class I alleles (such as B*57, B*8101 and B*5801) are associated with greater fitness costs. The HLA-association analysis revealed that in patients with low CD4 counts HLA B*57/*5801 was associated with persistence of T242N in the TW10 epitope, whereas HLA B*8101 was associated with a potential loss of selection pressure reflected by the decreased prevalence of T186S mutants in the TL9 epitope. We therefore targeted these specific associations to see how they impacted viral fitness at different stages of disease progression.

For HLA-B*8101, we excluded from the analysis any viruses with key mutations restricted by other HLA Class I alleles (e.g. T242N and A163G) which were known to impact fitness. Viruses from patients with HLA B*8101 were significantly less fit than viruses from patients in the rest of the cohort ($p=0.006$, Mann-Whitney test; Figure 6.3a) and mutations in the B*8101-restricted Gag TL9 epitope were associated with lower VRCs ($p=0.024$) both in the whole cohort and in patients with HLA-B*8101 ($p=0.025$). The main single mutation conferring loss of replicative fitness was the T186S immune escape mutation inside the TL9 epitope ($p<0.0001$). Neither the Q182S or T190X mutations had an impact on VRC, in this cohort (data not shown). Although patient numbers were small, there was no significant difference between the VRCs of viruses with the T186S mutation according to high and low CD4 count stratification. These data show that B*8101 imposes a fitness cost on the virus through selection of the T186S mutation in the TL9 epitope, we found no

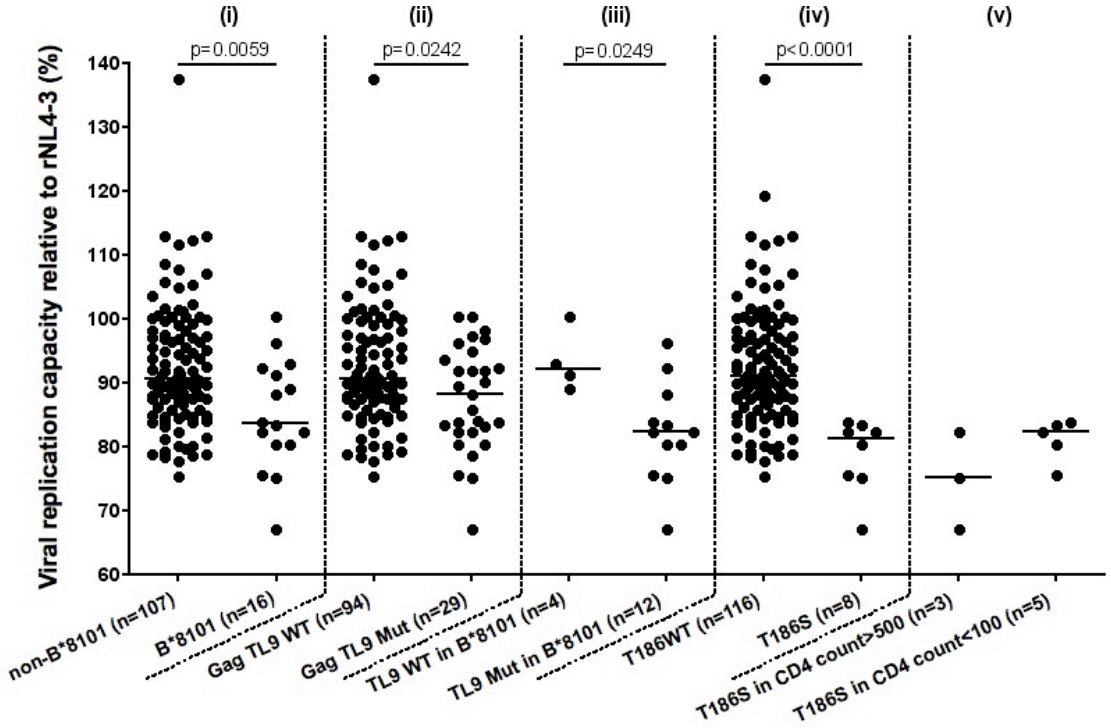
evidence that, when this mutation persists, the fitness cost is restored at low CD4 cell counts.

We compared the fitness costs associated with HLA B*8101 with those associated with B*57 and B*5801, as the HLA-association analysis suggested that for these latter alleles escape mutations persisted in advanced disease. We pooled the data for patients with HLA B*57 and B*5801 as these alleles have highly related binding motifs, and both present the Gag TW10 epitope and select the T242N escape mutation (Leslie, et al. 2004). As for HLA B*8101, viruses from patients with HLA-B*57 and HLA-B*5801 were significantly less fit than those from the rest of the cohort ($p=0.030$, Mann-Whitney test; Figure 6.3b), supporting the association between protective HLA Class I alleles and viral replicative cost (Figure 6.2a). In the whole cohort there was no significant effect of all mutations in the TW10 epitope on VRC. For patients with HLA B*57 or B*5801, there was a non-significant trend for mutations within the TW10 epitope to impair fitness, although the power of this analysis was limited as only one of the 22 patients had not selected a mutant epitope. However, the fitness cost of the T242N mutation was confirmed when the whole cohort was analysed, when it was associated with a significant drop in VRC ($p=0.029$). This result was independent of the effect of mutations in other B*57/B*5801 associated epitopes such as Gag KF11 or IW9 ($p=0.024$; data not shown), showing that the fitness cost was associated with T242N, rather than being positive for HLA-B*57/B*5801.

In the HLA-association analysis we had found that T242N was maintained in late disease, even though these data, as well as previous reports, suggest that this mutation

is associated with a fitness cost. We explored this apparent discrepancy by stratifying the fitness analysis of T242N according to CD4 count (Figure 4b), and found that in patients with advanced disease (CD4 counts <100 cells/μl), viral fitness had been restored even though the T242N persisted (p=0.003).

(a) Viral replication capacity of HIV isolates stratified by host HLA-B*8101, CD4 counts and viral Gag TL9 epitope mutation status



(b) Viral replication capacity of HIV isolates stratified by HLA-B*57, B*5801, CD4 counts, and viral Gag TW10 epitope mutations

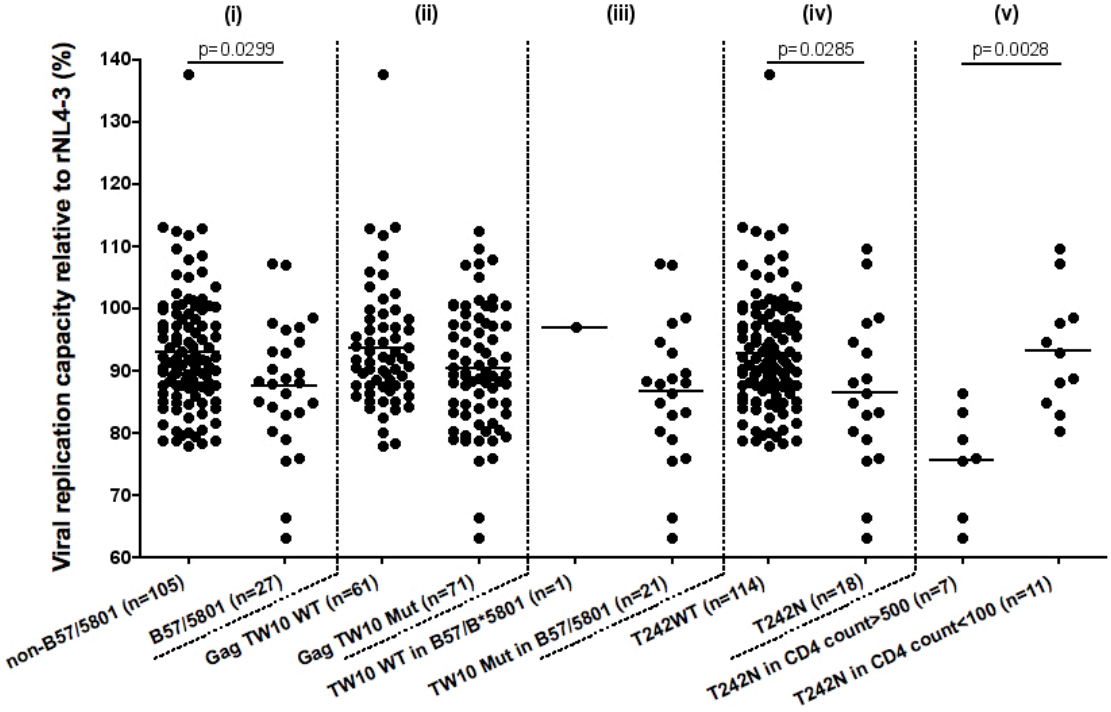


Figure 6.3: Viral replication capacity (VRC) in the viruses with mutants in (a) Gag TL9 and (b) Gag TW10 according to HLA Class I and CD4 cell count

In panel (a), the y-axis depicts viral replication fitness (VRC) of chimeric NL4-3 viruses containing patient autologous *gag-protease*, expressed as percentage growth rate compared to a control recombinant NL4-3 strain. The x-axis depicts the stratification of patients according to (i) HLA-B*8101, (ii) polymorphic status within the *gag* TL9 epitope in all patients, (iii) polymorphic status of *gag* TL9 in patients with HLA B*8101, (iv) presence of T186S in *gag* in Gag TL9, and (v) T186S further stratified according to the CD4 T cell count (<100 cells/μl or > 500 cells/μl). The p-value is calculated using Mann-Whitney test. In (b), patients are stratified according to (i) HLA-B*57 or B*5801 alleles, (ii) polymorphic status within the *gag* TW10 epitope, (iii) polymorphic status of TW10 in patients with HLA-B*57 or B*5801 alleles, (iv) polymorphic status of amino acid position 242 (within TW10 epitope – T242N) in all patients and (v) then by T242N in patients with HLA-B*57 or B*5801 according to CD4 count. The viral isolates possessing other known impairing escape mutations (A163G and T186S) are excluded from this analysis. The p-value is calculated using Mann Whitney test.

*6.4.4 Compensatory mutations accrue in advanced HIV infection in patients with HLA B*57 and B*5801*

As the T242N mutation has been previously linked with compensatory mutations we analysed our cohort to see if these accrued in late stage disease to explain the rise in viral fitness despite the maintenance of T242N. For the T242N mutation in TW10, compensatory mutations have previously been published (H219Q, I223V and M228, (Brockman, et al. 2007)), and so we looked for these mutations in the different CD4 T cell counts strata (Figure 6.4). As expected, we found a cumulative increase in the number of compensatory mutations in patients with T242N in individuals with lower CD4 cell counts, consistent with the increase in fitness in the previous analyses ($p=0.013$, Mann-Whitney test).

Together these data suggest that advanced HIV disease is associated with an increase in viral fitness, which may contribute to the rise in plasma viral load frequently seen in AIDS. However, this increase in fitness is multifactorial and we report two processes specific to different HLA Class I alleles – immune relaxation and compensatory mutations – to potentially explain these findings.

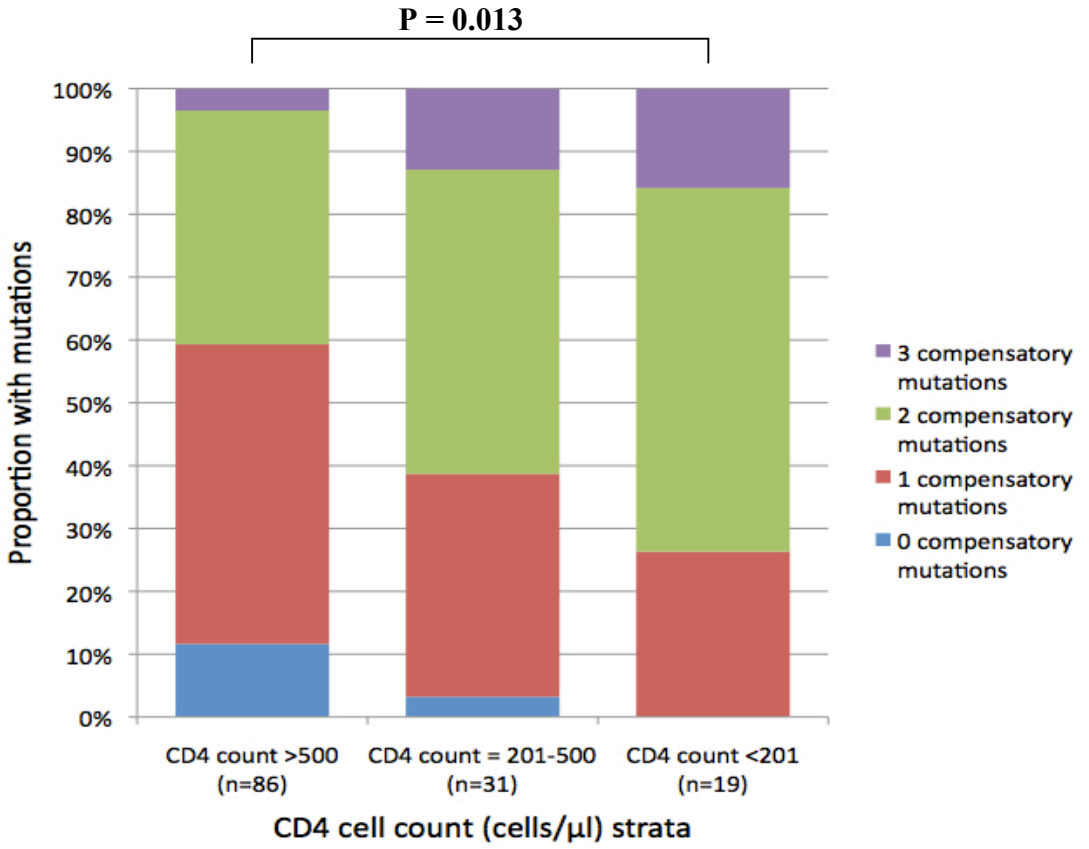


Figure 6.4: Prevalence of compensatory mutations with the T242N mutation in the HLA-B*57 and B*5801 restricted epitope, Gag TW10

In the 100% stacked column, the y-axis shows the number of mutations accumulated at the three known Gag protein amino acid sites (H219Q, I223V and M228I) that compensate for the T242N mutation (Brockman, et al. 2007). The x-axis depicts the categorical CD4 T cell counts strata and the number of patients within each strata (the "High CD4" group refers to patients with CD4 T cell counts >500 cells/µl, the "Intermediate CD4" group refers to patients with CD4 T cell counts between 200 and 400 cells/µl, and the "Low CD4" group refers to patients with CD4 T cell counts <100 cells/µl).

6.4.5 Preliminary screening of potential mutational fitness cost at sites selected by HLA mediated selection using stratification of VRC and viral load

The impact of mutations in individual amino acids in the HIV-1 Gag protein on VRC and plasma viral load were measured (preliminary data). Stratification of VRC and plasma viral load according to mutational status of the amino acid site in HIV-Gag under CTL selection were presented in Figure 6.5 and 6.6, respectively. Whilst results from VRC (Figure 6.5) suggest fitness impact of T186S (mainly selected by B*8101, p=0.006, Mann Whitney test), A163G (mainly selected by B*57, p=0.032) and T242N (mainly selected by B*57 and B*5801, p=0.030), the viral load according to amino acid polymorphism showed that T242N, but not A163G or T186S, were associated with a significantly lowered viral load (Figure 6.6, p=0.004, Mann-Whitney test). Although VRC was not evaluated in the viruses derived from HLA-B*27 positive hosts, the viral load of viruses possessing mutations in HLA-B*27 mediated R264K mutations (in KK10 epitope) is significantly lowered than that of population median (p=0.049, one tailed Mann-Whitney test).

However, it is worth noting that the analytic results presented here are preliminary only. Several limitations exist in the current analysis used: firstly, the stratification of VRC and viral load according to site-specific polymorphism did not control for the polymorphic status at other *gag* amino acids and statistic multiple comparisons. Secondly, bias exists in the sampling of host for VRC experiment (Table 6.1), results presented in Figure 6.5 are enriched with hosts of specific HLA class I alleles and CD4 count strata, and is therefore underpowered for many HLA related site selections and may not be representative of the entire cohort. VL, on the other hand is not biased in its sampling, apart from the cohort set up (Chapter 2.1). Thirdly, the causal

relationship to explain the result of stratification is complicated by the uneven HLA class I distribution across CD4 counts, which is known to be independently associated with increased VL (and VRC).

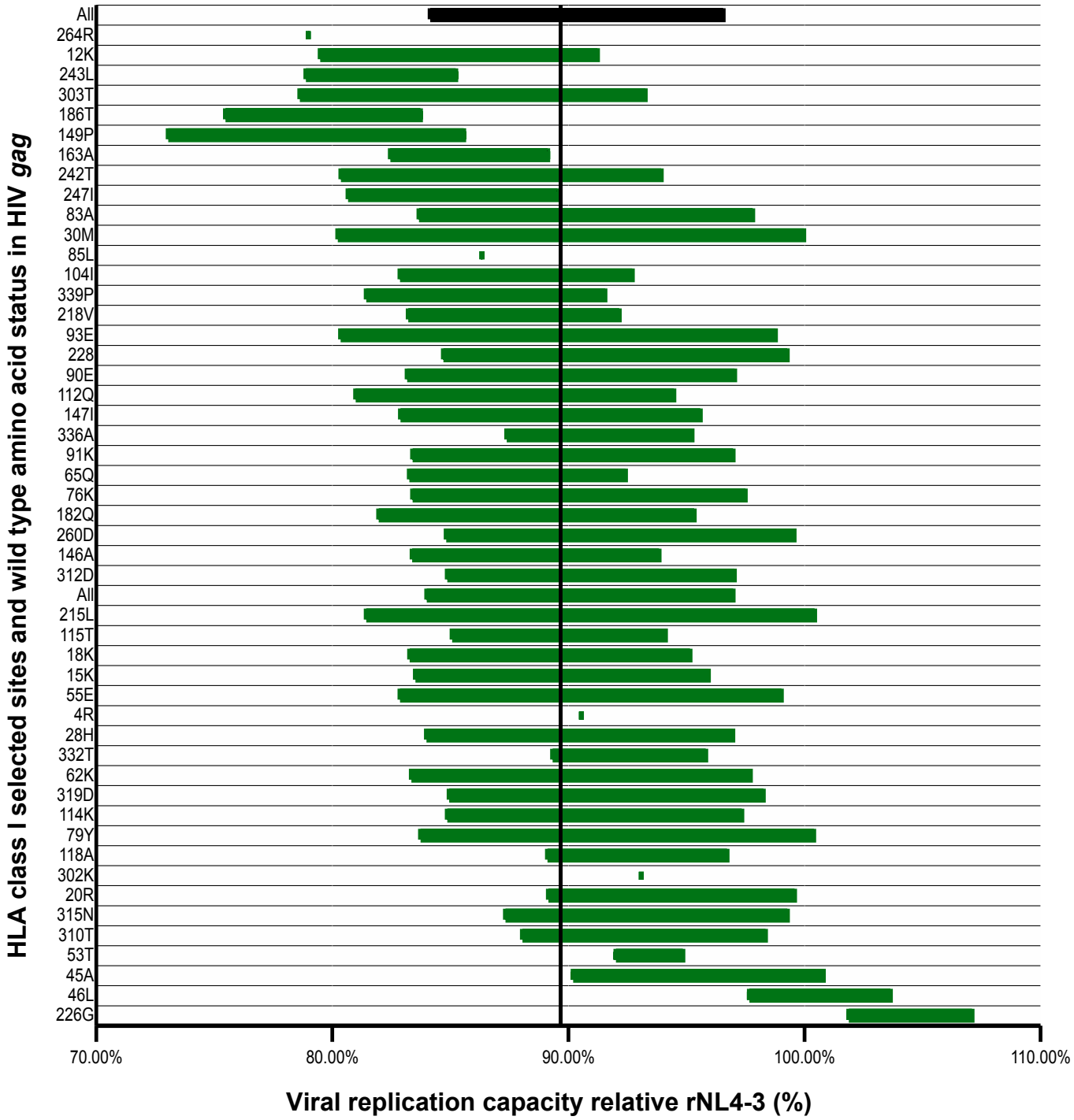


Figure 6.5 The viral replication capacity stratified according to *gag* amino acid sites that are under evident HLA mediated selection

In the floating bar chart, the y-axis depicts the Gag protein amino acid site under HLA mediated selection (identified in Chapter 5) and its majority consensus amino acid residue. The x-axis depicts viral replication capacity (as % relative to laboratory recombinant NL4-3 virus) in the patients who are polymorphic at the respective Gag amino acid position. Each floating bar presents the range between 25^{th} and 75^{th} percentile of the VRC in polymorphic patients. The black floating bar reflects the range between 25^{th} and 75^{th} percentile of the VRC in entire cohort (Bloemfontein patients with VRC measured in Table 6.1), with the black vertical line displaying the population median VRC.

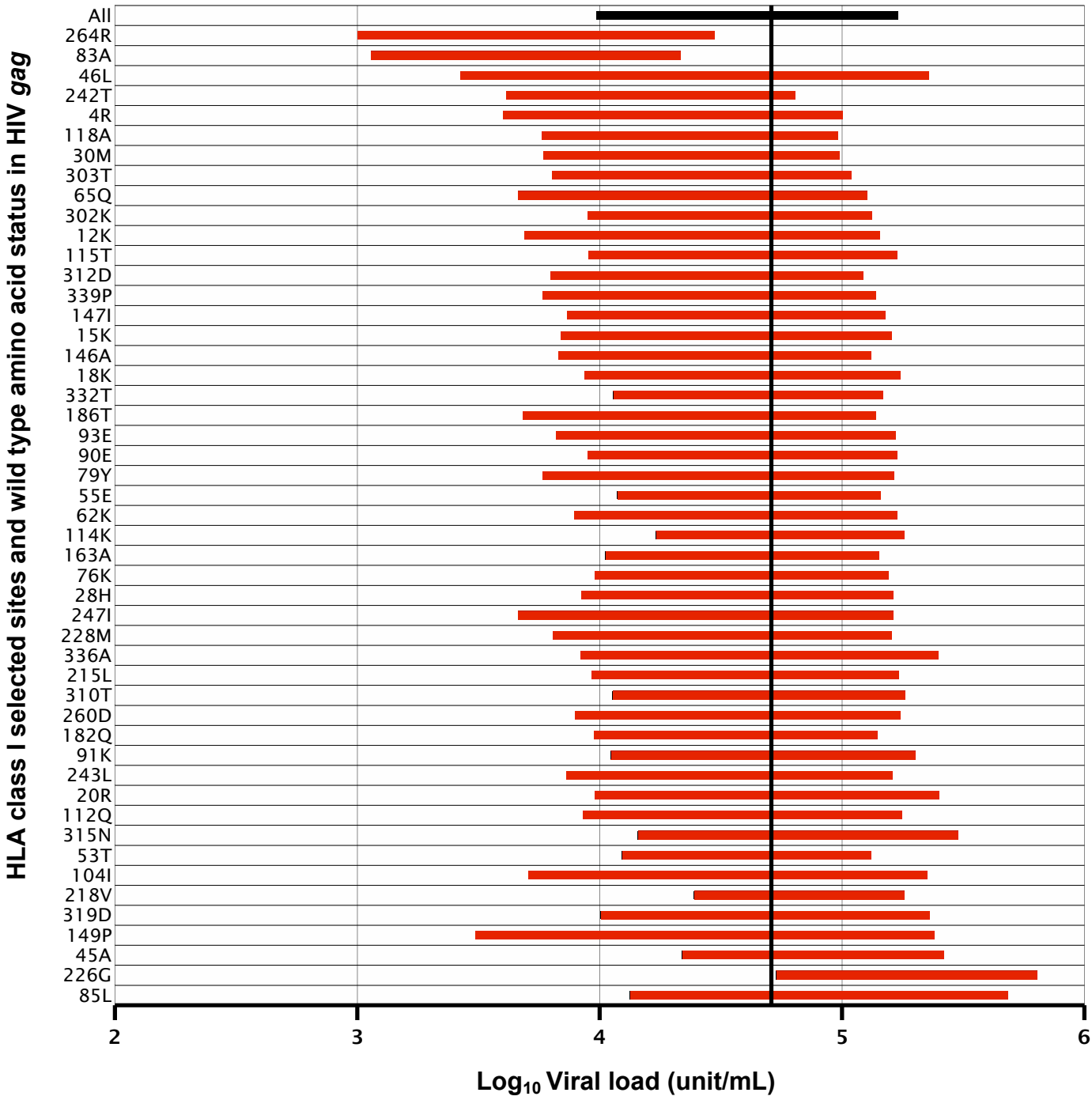


Figure 6.6 The viral load stratified according to *gag* amino acid site sites that are under evident HLA mediated selection

In the floating bar chart, the y-axis depicts the Gag protein amino acid site under HLA mediated selection (identified in Chapter 5) and its majority consensus amino acid

residue. The x-axis depicts viral load relative of patients who are polymorphic at the respective Gag amino acid position. Each floating bar presents the range between 25th and 75th percentile of the viral load in polymorphic patients. The black floating bar reflects the range between 25th and 75th percentile of the viral load in entire cohort (Combined Durban and Bloemfontein cohorts listed in Table 5.1), with the black vertical line displaying the population median viral load.

6.5 Discussion

In this study, viral fitness was measured using a recombinant assay in which a patient-derived PCR amplicon of HIV-1 gag-protease was inserted into a gag-protease-deleted pNL4-3 plasmid. The advantage of this approach is that it uses the full *gag* genome from the patient sequence as well as the associated protease, which encodes the enzyme for cleavage of the Gag-Pol protein. The disadvantage of this assay is that it is in a parallel rather than competitive format, and therefore may lack sensitivity for small fitness differences. In addition, the construct is a recombinant of a subtype B backbone with a subtype C insert, which evidently grow at a slower rate compared to subtype B recombinant viruses (data unpublished). Although this approach has been published elsewhere, any data should be interpreted relative to the control virus rather than as an absolute measure of viral fitness (Miura, et al. 2009, Miura, et al. 2010, Wright, et al. 2010). Choice of control is important. We used a subtype B wild-type backbone (pNL4-3Δgag-pro) into which we recombined the pNL4-3 gag-protease. This controlled for any change in fitness due to the recombination step. Furthermore, the NL4-3 laboratory strain viruses are CXCR4-tropic, whereas the CEM-GXR cell lines expresses both CXCR4 and CCR5. Although the switching in co-receptor usage in natural infection (from CCR5 tropic in earlier disease to CXCR4 or dual tropic in advanced disease) may interfere with our observation in this study, here we have only introduced the autologous *gag-protease* genes from the patients. The co-receptor switching phenomenon also occurs less frequently in subtype C HIV infections.

We found that VRC was significantly greater in patients with low CD4 cell counts and there was a significant, although weak, correlation with plasma viral load. These are amongst the first data showing that viral fitness impacts plasma viral load and that

this ex vivo assay reflects viral replication in vivo. Analysing our data by HLA Class I, reveals that HLA Class I alleles that are associated with clinical advantage (B*8101, B*57, B*5801) were associated with impaired viral fitness as had been inferred from previous analyses of the reversion of transmitted escape mutations in HLA mismatched hosts (Brumme, et al. 2008, Duda, et al. 2009). Interestingly, however, the increase in viral fitness with lower CD4 cell counts did not appear to be restricted to beneficial HLA alleles and was also present in a selection of viruses from patients with 'neutral/other' alleles.

The HLA-amino acid association analysis showed that for HLA B*8101 there was a decrease in mutations in the Gag TL9 epitope at low CD4 cell counts, whereas for HLA B*57/B*5801 there was persistence of the T242N escape mutation in TW10. The fitness assays showed that the T186S mutation in TL9 was associated with a fitness cost ($P<0.0001$) and it is therefore possible that in patients with B*8101 and low CD4 counts, the decrease in prevalence of the T186S mutation reflects reversion to a fitter strain coincident with the weakening of T cell responses. For B*57/B*5801 we saw a different pattern. Here, possession of HLA B*57/B*5801 was associated with less fit viruses in conjunction with the T242N mutation in the Gag TW10 epitope, but in patients with low CD4 counts fitness was restored even though T242N persisted. We therefore sought compensatory mutations to explain this finding. For T242N there are four documented compensatory mutations at Gag codons 219, 223, 228 and 248 (Brockman, et al. 2007). For this analysis of subtype C HIV-1 we excluded codon 248, as this is already variant in the subtype C consensus. When we stratified patients with T242N according to CD4 strata we found a significant

(p=0.013) increase in numbers of compensatory mutations as CD4 cell count fell, consistent with the rise in VRC.

These data show that viral fitness increases with progression to AIDS. Whether this is cause or effect is difficult to determine, although the fact that we see evidence for both immune relaxation and compensatory mutations in the same cohort indicates that progression may be a mixture of the two. What is clear is that viral fitness is a significant component of progression to AIDS and, as such, should be considered as a target for intervention, for example through vaccines aimed at epitopes which escape with high fitness costs. Trials such as DART have shown that HAART can be extremely effective in regions such as sub-Saharan Africa (Walker 2009), however the costs and logistics of provision are complex. Any intervention that could delay a requirement for HAART could have major economic and health implications and therefore combining such a vaccine strategy with HAART provision could be a productive coalition.

CHAPTER 7:

Conclusion

In this thesis I have explored different selection pressures imposed on HIV, and how the virus adapts to them. In different chapters, I have reported how HIV responds to both the humoral and T cell mediated arms of the immune system, as well as to the pressure imposed by antiretroviral therapy. I have focused my study on the chronic stages of HIV and advanced AIDS, which remain a significant clinical problem in the developing world despite improved access to HAART. The pathogenesis of progression to AIDS remains unclear, and little is known regarding immune responses during clinical CD4 cell depletion. The work I have presented here adds to the field by improving our understanding of the immunological and virological correlates of disease progression.

I hypothesised that a common adaptation strategy exists between different sources of HIV selection, and examined multiple selection forces to demonstrate the host-pathogen interplay, ranging across the role of B cell mediated NAb selection during chronic HIV infection, the molecular and pre-therapy epidemiological study of HIV in a large South African cohort, the bioinformatic analysis of HLA mediated CTL selection on viral sequence evolution in AIDS, and the assessment of viral replicative fitness of isolates derived from different HLA class I, *gag-protease* polymorphisms, pVL and CD4 cell counts. I confirmed the role of each of the selection pressures by examining the effect of their attenuation, outlined the specific pattern of induced viral

mutations and, in the study of CTL selection, demonstrated the fitness impact of the escape mutations.

HIV demonstrates a remarkable ability to rapidly mutate and escape each specific selection pressure. We have shown that many escape mutations confer little fitness cost, whilst some costly escape mutations can evolve secondary mutations that partially compensate the fitness impairment seen in chronic HIV infection. However, our results show that escape mutations such as T186S (driven by the targeting of the TL9 epitope by HLA-B*8101) may confer a significant viral fitness cost, without adequate compensatory mechanisms. Identification of other such detrimental mutations could be informative for CTL based vaccines, which may be designed to elicit costly viral escape.

Although AIDS pathophysiology remains unclear and difficult to investigate (it would be unethical to conduct a longitudinal study of untreated progression to AIDS), we have shown that disease progression is associated with uneven HLA Class I allelic distribution, altered CTL selection and increased viral fitness. Using molecular epidemiology and phylogenetic methods, I have shown that the HIV found in AIDS patients evolves within the host instead of resulting from a resurgence of virulent founder strains. Currently, expert opinion and clinical guidelines continue to vary regarding the best timing for the initiation of HAART (Gatell 2010). It is hoped that our characterisation of the host-pathogen interaction during advanced HIV and AIDS might inform such decision-making. In particular, we envisage that the temporal and functional attenuation of CTL and Nab, the development of ‘immune relaxation’, the

accumulation of compensatory mutations, and increases in viral fitness will all contribute to AIDS pathogenesis and therefore, warrant prompt intervention.

Similarly, our findings in Chapter 3 have provided a cautious message regarding the use of Rituximab in HIV patients who are not covered by HAART. The findings in Chapter 4 provided examples of the public health implications of the clustering of drug-associated mutations in patients with very low CD4 counts. The studies in Chapters 5 and 6 have also highlighted the host factors relating to viral evolution and disease progression, which may warrant a more individualised clinical and public health intervention. In resource-limited settings, such as South Africa, our findings may contribute towards guidelines that help to identify vulnerable individuals and offer prioritised investigation or intervention.

The findings presented and discussed in this thesis represent part of a larger body of studies. Many of the study topics above are on-going at the time of this thesis submission. However, the results provided by this thesis have informed the further work now being carried out in the laboratory. In particular, the case report findings in Chapter 3 prompt the prospective study to investigate the impact of Rituximab (and therefore CD20+ cell subsets) on the antiviral T cell. The molecular epidemiology and pre-therapy drug resistance studies in Chapter 4 have provided baseline information for the regional drug resistance surveillance program. The virological and preliminary ELISPOT findings in Chapter 5 and 6 have prompted us to focus our ELISPOT assay within specific HLA Class I subset and epitope interactions.

Ultimately, the impact and short-term future goals of the thesis outlined above aims to address some of the following larger challenges: Firstly, compare and contrast the strategies of HIV evolution under different host and drug-mediate selection. This is best exemplified by our observation of similar reversion strategies virus utilised to regain replication fitness during post-retuximab (Chapter 3) and AIDS phases (Chapter 5 and 6). Many more selection strategies offered by innate immunity players are also important for further investigation. Secondly, identifying potentially synergistic selection combination that will minimise viral escape, and hopefully, achieve viral elimination. One of the potential strategies we have identified in this thesis include targeting of the multiple sites that confers important replicative fitness implication and low potential for immune escape mutations. The development of HAART provides proof of principle that combining antiviral targets may result in synergistic viral suppression and prevent the emergence of stable escape mutants (Frater, et al. 2002, Palella, et al. 1998). Therefore, we are optimistic that the design of an antigenic vaccine capable of eliciting multiple immune responses (mediated by B and T cells) against the multiple "Achilles tendons" of HIV, may achieve favourable viral protection and control.

In summary, our collective approaches to HIV evolution have demonstrated that viral adaptation is dependent on the nature of the imposing selection pressures, the pathway and impact of mutational escape, and the existence of archived variants. Depending on the circumstances, viral adaptation to drug and immune selection pressure is achieved mainly through escape mutations, compensatory mutations, or reversion.

The models below summarise and integrate the key themes emerging from the thesis, featuring both within-host (Figure 7.1) and between-host (Figure 7.2) HIV evolution. Additionally there is likely to be an important interplay between the T cell mediated selection/control and antibody mediated selection/control. The balance between these forces has not yet been addressed in prospective studies.

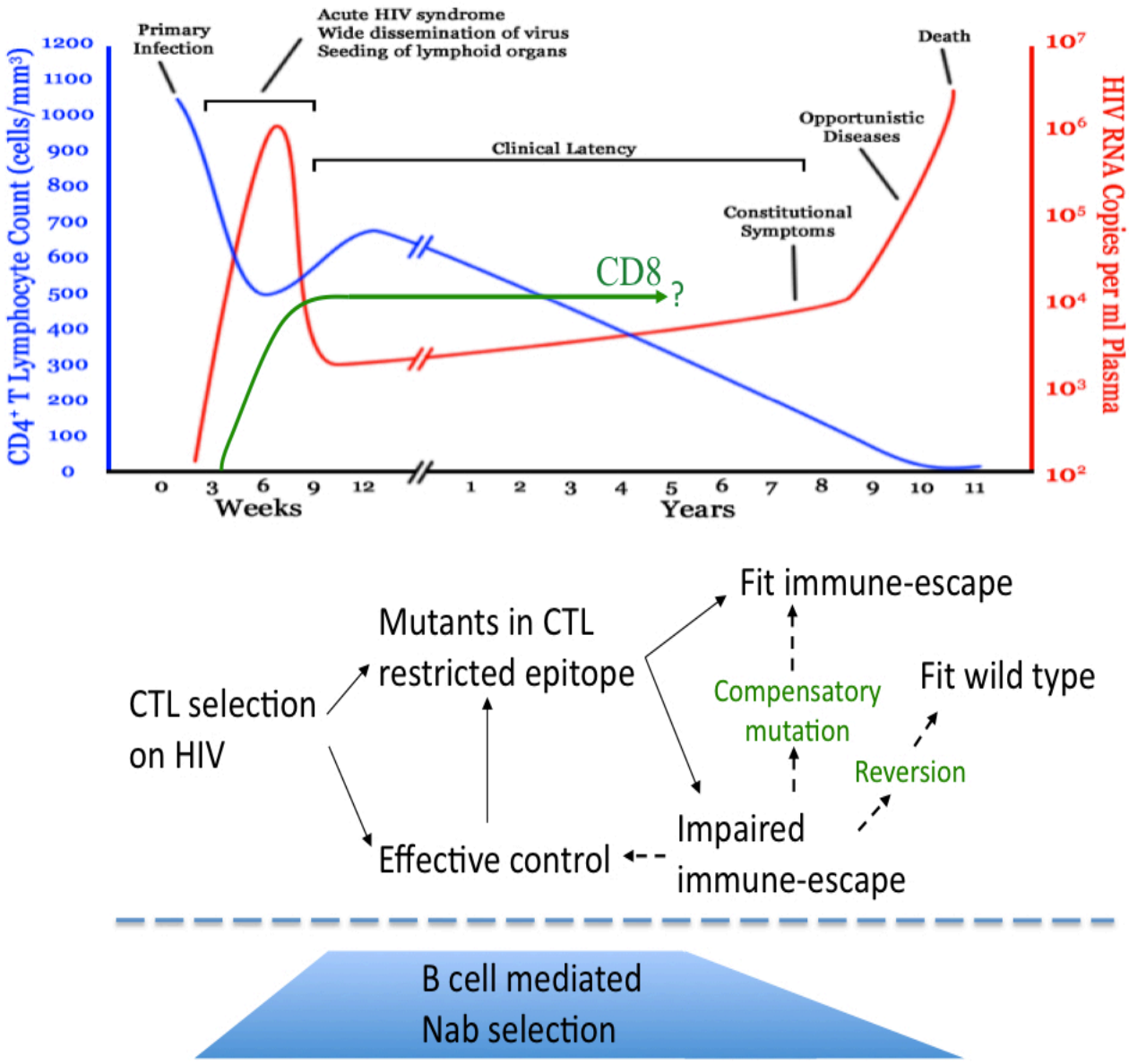


Figure 7.1: Schematic summary of within host HIV selection over time

(Image modified from Pantaleo G. et al, 2003 (Pantaleo, et al. 1993))

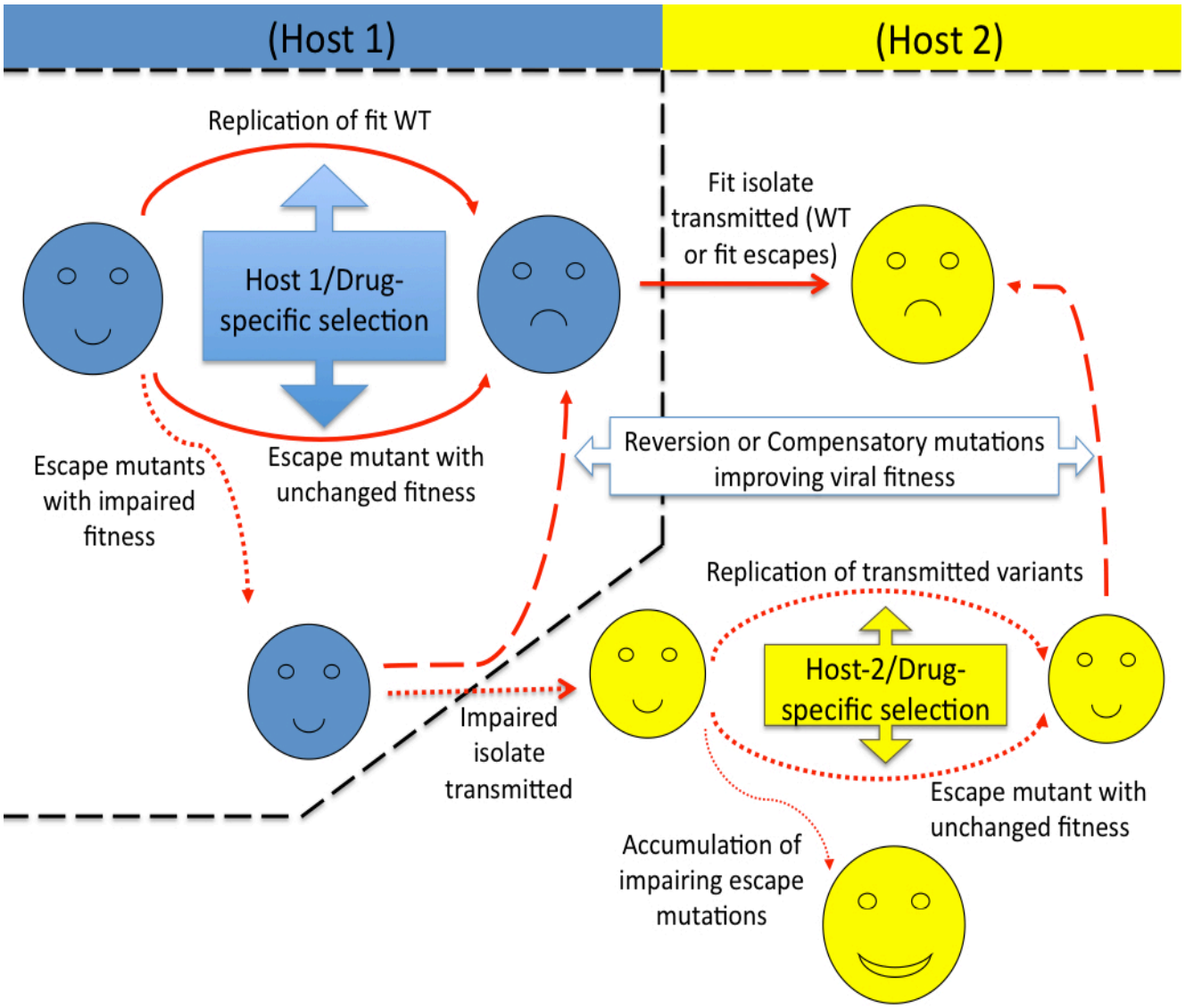


Figure 7.2: Cartoon summary of inter-host HIV selection within population

REFERENCES

1. **Aboulker, J. P., and A. M. Swart.** 1993. Preliminary analysis of the Concorde trial. Concorde Coordinating Committee. Lancet **341:**889-90.
2. **Adachi, A., H. E. Gendelman, S. Koenig, T. Folks, R. Willey, A. Rabson, and M. A. Martin.** 1986. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. J Virol **59:**284-91.
3. **Addo, M. M., R. Draenert, A. Rathod, C. L. Verrill, B. T. Davis, R. T. Gandhi, G. K. Robbins, N. O. Basgoz, D. R. Stone, D. E. Cohen, M. N. Johnston, T. Flynn, A. G. Wurcel, E. S. Rosenberg, M. Altfeld, and B. D. Walker.** 2007. Fully differentiated HIV-1 specific CD8+ T effector cells are more frequently detectable in controlled than in progressive HIV-1 infection. PLoS One **2:**e321.
4. **Aksoy, S., H. Harputluoglu, S. Kilickap, D. S. Dede, O. Dizdar, K. Altundag, and I. Barista.** 2007. Rituximab-related viral infections in lymphoma patients. Leuk Lymphoma **48:**1307-12.
5. **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.

6. **Altfeld, M., and E. S. Rosenberg.** 2000. The role of CD4(+) T helper cells in the cytotoxic T lymphocyte response to HIV-1. Curr Opin Immunol **12:**375-80.

7. **Arien, K. K., G. Vanham, and E. J. Arts.** 2007. Is HIV-1 evolving to a less virulent form in humans? Nat Rev Microbiol **5:**141-51.

8. **Ariyoshi, K., E. Harwood, R. Chiengsong-Popov, and J. Weber.** 1992. Is clearance of HIV-1 viraemia at seroconversion mediated by neutralising antibodies? Lancet **340:**1257-8.

9. **Armstrong, K. L., T. H. Lee, and M. Essex.** 2009. Replicative capacity differences of thymidine analog resistance mutations in subtype B and C human immunodeficiency virus type 1. J Virol **83:**4051-9.

10. **Arthos, J., C. Cicala, E. Martinelli, K. Macleod, D. Van Ryk, D. Wei, Z. Xiao, T. D. Veenstra, T. P. Conrad, R. A. Lempicki, S. McLaughlin, M. Pascuccio, R. Gopaul, J. McNally, C. C. Cruz, N. Censoplano, E. Chung, K. N. Reitano, S. Kottilil, D. J. Goode, and A. S. Fauci.** 2008. HIV-1 envelope protein binds to and signals through integrin alpha4beta7, the gut mucosal homing receptor for peripheral T cells. Nat Immunol **9:**301-9.

11. **Asquith, B., C. T. Edwards, M. Lipsitch, and A. R. McLean.** 2006. Inefficient cytotoxic T lymphocyte-mediated killing of HIV-1-infected cells in vivo. PLoS Biol **4:**e90.

12. **ASSA.** 2007. ASSA 2003 Demographic Model Produced by the Actuarial Society of South Africa. Demographic Model. ASSA.

13. **AVERT.org** 26 April 2010 2009, posting date. South Africa HIV & AIDS Statistics (http://www.avert.org/safricastats.htm). AVERT.org. [Online.]

14. **Baba, T. W., V. Liska, R. Hofmann-Lehmann, J. Vlasak, W. Xu, S. Ayehunie, L. A. Cavacini, M. R. Posner, H. Katinger, G. Stiegler, B. J.**

Bernacky, T. A. Rizvi, R. Schmidt, L. R. Hill, M. E. Keeling, Y. Lu, J. E. Wright, T. C. Chou, and R. M. Ruprecht. 2000. Human neutralizing monoclonal antibodies of the IgG1 subtype protect against mucosal simian-human immunodeficiency virus infection. Nat Med **6:**200-6.

15. **Bailey, J. R., K. G. Lassen, H. C. Yang, T. C. Quinn, S. C. Ray, J. N. Blankson, and R. F. Siliciano.** 2006. Neutralizing antibodies do not mediate suppression of human immunodeficiency virus type 1 in elite suppressors or selection of plasma virus variants in patients on highly active antiretroviral therapy. J Virol **80:**4758-70.

16. **Baker, B. M., B. L. Block, A. C. Rothchild, and B. D. Walker.** 2009. Elite control of HIV infection: implications for vaccine design. Expert Opin Biol Ther **9:**55-69.

17. **Bank, T. W.** 2009, posting date. The World Bank: South Africa (http://data.worldbank.org/country/south-africa). The World Bank. [Online.]

18. **Barbaro, G., A. Scozzafava, A. Mastrolorenzo, and C. T. Supuran.** 2005. Highly active antiretroviral therapy: current state of the art, new agents and their pharmacological interactions useful for improving therapeutic outcome. Curr Pharm Des **11:**1805-43.

19. **Barouch, D. H., and B. Korber.** 2010. HIV-1 vaccine development after STEP. Annu Rev Med **61:**153-67.

20. **Bennett, D. E., R. J. Camacho, D. Otelea, D. R. Kuritzkes, H. Fleury, M. Kiuchi, W. Heneine, R. Kantor, M. R. Jordan, J. M. Schapiro, A. M. Vandamme, P. Sandstrom, C. A. Boucher, D. van de Vijver, S. Y. Rhee, T. F. Liu, D. Pillay, and R. W. Shafer.** 2009. Drug resistance mutations for

surveillance of transmitted HIV-1 drug-resistance: 2009 update. PLoS ONE **4:**e4724.

21. **Bennett, D. E., M. Myatt, S. Bertagnolio, D. Sutherland, and C. F. Gilks.** 2008. Recommendations for surveillance of transmitted HIV drug resistance in countries scaling up antiretroviral treatment. Antivir Ther **13 Suppl 2:**25-36.

22. **Bessong, P. O., J. Mphahlele, I. A. Choge, L. C. Obi, L. Morris, M. L. Hammarskjold, and D. M. Rekosh.** 2006. Resistance mutational analysis of HIV type 1 subtype C among rural South African drug-naive patients prior to large-scale availability of antiretrovirals. AIDS Res Hum Retroviruses **22:**1306-12.

23. **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.

24. **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.

25. **Brenchley, J. M., and D. C. Douek.** 2008. The mucosal barrier and immune activation in HIV pathogenesis. Curr Opin HIV AIDS **3:**356-61.

26. **Brenchley, J. M., T. W. Schacker, L. E. Ruff, D. A. Price, J. H. Taylor, G. J. Beilman, P. L. Nguyen, A. Khoruts, M. Larson, A. T. Haase, and D. C. Douek.** 2004. CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. J Exp Med **200:**749-59.

27. **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.

28. **Brockman, M. A., G. O. Tanzi, B. D. Walker, and T. M. Allen.** 2006. Use of a novel GFP reporter cell line to examine replication capacity of CXCR4- and CCR5-tropic HIV-1 by flow cytometry. J Virol Methods **131:**134-42.

29. **Brumme, Z., B. Wang, K. Nair, C. Brumme, C. de Pierres, S. Reddy, B. Julg, E. Moodley, C. Thobakgale, Z. Lu, M. van der Stok, K. Bishop, Z. Mncube, F. Chonco, Y. Yuki, N. Frahm, C. Brander, M. Carrington, K. Freedberg, P. Kiepiela, P. Goulder, B. Walker, T. Ndung'u, and E. Losina.** 2009. Impact of select immunologic and virologic biomarkers on CD4 cell count decrease in patients with chronic HIV-1 subtype C infection: results from Sinikithemba Cohort, Durban, South Africa. Clin Infect Dis **49:**956-64.

30. **Brumme, Z. L., C. J. Brumme, J. Carlson, H. Streeck, M. John, Q. Eichbaum, B. L. Block, B. Baker, C. Kadie, M. Markowitz, H. Jessen, A. D. Kelleher, E. Rosenberg, J. Kaldor, Y. Yuki, M. Carrington, T. M. Allen, S. Mallal, M. Altfeld, D. Heckerman, and B. D. Walker.** 2008. Marked epitope and allele-specific differences in rates of mutation in HIV-1 Gag, Pol and Nef CTL epitopes in acute/early HIV-1 infection. J Virol.

31. **Brumme, Z. L., C. J. Brumme, J. Carlson, H. Streeck, M. John, Q. Eichbaum, B. L. Block, B. Baker, C. Kadie, M. Markowitz, H. Jessen, A. D. Kelleher, E. Rosenberg, J. Kaldor, Y. Yuki, M. Carrington, T. M. Allen, S.**

Mallal, M. Altfeld, D. Heckerman, and B. D. Walker. 2008. Marked epitope- and allele-specific differences in rates of mutation in human immunodeficiency type 1 (HIV-1) Gag, Pol, and Nef cytotoxic T-lymphocyte epitopes in acute/early HIV-1 infection. J Virol **82:**9216-27.

32. **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.

33. **Brumme, Z. L., I. Tao, S. Szeto, C. J. Brumme, J. M. Carlson, D. Chan, C. Kadie, N. Frahm, C. Brander, B. Walker, D. Heckerman, and P. R. Harrigan.** 2008. Human leukocyte antigen-specific polymorphisms in HIV-1 Gag and their association with viral load in chronic untreated infection. AIDS **22:**1277-86.

34. **Buchbinder, S.** 2009. Clinical Endpoints in the Step Study. *In* B. Autran and S. Hammer (ed.), Keystone Symposium - HIV Vaccine. Keystone, Colorado.

35. **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, and M. N. Robertson.** 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.

36. **Burton, D. R., R. L. Stanfield, and I. A. Wilson.** 2005. Antibody vs. HIV in a clash of evolutionary titans. Proc Natl Acad Sci U S A **102:**14943-8.

37. **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.

38. **Carmichael, A., X. Jin, P. Sissons, and L. Borysiewicz.** 1993. Quantitative analysis of the human immunodeficiency virus type 1 (HIV-1)-specific cytotoxic T lymphocyte (CTL) response at different stages of HIV-1 infection: differential CTL responses to HIV-1 and Epstein-Barr virus in late disease. J Exp Med **177:**249-56.

39. **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.

40. **Carrington, M., and S. J. O'Brien.** 2003. The influence of HLA genotype on AIDS. Annu Rev Med **54:**535-51.

41. **Carvajal-Rodriguez, A., K. A. Crandall, and D. Posada.** 2007. Recombination favors the evolution of drug resistance in HIV-1 during antiretroviral therapy. Infect Genet Evol **7:**476-83.

42. **Choe, H., M. Farzan, Y. Sun, N. Sullivan, B. Rollins, P. D. Ponath, L. Wu, C. R. Mackay, G. LaRosa, W. Newman, N. Gerard, C. Gerard, and J. Sodroski.** 1996. The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. Cell **85:**1135-48.

43. **Cohen, M. S.** 2004. HIV and sexually transmitted diseases: lethal synergy. Top HIV Med **12:**104-7.

44. **Commons, R.-a. a. W.** 2010, posting date. South Africa (http://en.wikipedia.org/wiki/South_Africa). Multi-license with GFDL and Creative Commons CC-BY-SA-2.5 and older versions (2.0 and 1.0). [Online.]

45. **Corti, D., J. P. Langedijk, A. Hinz, M. S. Seaman, F. Vanzetta, B. M. Fernandez-Rodriguez, C. Silacci, D. Pinna, D. Jarrossay, S. Balla-Jhagjhoorsingh, B. Willems, M. J. Zekveld, H. Dreja, E. O'Sullivan, C. Pade, C. Orkin, S. A. Jeffs, D. C. Montefiori, D. Davis, W. Weissenhorn, A. McKnight, J. L. Heeney, F. Sallusto, Q. J. Sattentau, R. A. Weiss, and A. Lanzavecchia.** 2010. Analysis of memory B cell responses and isolation of novel monoclonal antibodies with neutralizing breadth from HIV-1-infected individuals. PLoS ONE **5:**e8805.

46. **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.

47. **Cross, A. H., J. L. Stark, J. Lauber, M. J. Ramsbottom, and J. A. Lyons.** 2006. Rituximab reduces B cells and T cells in cerebrospinal fluid of multiple sclerosis patients. J Neuroimmunol **180:**63-70.

48. **Dahab, M., S. Charalambous, R. Hamilton, K. Fielding, K. Kielmann, G. J. Churchyard, and A. D. Grant.** 2008. "That is why I stopped the ART": patients' & providers' perspectives on barriers to and enablers of HIV treatment adherence in a South African workplace programme. BMC Public Health **8:**63.

49. **de Oliveira, T., and S. Cassol.** 2007. BioAfrica - HIV Bioinformatics in Africa. South African National Bioinformatics Institute.

50. **de Oliveira, T., Deforche, K., Cassol, S., Rambaut, A., Vandamme, A.** 2007. REGA HIV Subtyping Tool Version 2.0. Developed in cooperation with the Evolutionary Biology Group at University of Oxford, UK., the HIV-1 Pathogenesis and Immunotherapeutics Program at University of Pretoria, South Africa, and the REGA Institute at the Katholieke Universiteit Leuven, Belgium.

51. **Department of Health, S. A.** 2003. Operational Plan for Comprehensive HIV and Aids Care, Management and Treatment for South Africa. *In* S. A. Department of Health (ed.).

52. **Dinges, W. L., J. Richardt, D. Friedrich, E. Jalbert, Y. Liu, C. E. Stevens, J. Maenza, A. C. Collier, D. E. Geraghty, J. Smith, Z. Moodie, J. I. Mullins, M. J. McElrath, and H. Horton.** 2010. Virus-specific CD8+ T-cell responses better define HIV disease progression than HLA genotype. J Virol **84:**4461-8.

53. **Dittmar, M. T., A. McKnight, G. Simmons, P. R. Clapham, R. A. Weiss, and P. Simmonds.** 1997. HIV-1 tropism and co-receptor use. Nature **385:**495-6.

54. **DoHaH, S.** 2009. Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. (http://wwwaidsinfonihgov/ContentFiles/AdultandAdolescentGLpdf), p. 1-161. *In* D. o. H. a. H. Services (ed.).

55. **Douek, D.** 2007. HIV disease progression: immune activation, microbes, and a leaky gut. Top HIV Med **15:**114-7.

56. **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.

57. **Dragic, T., V. Litwin, G. P. Allaway, S. R. Martin, Y. Huang, K. A. Nagashima, C. Cayanan, P. J. Maddon, R. A. Koup, J. P. Moore, and W. A. Paxton.** 1996. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. Nature **381:**667-73.

58. **Drummond, A. J., and A. Rambaut.** 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol **7:**214.

59. **Drummond, A. J., A. Rambaut, B. Shapiro, and O. G. Pybus.** 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol Biol Evol **22:**1185-92.

60. **Duda, A., L. Lee-Turner, J. Fox, N. Robinson, S. Dustan, S. Kaye, H. Fryer, M. Carrington, M. McClure, A. R. McLean, S. Fidler, J. Weber, R. E. Phillips, and A. J. Frater.** 2009. HLA-associated clinical progression correlates with epitope reversion rates in early human immunodeficiency virus infection. J Virol **83:**1228-39.

61. **Egger, M., B. Hirschel, P. Francioli, P. Sudre, M. Wirz, M. Flepp, M. Rickenbach, R. Malinverni, P. Vernazza, and M. Battegay.** 1997. Impact of new antiretroviral combination therapies in HIV infected patients in Switzerland: prospective multicentre study. Swiss HIV Cohort Study. Bmj **315:**1194-9.

62. **Fairall, L. R., M. O. Bachmann, G. M. Louwagie, C. van Vuuren, P. Chikobvu, D. Steyn, G. H. Staniland, V. Timmerman, M. Msimanga, C. J. Seebregts, A. Boulle, R. Nhiwatiwa, E. D. Bateman, M. F. Zwarenstein, and R. D. Chapman.** 2008. Effectiveness of antiretroviral treatment in a South African program: a cohort study. Arch Intern Med **168:**86-93.

63. **Fauci, A. S., M. I. Johnston, C. W. Dieffenbach, D. R. Burton, S. M. Hammer, J. A. Hoxie, M. Martin, J. Overbaugh, D. I. Watkins, A. Mahmoud, and W. C. Greene.** 2008. HIV vaccine research: the way forward. Science **321:**530-2.

64. **Feeney, M. E., Y. Tang, K. A. Roosevelt, A. J. Leslie, K. McIntosh, N. Karthas, B. D. Walker, and P. J. Goulder.** 2004. Immune escape precedes breakthrough human immunodeficiency virus type 1 viremia and broadening of the cytotoxic T-lymphocyte response in an HLA-B27-positive long-term-nonprogressing child. J Virol **78:**8927-30.

65. **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, and D. B. Goldstein.** 2009. Common genetic variation and the control of HIV-1 in humans. PLoS Genet **5:**e1000791.

66. **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.

67. **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.

68. **Fidler, S., A. Oxenius, M. Brady, J. Clarke, I. Cropley, A. Babiker, H. T. Zhang, D. Price, R. Phillips, and J. Weber.** 2002. Virological and immunological effects of short-course antiretroviral therapy in primary HIV infection. AIDS **16:**2049-54.

69. **Fiebig, E. W., D. J. Wright, B. D. Rawal, P. E. Garrett, R. T. Schumacher, L. Peddada, C. Heldebrant, R. Smith, A. Conrad, S. H. Kleinman, and M. P. Busch.** 2003. Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. AIDS **17:**1871-9.

70. **Frater, A. J., A. Beardall, K. Ariyoshi, D. Churchill, S. Galpin, J. R. Clarke, J. N. Weber, and M. O. McClure.** 2001. Impact of baseline polymorphisms in RT and protease on outcome of highly active antiretroviral therapy in HIV-1-infected African patients. Aids **15:**1493-502.

71. **Frater, A. J., H. Brown, A. Oxenius, H. F. Gunthard, B. Hirschel, N. Robinson, A. J. Leslie, R. Payne, H. Crawford, A. Prendergast, C. Brander, P. Kiepiela, B. D. Walker, P. J. Goulder, A. McLean, and R. E.**

Phillips. 2007. Effective T-cell responses select human immunodeficiency virus mutants and slow disease progression. J Virol **81:**6742-51.

72. **Frater, A. J., D. T. Dunn, A. J. Beardall, K. Ariyoshi, J. R. Clarke, M. O. McClure, and J. N. Weber.** 2002. Comparative response of African HIV-1-infected individuals to highly active antiretroviral therapy. Aids **16:**1139-46.

73. **Frater, A. J., C. T. Edwards, N. McCarthy, J. Fox, H. Brown, A. Milicic, N. Mackie, T. Pillay, J. W. Drijfhout, S. Dustan, J. R. Clarke, E. C. Holmes, H. T. Zhang, K. Pfafferott, P. J. Goulder, M. O. McClure, J. Weber, R. E. Phillips, and S. Fidler.** 2006. Passive sexual transmission of human immunodeficiency virus type 1 variants and adaptation in new hosts. J Virol **80:**7226-34.

74. **Gandhi, R. T., A. Wurcel, E. S. Rosenberg, M. N. Johnston, N. Hellmann, M. Bates, M. S. Hirsch, and B. D. Walker.** 2003. Progressive reversion of human immunodeficiency virus type 1 resistance mutations in vivo after transmission of a multiply drug-resistant virus. Clin Infect Dis **37:**1693-8.

75. **Gao, X., A. Bashirova, A. K. Iversen, J. Phair, J. J. Goedert, S. Buchbinder, K. Hoots, D. Vlahov, M. Altfeld, S. J. O'Brien, and M. Carrington.** 2005. AIDS restriction HLA allotypes target distinct intervals of HIV-1 pathogenesis. Nat Med **11:**1290-2.

76. **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.

77. **Garcia-Diaz, A., C. Blok, S. Madge, C. Booth, M. Tyrer, S. Bonora, T. Mahungu, A. Owen, M. Johnson, and A. M. Geretti.** 2008. Detection of

low-frequency K103N mutants after unstructured discontinuation of efavirenz in the presence of the CYP2B6 516 TT polymorphism. J Antimicrob Chemother **62:**1188-90.

78. **Garcia-Lerma, J. G., H. MacInnes, D. Bennett, H. Weinstock, and W. Heneine.** 2004. Transmitted human immunodeficiency virus type 1 carrying the D67N or K219Q/E mutation evolves rapidly to zidovudine resistance in vitro and shows a high replicative fitness in the presence of zidovudine. J Virol **78:**7545-52.

79. **Garcia-Rodriguez, M. J., M. A. Canales, D. Hernandez-Maraver, and F. Hernandez-Navarro.** 2008. Late reactivation of resolved hepatitis B virus infection: an increasing complication post rituximab-based regimens treatment? Am J Hematol **83:**673-5.

80. **Gatell, J. M.** 2010. When and why to start antiretroviral therapy? Journal of Antimicrobial Chemotherapy **65:**3.

81. **Gaufin, T., M. Pattison, R. Gautam, C. Stoulig, J. Dufour, J. MacFarland, D. Mandell, C. Tatum, M. H. Marx, R. M. Ribeiro, D. Montefiori, C. Apetrei, and I. Pandrea.** 2009. Effect of B-cell depletion on viral replication and clinical outcome of simian immunodeficiency virus infection in a natural host. J Virol **83:**10347-57.

82. **Gervaix, A., D. West, L. M. Leoni, D. D. Richman, F. Wong-Staal, and J. Corbeil.** 1997. A new reporter cell line to monitor HIV infection and drug susceptibility in vitro. Proc Natl Acad Sci U S A **94:**4653-8.

83. **Geskus, R. B., M. Prins, J. B. Hubert, F. Miedema, B. Berkhout, C. Rouzioux, J. F. Delfraissy, and L. Meyer.** 2007. The HIV RNA setpoint theory revisited. Retrovirology **4:**65.

84. **Gifford, R. J., T. de Oliveira, A. Rambaut, O. G. Pybus, D. Dunn, A. M. Vandamme, P. Kellam, and D. Pillay.** 2007. Phylogenetic surveillance of viral genetic diversity and the evolving molecular epidemiology of human immunodeficiency virus type 1. J Virol **81:**13050-6.

85. **Gifford, R. J., A. Katzourakis, M. Tristem, O. G. Pybus, M. Winters, and R. W. Shafer.** 2008. A transitional endogenous lentivirus from the genome of a basal primate and implications for lentivirus evolution. Proc Natl Acad Sci U S A **105:**20362-7.

86. **Gilks, C. F., S. Crowley, R. Ekpini, S. Gove, J. Perriens, Y. Souteyrand, D. Sutherland, M. Vitoria, T. Guerma, and K. De Cock.** 2006. The WHO public-health approach to antiretroviral treatment against HIV in resource-limited settings. Lancet **368:**505-10.

87. **Goepfert, P. A., W. Lumm, P. Farmer, P. Matthews, A. Prendergast, J. M. Carlson, C. A. Derdeyn, J. Tang, R. A. Kaslow, A. Bansal, K. Yusim, D. Heckerman, J. Mulenga, S. Allen, P. J. Goulder, and E. Hunter.** 2008. Transmission of HIV-1 Gag immune escape mutations is associated with reduced viral load in linked recipients. J Exp Med **205:**1009-17.

88. **Gomez, C., and T. J. Hope.** 2005. The ins and outs of HIV replication. Cell Microbiol **7:**621-6.

89. **Gonzalez, N., A. Alvarez, and J. Alcami.** 2010. Broadly Neutralizing Antibodies and their Significance for HIV-1 Vaccines. Curr HIV Res.

90. **Gordon, M., T. De Oliveira, K. Bishop, H. M. Coovadia, L. Madurai, S. Engelbrecht, E. Janse van Rensburg, A. Mosam, A. Smith, and S. Cassol.** 2003. Molecular characteristics of human immunodeficiency virus type 1

subtype C viruses from KwaZulu-Natal, South Africa: implications for vaccine and antiretroviral control strategies. J Virol **77:**2587-99.

91. **Goulder, P., D. Price, M. Nowak, S. Rowland-Jones, R. Phillips, and A. McMichael.** 1997. Co-evolution of human immunodeficiency virus and cytotoxic T-lymphocyte responses. Immunol Rev **159:**17-29.

92. **Goulder, P. J., M. A. Altfeld, E. S. Rosenberg, T. Nguyen, Y. Tang, R. L. Eldridge, M. M. Addo, S. He, J. S. Mukherjee, M. N. Phillips, M. Bunce, S. A. Kalams, R. P. Sekaly, B. D. Walker, and C. Brander.** 2001. Substantial differences in specificity of HIV-specific cytotoxic T cells in acute and chronic HIV infection. J Exp Med **193:**181-94.

93. **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.

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

95. **Goulder, P. J., and D. I. Watkins.** 2008. Impact of MHC class I diversity on immune control of immunodeficiency virus replication. Nat Rev Immunol **8:**619-30.

96. **Greene, W. C., and B. M. Peterlin.** 2002. Charting HIV's remarkable voyage through the cell: Basic science as a passport to future therapy. Nat Med **8:**673-80.

97. **Grenfell, B. T., O. G. Pybus, J. R. Gog, J. L. Wood, J. M. Daly, J. A. Mumford, and E. C. Holmes.** 2004. Unifying the epidemiological and evolutionary dynamics of pathogens. Science **303:**327-32.

98. **Grossman, Z., V. Istomin, D. Averbuch, M. Lorber, K. Risenberg, I. Levi, M. Chowers, M. Burke, N. Bar Yaacov, and J. M. Schapiro.** 2004. Genetic variation at NNRTI resistance-associated positions in patients infected with HIV-1 subtype C. Aids **18:**909-15.

99. **Grossman, Z., M. Meier-Schellersheim, W. E. Paul, and L. J. Picker.** 2006. Pathogenesis of HIV infection: what the virus spares is as important as what it destroys. Nat Med **12:**289-95.

100. **Guindon, S., and O. Gascuel.** 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic Biology**:**8.

101. **Hahn, B. H., G. M. Shaw, K. M. De Cock, and P. M. Sharp.** 2000. AIDS as a zoonosis: scientific and public health implications. Science **287:**607-14.

102. **Hay, C. M., D. J. Ruhl, N. O. Basgoz, C. C. Wilson, J. M. Billingsley, M. P. DePasquale, R. T. D'Aquila, S. M. Wolinsky, J. M. Crawford, D. C. Montefiori, and B. D. Walker.** 1999. Lack of viral escape and defective in vivo activation of human immunodeficiency virus type 1-specific cytotoxic T lymphocytes in rapidly progressive infection. J Virol **73:**5509-19.

103. **Hecht, F. M., W. Hartogensis, L. Bragg, P. Bacchetti, R. Atchison, R. Grant, J. Barbour, and S. G. Deeks.** 2010. HIV RNA level in early infection is predicted by viral load in the transmission source. AIDS **24:**941-5.

104. **Hellerstein, M. K., R. A. Hoh, M. B. Hanley, D. Cesar, D. Lee, R. A. Neese, and J. M. McCune.** 2003. Subpopulations of long-lived and short-lived T cells in advanced HIV-1 infection. J Clin Invest **112:**956-66.

105. **Herbeck, J. T., G. S. Gottlieb, C. A. Winkler, G. W. Nelson, P. An, B. S. Maust, K. G. Wong, J. L. Troyer, J. J. Goedert, B. D. Kessing, R. Detels, S. M. Wolinsky, J. Martinson, S. Buchbinder, G. D. Kirk, L. P. Jacobson, J. B. Margolick, R. A. Kaslow, S. J. O'Brien, and J. I. Mullins.** Multistage genomewide association study identifies a locus at 1q41 associated with rate of HIV-1 disease progression to clinical AIDS. J Infect Dis **201:**618-26.

106. **Hessell, A. J., E. G. Rakasz, D. M. Tehrani, M. Huber, K. L. Weisgrau, G. Landucci, D. N. Forthal, W. C. Koff, P. Poignard, D. I. Watkins, and D. R. Burton.** 2010. Broadly neutralizing monoclonal antibodies 2F5 and 4E10 directed against the human immunodeficiency virus type 1 gp41 membrane-proximal external region protect against mucosal challenge by simian-human immunodeficiency virus SHIVBa-L. J Virol **84:**1302-13.

107. **Hirsch, V. M.** 2004. What can natural infection of African monkeys with simian immunodeficiency virus tell us about the pathogenesis of AIDS? AIDS Rev **6:**40-53.

108. **HIV/AIDS, N. C. f. I. D. D. o., K. G. Castro, J. W. Ward, L. Slutsker, J. W. Buehler, H. W. Jaffe, and R. L. Berkelman.** 1993. 1993 Revised Classification System for HIV Infection and Expanded Surveillance Case Definition for AIDS Among Adolescents and Adults. Recommendation and Reports. Centers for Disease Ccontrol and Prevention (CDC).

109. **Holl, V., M. Peressin, T. Decoville, S. Schmidt, S. Zolla-Pazner, A. M. Aubertin, and C. Moog.** 2006. Nonneutralizing antibodies are able to inhibit human immunodeficiency virus type 1 replication in macrophages and immature dendritic cells. J Virol **80:**6177-81.

110. **Holl, V., M. Peressin, S. Schmidt, T. Decoville, S. Zolla-Pazner, A. M. Aubertin, and C. Moog.** 2006. Efficient inhibition of HIV-1 replication in human immature monocyte-derived dendritic cells by purified anti-HIV-1 IgG without induction of maturation. Blood **107:**4466-74.

111. **Huang, K. H., D. Goedhals, H. Fryer, C. van Vuuren, A. Katzourakis, T. De Oliveira, H. Brown, S. Cassol, C. Seebregts, A. McLean, P. Klenerman, R. Phillips, and J. Frater.** 2009. Prevalence of HIV type-1 drug-associated mutations in pre-therapy patients in the Free State, South Africa. Antivir Ther **14:**975-84.

112. **Huang, K. H. G., D. Bonsall, A. Katzourakis, E. Thomson, S. J. Fidler, J. Main, D. Muir, J. N. Weber, A. J. Frater, R. E. Phillips, O. Pybus, P. Goulder, M. O. McClure, G. S. Cooke, and P. Klenerman.** 2010. B cell depletion in vivo reveals a role for antibody control of HIV. Nature Communication.

113. **Hutnick, N. A., D. G. Carnathan, S. A. Dubey, G. Makedonas, K. S. Cox, L. Kierstead, S. J. Ratcliffe, M. N. Robertson, D. R. Casimiro, H. C. Ertl, and M. R. Betts.** 2009. Baseline Ad5 serostatus does not predict Ad5 HIV vaccine-induced expansion of adenovirus-specific CD4+ T cells. Nat Med **15:**876-8.

114. **Janeway, C., K. M. Murphy, P. Travers, and M. Walport.** 2005. Immunobiology - The Immune System in Health and Disease, 6th ed. Garland Science, New York and London.

115. **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.

116. **Jin, X., M. Wills, J. G. Sissons, and A. Carmichael.** 1998. Progressive loss of IL-2-expandable HIV-1-specific cytotoxic T lymphocytes during asymptomatic HIV infection. Eur J Immunol **28:**3564-76.

117. **Johnson, V. A., F. Brun-Vezinet, B. Clotet, H. F. Gunthard, D. R. Kuritzkes, D. Pillay, J. M. Schapiro, and D. D. Richman.** 2008. Update of the Drug Resistance Mutations in HIV-1. Top HIV Med **16:**138-45.

118. **Jolles, S., M. Tyrer, M. Johnson, and D. Webster.** 2001. Long term recovery of IgG and IgM production during HIV infection in a patient with common variable immunodeficiency (CVID). J Clin Pathol **54:**713-5.

119. **Jourdain, G., N. Ngo-Giang-Huong, S. Le Coeur, C. Bowonwatanuwong, P. Kantipong, P. Leechanachai, S. Ariyadej, P. Leenasirimakul, S. Hammer, and M. Lallemant.** 2004. Intrapartum exposure to nevirapine and subsequent maternal responses to nevirapine-based antiretroviral therapy. N Engl J Med **351:**229-40.

120. **Kalams, S. A., S. P. Buchbinder, E. S. Rosenberg, J. M. Billingsley, D. S. Colbert, N. G. Jones, A. K. Shea, A. K. Trocha, and B. D. Walker.** 1999. Association between virus-specific cytotoxic T-lymphocyte and helper responses in human immunodeficiency virus type 1 infection. J Virol **73:**6715-20.

121. **Kantor, R., D. A. Katzenstein, B. Efron, A. P. Carvalho, B. Wynhoven, P. Cane, J. Clarke, S. Sirivichayakul, M. A. Soares, J. Snoeck, C. Pillay, H. Rudich, R. Rodrigues, A. Holguin, K. Ariyoshi, M. B. Bouzas, P. Cahn, W. Sugiura, V. Soriano, L. F. Brigido, Z. Grossman, L. Morris, A. M.**

Vandamme, A. Tanuri, P. Phanuphak, J. N. Weber, D. Pillay, P. R. Harrigan, R. Camacho, J. M. Schapiro, and R. W. Shafer. 2005. Impact of HIV-1 subtype and antiretroviral therapy on protease and reverse transcriptase genotype: results of a global collaboration. PLoS Med **2:**e112.

122. **Karlsson, A. C., S. R. Younger, J. N. Martin, Z. Grossman, E. Sinclair, P. W. Hunt, E. Hagos, D. F. Nixon, and S. G. Deeks.** 2004. Immunologic and virologic evolution during periods of intermittent and persistent low-level viremia. AIDS **18:**981-9.

123. **Kaslow, R. A., M. Carrington, R. Apple, L. Park, A. Munoz, A. J. Saah, J. J. Goedert, C. Winkler, S. J. O'Brien, C. Rinaldo, R. Detels, W. Blattner, J. Phair, H. Erlich, and D. L. Mann.** 1996. Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection. Nat Med **2:**405-11.

124. **Kaufmann, G. R., P. Cunningham, A. D. Kelleher, J. Zaunders, A. Carr, J. Vizzard, M. Law, and D. A. Cooper.** 1998. Patterns of viral dynamics during primary human immunodeficiency virus type 1 infection. The Sydney Primary HIV Infection Study Group. J Infect Dis **178:**1812-5.

125. **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.

126. **Keele, B. F., E. E. Giorgi, J. F. Salazar-Gonzalez, J. M. Decker, K. T. Pham, M. G. Salazar, C. Sun, T. Grayson, S. Wang, H. Li, X. Wei, C. Jiang, J. L. Kirchherr, F. Gao, J. A. Anderson, L. H. Ping, R. Swanstrom, G. D. Tomaras, W. A. Blattner, P. A. Goepfert, J. M. Kilby, M. S. Saag, E. L. Delwart, M. P. Busch, M. S. Cohen, D. C. Montefiori, B. F. Haynes, B. Gaschen, G. S. Athreya, H. Y. Lee, N. Wood, C. Seoighe, A. S. Perelson, T. Bhattacharya, B. T. Korber, B. H. Hahn, and G. M. Shaw.** 2008. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. Proc Natl Acad Sci U S A **105:**7552-7.

127. **Kelleher, A. D., C. Long, E. C. Holmes, R. L. Allen, J. Wilson, C. Conlon, C. Workman, S. Shaunak, K. Olson, P. Goulder, C. Brander, G. Ogg, J. S. Sullivan, W. Dyer, I. Jones, A. J. McMichael, S. Rowland-Jones, and R. E. Phillips.** 2001. Clustered mutations in HIV-1 gag are consistently required for escape from HLA-B27-restricted cytotoxic T lymphocyte responses. J Exp Med **193:**375-86.

128. **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.

129. **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.

130. **Klenerman, P., and A. Hill.** 2005. T cells and viral persistence: lessons from diverse infections. Nat Immunol **6:**873-9.

131. **Klotman, M. E., S. Kim, A. Buchbinder, A. DeRossi, D. Baltimore, and F. Wong-Staal.** 1991. Kinetics of expression of multiply spliced RNA in early human immunodeficiency virus type 1 infection of lymphocytes and monocytes. Proc Natl Acad Sci U S A **88:**5011-5.

132. **Koff, W. C., and S. F. Berkley.** 2010. The Renaissance in HIV Vaccine Development - Future Directions. N Engl J Med **363:**e7.

133. **Koibuchi, T., T. M. Allen, M. Lichterfeld, S. K. Mui, K. M. O'Sullivan, A. Trocha, S. A. Kalams, R. P. Johnson, and B. D. Walker.** 2005. Limited sequence evolution within persistently targeted CD8 epitopes in chronic human immunodeficiency virus type 1 infection. J Virol **79:**8171-81.

134. **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.

135. **Korber, B., M. Muldoon, J. Theiler, F. Gao, R. Gupta, A. Lapedes, B. H. Hahn, S. Wolinsky, and T. Bhattacharya.** 2000. Timing the ancestor of the HIV-1 pandemic strains. Science **288:**1789-96.

136. **Kosmrlj, A., E. L. Read, Y. Qi, T. M. Allen, M. Altfeld, S. G. Deeks, F. Pereyra, M. Carrington, B. D. Walker, and A. K. Chakraborty.** 2010. Effects of thymic selection of the T-cell repertoire on HLA class I-associated control of HIV infection. Nature.

137. **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.

138. **Kowalski, M., J. Potz, L. Basiripour, T. Dorfman, W. C. Goh, E. Terwilliger, A. Dayton, C. Rosen, W. Haseltine, and J. Sodroski.** 1987. Functional regions of the envelope glycoprotein of human immunodeficiency virus type 1. Science **237:**1351-5.

139. **Kwong, P. D., R. Wyatt, J. Robinson, R. W. Sweet, J. Sodroski, and W. A. Hendrickson.** 1998. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. Nature **393:**648-59.

140. **Larder, B. A., P. Kellam, and S. D. Kemp.** 1993. Convergent combination therapy can select viable multidrug-resistant HIV-1 in vitro. Nature **365:**451-3.

141. **Larder, B. A., S. D. Kemp, and P. R. Harrigan.** 1995. Potential mechanism for sustained antiretroviral efficacy of AZT-3TC combination therapy. Science **269:**696-9.

142. **Lee, S. K., Z. Xu, J. Lieberman, and P. Shankar.** 2002. The functional CD8 T cell response to HIV becomes type-specific in progressive disease. J Clin Invest **110:**1339-47.

143. **Lemey, P., O. G. Pybus, B. Wang, N. K. Saksena, M. Salemi, and A. M. Vandamme.** 2003. Tracing the origin and history of the HIV-2 epidemic. Proc Natl Acad Sci U S A **100:**6588-92.

144. **Leslie, A., D. Kavanagh, I. Honeyborne, K. Pfafferott, C. Edwards, T. Pillay, L. Hilton, C. Thobakgale, D. Ramduth, R. Draenert, S. Le Gall, G. Luzzi, A. Edwards, C. Brander, A. K. Sewell, S. Moore, J. Mullins, C. Moore, S. Mallal, N. Bhardwaj, K. Yusim, R. Phillips, P. Klenerman, B. Korber, P. Kiepiela, B. Walker, and P. Goulder.** 2005. Transmission and accumulation of CTL escape variants drive negative associations between HIV polymorphisms and HLA. J Exp Med **201:**891-902.

145. **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.

146. **Letvin, N. L.** 2006. Progress and obstacles in the development of an AIDS vaccine. Nat Rev Immunol **6:**930-9.

147. **Levy, J. A.** 2007. HIV and the Pathogenesis of AIDS, 3 ed. ASM Press, Washington, DC.

148. **Li, Q., L. Duan, J. D. Estes, Z. M. Ma, T. Rourke, Y. Wang, C. Reilly, J. Carlis, C. J. Miller, and A. T. Haase.** 2005. Peak SIV replication in resting memory CD4+ T cells depletes gut lamina propria CD4+ T cells. Nature **434:**1148-52.

149. **Liossis, S. N., and P. P. Sfikakis.** 2008. Rituximab-induced B cell depletion in autoimmune diseases: potential effects on T cells. Clin Immunol **127:**280-5.

150. **Los Alamos National Laboratory, L.** 2010. HIV Databases. Los Alamos National Security, LLC.

151. **Malim, M. H., and M. Emerman.** 2001. HIV-1 sequence variation: drift, shift, and attenuation. Cell **104:**469-72.

152. **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.

153. **Mascola, J. R., and D. C. Montefiori.** 2003. HIV-1: nature's master of disguise. Nat Med **9:**393-4.

154. **Matano, T., R. Shibata, C. Siemon, M. Connors, H. C. Lane, and M. A. Martin.** 1998. Administration of an anti-CD8 monoclonal antibody interferes with the clearance of chimeric simian/human immunodeficiency virus during primary infections of rhesus macaques. J Virol **72:**164-9.

155. **Matloubian, M., R. J. Concepcion, and R. Ahmed.** 1994. CD4+ T cells are required to sustain CD8+ cytotoxic T-cell responses during chronic viral infection. J Virol **68:**8056-63.

156. **Matthews, P. C., A. Prendergast, A. Leslie, H. Crawford, R. Payne, C. Rousseau, M. Rolland, I. Honeyborne, J. Carlson, C. Kadie, C. Brander, K. Bishop, N. Mlotshwa, J. I. Mullins, H. Coovadia, T. Ndung'u, B. D. Walker, D. Heckerman, and P. J. Goulder.** 2008. Central role of reverting

mutations in HLA associations with human immunodeficiency virus set point. J Virol **82:**8548-59.

157. **McKeating, J. A., J. Gow, J. Goudsmit, L. H. Pearl, C. Mulder, and R. A. Weiss.** 1989. Characterization of HIV-1 neutralization escape mutants. AIDS **3:**777-84.

158. **McMichael, A. J., and T. Hanke.** 2003. HIV vaccines 1983-2003. Nat Med **9:**874-80.

159. **Mease, P. J.** 2008. B cell-targeted therapy in autoimmune disease: rationale, mechanisms, and clinical application. J Rheumatol **35:**1245-55.

160. **Melchior, J. C., T. Niyongabo, D. Henzel, I. Durack-Bown, S. C. Henri, and A. Boulier.** 1999. Malnutrition and wasting, immunodepression, and chronic inflammation as independent predictors of survival in HIV-infected patients. Nutrition **15:**865-9.

161. **Mellors, J. W., A. Munoz, J. V. Giorgi, J. B. Margolick, C. J. Tassoni, P. Gupta, L. A. Kingsley, J. A. Todd, A. J. Saah, R. Detels, J. P. Phair, and C. R. Rinaldo, Jr.** 1997. Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection. Ann Intern Med **126:**946-54.

162. **Mellors, J. W., C. R. Rinaldo, Jr., P. Gupta, R. M. White, J. A. Todd, and L. A. Kingsley.** 1996. Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. Science **272:**1167-70.

163. **Meyer-Olson, D., K. W. Brady, M. T. Bartman, K. M. O'Sullivan, B. C. Simons, J. A. Conrad, C. B. Duncan, S. Lorey, A. Siddique, R. Draenert, M. Addo, M. Altfeld, E. Rosenberg, T. M. Allen, B. D. Walker, and S. A. Kalams.** 2006. Fluctuations of functionally distinct CD8+ T-cell clonotypes demonstrate flexibility of the HIV-specific TCR repertoire. Blood **107:**2373-83.

164. **Miller, C., M. Genesca, K. Abel, D. Montefiori, D. Forthal, K. Bost, J. Li, D. Favre, and J. McCune.** 2007. Antiviral antibodies are necessary for control of simian immunodeficiency virus replication. Journal of Virology **81:**12.

165. **Miura, T., M. A. Brockman, Z. L. Brumme, C. J. Brumme, F. Pereyra, A. Trocha, B. L. Block, A. Schneidewind, T. M. Allen, D. Heckerman, and B. D. Walker.** 2009. HLA-associated alterations in replication capacity of chimeric NL4-3 viruses carrying gag-protease from elite controllers of human immunodeficiency virus type 1. J Virol **83:**140-9.

166. **Miura, T., Z. Brumme, M. A. Brockman, P. Rosato, J. Sela, C. Brumme, F. Pereyra, D. E. Kaufmann, A. Trocha, B. L. Block, E. S. Daar, E. Cornnick, H. Jessen, A. D. Kelleher, E. Rosenberg, M. Markowitz, K. Schafer, F. Vaida, A. Iwamoto, S. Little, and B. D. Walker.** 2010. Impaired replication capacity of acute/early viruses in persons who become HIV controllers. Journal of Virology **84:**10.

167. **Moir, S., and A. S. Fauci.** 2009. B cells in HIV infection and disease. Nat Rev Immunol **9:**235-45.

168. **Montefiori, D., Q. Sattentau, J. Flores, J. Esparza, and J. Mascola.** 2007. Antibody-based HIV-1 vaccines: recent developments and future directions. PLoS Med **4:**e348.

169. **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.

170. **Moore, G. E., R. E. Gerner, and H. A. Franklin.** 1967. Culture of normal human leukocytes. JAMA **199:**519-24.

171. **Nagot, N., A. Ouedraogo, V. Foulongne, I. Konate, H. A. Weiss, L. Vergne, M. C. Defer, D. Djagbare, A. Sanon, J. B. Andonaba, P. Becquart, M. Segondy, R. Vallo, A. Sawadogo, P. Van de Perre, and P. Mayaud.** 2007. Reduction of HIV-1 RNA levels with therapy to suppress herpes simplex virus. N Engl J Med **356:**790-9.

172. **Ndung'u, T., M. Brockman, T. Allen, Z. Brumme, and A. J. Frater.** 2009. Challenges of engineering a complete subtype C chemic virus for the use of in vitro viral growth assay.

173. **Ndung'u, T., B. Renjifo, and M. Essex.** 2001. Construction and analysis of an infectious human Immunodeficiency virus type 1 subtype C molecular clone. J Virol **75:**4964-72.

174. **Ng, C. T., J. P. Jaworski, P. Jayaraman, W. F. Sutton, P. Delio, L. Kuller, D. Anderson, G. Landucci, B. A. Richardson, D. R. Burton, D. N. Forthal, and N. L. Haigwood.** 2010. Passive neutralizing antibody controls SHIV viremia and enhances B cell responses in infant macaques. Nat Med **16:**1117-9.

175. **Novak, R. M., L. Chen, R. D. MacArthur, J. D. Baxter, K. Huppler Hullsiek, G. Peng, Y. Xiang, C. Henely, B. Schmetter, J. Uy, M. van den Berg-Wolf, and M. Kozal.** 2005. Prevalence of antiretroviral drug resistance mutations in chronically HIV-infected, treatment-naive patients: implications for routine resistance screening before initiation of antiretroviral therapy. Clin Infect Dis **40:**468-74.

176. **O'Brien, K. L., J. Liu, S. L. King, Y. H. Sun, J. E. Schmitz, M. A. Lifton, N. A. Hutnick, M. R. Betts, S. A. Dubey, J. Goudsmit, J. W. Shiver, M. N. Robertson, D. R. Casimiro, and D. H. Barouch.** 2009. Adenovirus-specific

immunity after immunization with an Ad5 HIV-1 vaccine candidate in humans. Nat Med **15:**873-5.

177. **O'Brien, S. J., X. Gao, and M. Carrington.** 2001. HLA and AIDS: a cautionary tale. Trends Mol Med **7:**379-81.

178. **Oelrichs, R. B., V. A. Lawson, K. M. Coates, C. Chatfield, N. J. Deacon, and D. A. McPhee.** 2000. Rapid full-length genomic sequencing of two cytopathically heterogeneous Australian primary HIV-1 isolates. J Biomed Sci **7:**128-35.

179. **Padiglione, A., E. Aleksic, M. French, A. Arnott, K. M. Wilson, E. Tippett, M. Kaye, L. Gray, A. Ellett, M. Crane, D. E. Leslie, S. R. Lewin, A. Breschkin, C. Birch, P. R. Gorry, D. A. McPhee, and S. M. Crowe.** 2010. Extremely prolonged HIV seroconversion associated with an MHC haplotype carrying disease susceptibility genes for antibody deficiency disorders. Clin Immunol **137:**199-208.

180. **Palella, F. J., Jr., K. M. Delaney, A. C. Moorman, M. O. Loveless, J. Fuhrer, G. A. Satten, D. J. Aschman, and S. D. Holmberg.** 1998. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. N Engl J Med **338:**853-60.

181. **Pantaleo, G.** 2008. HIV-1 T-cell vaccines: evaluating the next step. Lancet Infect Dis **8:**82-3.

182. **Pantaleo, G., C. Graziosi, and A. S. Fauci.** 1993. New concepts in the immunopathogenesis of human immunodeficiency virus infection. N Engl J Med **328:**327-35.

183. **Parren, P. W., D. R. Burton, and Q. J. Sattentau.** 1997. HIV-1 antibody--debris or virion? Nat Med **3:**366-7.

184. **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.

185. **Phillips, R. E., S. Rowland-Jones, D. F. Nixon, F. M. Gotch, J. P. Edwards, A. O. Ogunlesi, J. G. Elvin, J. A. Rothbard, C. R. Bangham, C. R. Rizza, and et al.** 1991. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. Nature **354:**453-9.

186. **Pieniazek, D., M. Rayfield, D. J. Hu, J. Nkengasong, S. Z. Wiktor, R. Downing, B. Biryahwaho, T. Mastro, A. Tanuri, V. Soriano, R. Lal, and T. Dondero.** 2000. Protease sequences from HIV-1 group M subtypes A-H reveal distinct amino acid mutation patterns associated with protease resistance in protease inhibitor-naive individuals worldwide. HIV Variant Working Group. Aids **14:**1489-95.

187. **Pillay, C., H. Bredell, J. McIntyre, G. Gray, and L. Morris.** 2002. HIV-1 subtype C reverse transcriptase sequences from drug-naive pregnant women in South Africa. AIDS Res Hum Retroviruses **18:**605-10.

188. **Pillay, D.** 2007. The priorities for antiviral drug resistance surveillance and research. J Antimicrob Chemother **60 Suppl 1:**i57-8.

189. **Poignard, P., R. Sabbe, G. R. Picchio, M. Wang, R. J. Gulizia, H. Katinger, P. W. Parren, D. E. Mosier, and D. R. Burton.** 1999. Neutralizing antibodies

have limited effects on the control of established HIV-1 infection in vivo. Immunity **10:**431-8.

190. **Prado, J. G., T. Wrin, J. Beauchaine, L. Ruiz, C. J. Petropoulos, S. D. Frost, B. Clotet, R. T. D'Aquila, and J. Martinez-Picado.** 2002. Amprenavir-resistant HIV-1 exhibits lopinavir cross-resistance and reduced replication capacity. AIDS **16:**1009-17.

191. **Rambaut, A., D. Posada, K. A. Crandall, and E. C. Holmes.** 2004. The causes and consequences of HIV evolution. Nat Rev Genet **5:**52-61.

192. **Rangel-Moreno, J., D. M. Carragher, R. S. Misra, K. Kusser, L. Hartson, A. Moquin, F. E. Lund, and T. D. Randall.** 2008. B cells promote resistance to heterosubtypic strains of influenza via multiple mechanisms. J Immunol **180:**454-63.

193. **Report, N.** 2007. HIV vaccine failure prompts Merck to halt trial. Nature **449:**390.

194. **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, and J. H. Kim.** 2009. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. N Engl J Med **361:**2209-20.

195. **Richman, D. D., S. C. Morton, T. Wrin, N. Hellmann, S. Berry, M. F. Shapiro, and S. A. Bozzette.** 2004. The prevalence of antiretroviral drug resistance in the United States. Aids **18:**1393-401.

196. **Richman, D. D., T. Wrin, S. J. Little, and C. J. Petropoulos.** 2003. Rapid evolution of the neutralizing antibody response to HIV type 1 infection. Proc Natl Acad Sci U S A **100:**4144-9.

197. **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.

198. **Rockstroh, J. K.** 2006. Influence of viral hepatitis on HIV infection. J Hepatol **44:**S25-7.

199. **Rosenberg, E. S., J. M. Billingsley, A. M. Caliendo, S. L. Boswell, P. E. Sax, S. A. Kalams, and B. D. Walker.** 1997. Vigorous HIV-1-specific CD4+ T cell responses associated with control of viremia. Science **278:**1447-50.

200. **Rousseau, C. M., M. G. Daniels, J. M. Carlson, C. Kadie, H. Crawford, A. Prendergast, P. Matthews, R. Payne, M. Rolland, D. N. Raugi, B. S. Maust, G. H. Learn, D. C. Nickle, H. Coovadia, T. Ndung'u, N. Frahm, C. Brander, B. D. Walker, P. J. Goulder, T. Bhattacharya, D. E. Heckerman, B. T. Korber, and J. I. Mullins.** 2008. HLA class I-driven evolution of human immunodeficiency virus type 1 subtype c proteome: immune escape and viral load. J Virol **82:**6434-46.

201. **S. Georgia Centre for AIDS Research, D. U. h. w. d. o. S. D.** 2005. Powerpoint presentation.

202. **Saksena, N. K., B. Rodes, B. Wang, and V. Soriano.** 2007. Elite HIV controllers: myth or reality? AIDS Rev **9:**195-207.

203. **Sanches, M., S. Krauchenco, N. H. Martins, A. Gustchina, A. Wlodawer, and I. Polikarpov.** 2007. Structural characterization of B and non-B subtypes of HIV-protease: insights into the natural susceptibility to drug resistance development. J Mol Biol **369:**1029-40.

204. **Sarngadharan, M. G., M. Popovic, L. Bruch, J. Schupbach, and R. C. Gallo.** 1984. Antibodies reactive with human T-lymphotropic retroviruses (HTLV-III) in the serum of patients with AIDS. Science **224:**506-8.

205. **Sattentau, Q.** 2008. Correlates of antibody-mediated protection against HIV infection. Curr Opin HIV AIDS **3:**368-74.

206. **Scanlan, C. N., R. Pantophlet, M. R. Wormald, E. Ollmann Saphire, R. Stanfield, I. A. Wilson, H. Katinger, R. A. Dwek, P. M. Rudd, and D. R. Burton.** 2002. The broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2G12 recognizes a cluster of alpha1-->2 mannose residues on the outer face of gp120. J Virol **76:**7306-21.

207. **Scheid, J. F., H. Mouquet, N. Feldhahn, M. S. Seaman, K. Velinzon, J. Pietzsch, R. G. Ott, R. M. Anthony, H. Zebroski, A. Hurley, A. Phogat, B. Chakrabarti, Y. Li, M. Connors, F. Pereyra, B. D. Walker, H. Wardemann, D. Ho, R. T. Wyatt, J. R. Mascola, J. V. Ravetch, and M. C. Nussenzweig.** 2009. Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. Nature **458:**636-40.

208. **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.

209. **Schmitz, J. E., M. J. Kuroda, S. Santra, M. A. Simon, M. A. Lifton, W. Lin, R. Khunkhun, M. Piatak, J. D. Lifson, G. Grosschupff, R. S. Gelman, P. Racz, K. Tenner-Racz, K. A. Mansfield, N. L. Letvin, D. C. Montefiori, and K. A. Reimann.** 2003. Effect of humoral immune responses on controlling viremia during primary infection of rhesus monkeys with simian immunodeficiency virus. J Virol **77:**2165-73.

210. **Seebregts, C. J., C. van Vuuren, D. Goedhals, T. de Oliveira, P. Drew, P. Makhoahle, R. Nhiwatiwa, V. Timmerman, F. L., E. Kotze, R. Chapman, and S. Cassol.** 5-8 June 2007. Baseline Characterization and Resistance Genotyping of HIV from the Public Sector Antiretroviral Treatment Program in the Free State Province of South Africa, The Third South African AIDS Conference, Durban, South Africa.

211. **Shafer, R. W., S. Y. Rhee, and D. E. Bennett.** 2008. Consensus drug resistance mutations for epidemiological surveillance: basic principles and potential controversies. Antivir Ther **13 Suppl 2:**59-68.

212. **Shafer, R. W., S. Y. Rhee, D. Pillay, V. Miller, P. Sandstrom, J. M. Schapiro, D. R. Kuritzkes, and D. Bennett.** 2007. HIV-1 protease and reverse transcriptase mutations for drug resistance surveillance. Aids **21:**215-23.

213. **Shankarappa, R., J. B. Margolick, S. J. Gange, A. G. Rodrigo, D. Upchurch, H. Farzadegan, P. Gupta, C. R. Rinaldo, G. H. Learn, X. He, X. L. Huang, and J. I. Mullins.** 1999. Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. J Virol **73:**10489-502.

214. **Shekelle, P., M. Maglione, M. B. Geotz, G. Wagner, Z. Wang, L. Hilton, J. Carter, S. Chen, C. Tringle, W. Mojica, and S. Newberry.** 2007. Antiretroviral (ARV) drug resistance in the developing world. Evid Rep Technol Assess (Full Rep)**:**1-74.

215. **Sow, P. S., L. F. Otieno, E. Bissagnene, C. Kityo, R. Bennink, P. Clevenbergh, F. W. Wit, E. Waalberg, T. F. Rinke de Wit, and J. M. Lange.** 2007. Implementation of an antiretroviral access program for HIV-1-infected individuals in resource-limited settings: clinical results from 4 African countries. J Acquir Immune Defic Syndr **44:**262-7.

216. **Stanford University, U.** 2007. HIV Drug Resistance Database. Stanford University.

217. **Stevenson, M.** 2003. HIV-1 pathogenesis. Nat Med **9:**853-60.

218. **Streeck, H., J. S. Jolin, Y. Qi, B. Yassine-Diab, R. C. Johnson, D. S. Kwon, M. M. Addo, C. Brumme, J. P. Routy, S. Little, H. K. Jessen, A. D. Kelleher, F. M. Hecht, R. P. Sekaly, E. S. Rosenberg, B. D. Walker, M. Carrington, and M. Altfeld.** 2009. Human immunodeficiency virus type 1-specific CD8+ T-cell responses during primary infection are major determinants of the viral set point and loss of CD4+ T cells. J Virol **83:**7641-8.

219. **Tang, J., S. Tang, E. Lobashevsky, A. D. Myracle, U. Fideli, G. Aldrovandi, S. Allen, R. Musonda, and R. A. Kaslow.** 2002. Favorable and unfavorable HLA class I alleles and haplotypes in Zambians predominantly infected with clade C human immunodeficiency virus type 1. J Virol **76:**8276-84.

220. **Team, H. D. R. S.** 2003. Guidelines for Surveillance of HIV Drug Resistance. Guideline. WHO.

221. **Tenzer, S., E. Wee, A. Burgevin, G. Stewart-Jones, L. Friis, K. Lamberth, C. H. Chang, M. Harndahl, M. Weimershaus, J. Gerstoft, N. Akkad, P. Klenerman, L. Fugger, E. Y. Jones, A. J. McMichael, S. Buus, H. Schild, P. van Endert, and A. K. Iversen.** 2009. Antigen processing influences HIV-specific cytotoxic T lymphocyte immunodominance. Nat Immunol **10:**636-46.

222. **Trkola, A., H. Kuster, P. Rusert, B. Joos, M. Fischer, C. Leemann, A. Manrique, M. Huber, M. Rehr, A. Oxenius, R. Weber, G. Stiegler, B. Vcelar, H. Katinger, L. Aceto, and H. F. Gunthard.** 2005. Delay of HIV-1 rebound after cessation of antiretroviral therapy through passive transfer of human neutralizing antibodies. Nat Med **11:**615-22.

223. **UNAIDS/WHO** 2009, posting date. Report on the Global HIV/AIDS Epidemic. www.unaids.org. UNAIDS. [Online.]

224. **UNAIDS/WHO.** 2007. Report on the Global HIV/AIDS Epidemic. www.unaids.org. Report. UNAIDS.

225. **Uy, J., C. Armon, K. Buchacz, J. Brooks, and H. Investigators.** Jul 2007. Initiation of HAART at CD4 cell counts >= 350 cells/mm3 is associated with a lower prevalence of antiretroviral resistance mutations at virologic failure, 4th IAS Conference on HIV Pathogenesis, Treatment and Prevention. Kaiser Network, Sydney, Australia.

226. **Vallerskog, T., I. Gunnarsson, M. Widhe, A. Risselada, L. Klareskog, R. van Vollenhoven, V. Malmstrom, and C. Trollmo.** 2007. Treatment with rituximab affects both the cellular and the humoral arm of the immune system in patients with SLE. Clin Immunol **122:**62-74.

227. **van der Ryst, E., M. Kotze, G. Joubert, M. Steyn, H. Pieters, M. van der Westhuizen, M. van Staden, and C. Venter.** 1998. Correlation among total

lymphocyte count, absolute CD4+ count, and CD4+ percentage in a group of HIV-1-infected South African patients. J Acquir Immune Defic Syndr Hum Retrovirol **19:**238-44.

228. **VaxGen.** 2003. VaxGen announces results of its phase 3 HIV vaccine trial in Thailand: vaccine fails to meet endpoints (http://investor.vaxgen.com/profiles/investor/ResLibraryView.asp?ResLibraryID=5836&GoTopage=16&BzID=923&Category=214). VaxGen, Inc. (Nasdeq: VXGN).

229. **Velazquez-Campoy, A., M. J. Todd, S. Vega, and E. Freire.** 2001. Catalytic efficiency and vitality of HIV-1 proteases from African viral subtypes. Proc Natl Acad Sci U S A **98:**6062-7.

230. **Venter, W. D. F.** 2005. A critical evaluation of the South African state antiretroviral programme. The Southern African Journal of HIV Medicine.

231. **Walker, L. M., S. K. Phogat, P. Y. Chan-Hui, D. Wagner, P. Phung, J. L. Goss, T. Wrin, M. D. Simek, S. Fling, J. L. Mitcham, J. K. Lehrman, F. H. Priddy, O. A. Olsen, S. M. Frey, P. W. Hammond, S. Kaminsky, T. Zamb, M. Moyle, W. C. Koff, P. Poignard, and D. R. Burton.** 2009. Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. Science **326:**285-9.

232. **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**.

233. **Walker, S., DART Trial Team.** 2009. Presented at the 5th IAS Conference on HIV Pathogenesis, Treatment and Prevention, 19-22 July 2009, , Cape Town, South Africa.

234. **Wang, Y. E., B. Li, J. M. Carlson, H. Streeck, A. D. Gladden, R. Goodman, A. Schneidewind, K. A. Power, I. Toth, N. Frahm, G. Alter, C. Brander, M. Carrington, B. D. Walker, M. Altfeld, D. Heckerman, and T. M. Allen.** 2009. Protective HLA class I alleles that restrict acute-phase CD8+ T-cell responses are associated with viral escape mutations located in highly conserved regions of human immunodeficiency virus type 1. J Virol **83:**1845-55.

235. **Wei, X., J. M. Decker, S. Wang, H. Hui, J. C. Kappes, X. Wu, J. F. Salazar-Gonzalez, M. G. Salazar, J. M. Kilby, M. S. Saag, N. L. Komarova, M. A. Nowak, B. H. Hahn, P. D. Kwong, and G. M. Shaw.** 2003. Antibody neutralization and escape by HIV-1. Nature **422:**307-12.

236. **Weiss, R. A.** 2008. Special anniversary review: twenty-five years of human immunodeficiency virus research: successes and challenges. Clin Exp Immunol **152:**201-10.

237. **Worobey, M., M. Gemmel, D. E. Teuwen, T. Haselkorn, K. Kunstman, M. Bunce, J. J. Muyembe, J. M. Kabongo, R. M. Kalengayi, E. Van Marck, M. T. Gilbert, and S. M. Wolinsky.** 2008. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. Nature **455:**661-4.

238. **Worobey, M., M. L. Santiago, B. F. Keele, J. B. Ndjango, J. B. Joy, B. L. Labama, A. B. Dhed, A. Rambaut, P. M. Sharp, G. M. Shaw, and B. H. Hahn.** 2004. Origin of AIDS: contaminated polio vaccine theory refuted. Nature **428:**820.

239. **Wright, J., M. brockman, Z. Brumme, K. Nair, F. Chonco, M. van der Stok, Z. Mncube, K. Bishop, B. Walker, and T. Ndung'u.** 2010. Gag-Protease-mediated Fitness Is Associated with Viral Set Point in Individuals Acutely Infected with HIV-1 Subtype C, 17th Conference on Retroviruses and Opportunistic Infections, San Francisco.

240. **Xu, X. N., G. R. Screaton, F. M. Gotch, T. Dong, R. Tan, N. Almond, B. Walker, R. Stebbings, K. Kent, S. Nagata, J. E. Stott, and A. J. McMichael.** 1997. Evasion of cytotoxic T lymphocyte (CTL) responses by nef-dependent induction of Fas ligand (CD95L) expression on simian immunodeficiency virus-infected cells. J Exp Med **186:**7-16.

241. **Yang, Z., and R. Nielsen.** 1998. Synonymous and nonsynonymous rate variation in nuclear genes of mammals. J Mol Evol **46:**409-18.

242. **Yarchoan, R., R. W. Klecker, K. J. Weinhold, P. D. Markham, H. K. Lyerly, D. T. Durack, E. Gelmann, S. N. Lehrman, R. M. Blum, D. W. Barry, and et al.** 1986. Administration of 3'-azido-3'-deoxythymidine, an inhibitor of HTLV-III/LAV replication, to patients with AIDS or AIDS-related complex. Lancet **1:**575-80.

243. **Zhou, T., I. Georgiev, X. Wu, Z. Y. Yang, K. Dai, A. Finzi, Y. D. Kwon, J. F. Scheid, W. Shi, L. Xu, Y. Yang, J. Zhu, M. C. Nussenzweig, J. Sodroski, L. Shapiro, G. J. Nabel, J. R. Mascola, and P. D. Kwong.** 2010. Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. Science **329:**811-7.

244. **Zhu, T., B. T. Korber, A. J. Nahmias, E. Hooper, P. M. Sharp, and D. D. Ho.** 1998. An African HIV-1 sequence from 1959 and implications for the origin of the epidemic. Nature **391:**594-7.

245. **Zhuang, Y., Y. Sun, S. Zhai, D. Huang, S. Zhao, S. Wang, W. Kang, X. Li, B. D. Walker, M. Altfeld, and X. G. Yu.** 2008. Relative dominance of Env-gp41-specific cytotoxic T lymphocytes responses in HIV-1 advanced infection. Curr HIV Res **6:**239-45.

246. **Zinkernagel, R. M., and P. C. Doherty.** 1997. The discovery of MHC restriction. Immunol Today **18:**14-7.