Population-specific HLA Impact in Immune Control of HIV in Mexico and Non-Mexican HIV Infected Cohorts

Claudia Ivette Juarez Molina

Wolfson College and

Department of Paediatrics

University of Oxford

Michaelmas Term 2013

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Population-specific HLA Impact in Immune Control of HIV in Mexico and Non-Mexican HIV Infected Cohorts

Claudia Ivette Juarez Molina, Wolfson College

Submitted for the degree of DPhil at the University of Oxford, Michaelmas 2013

Abstract

HIV-1 persists to be a major health problem worldwide. A prophylactic or therapeutic vaccine offers the best hope to restrain the HIV-1 epidemic, however a consistent correlate of immune protection is yet to be found.

HLA class I expression and their restricting HIV-1 specific $CD8^+$ T cell responses have been shown to play a vital role in the control of HIV-1 infection. The interactions between HIV-1 and $CD8^+$ T cell responses are complex and the mechanisms involved in the success or failure to control viraemia remain uncertain. Thus, the aim of these studies was to help define what $CD8^+$ T cell responses a vaccine needs to induce to achieve durable immune control of HIV-1 infection. Focusing initially on HLA-B*35, an allele that has consistently been associated with rapid HIV-1 disease progression in the context of B clade infection, this study shows substantial differences in markers of HIV-1 disease outcome associated with different HLA-B*35 subtypes. Preliminary data suggest that effective targeting of a single epitope in Gag may be associated with HLA-B*35 mediated control of HIV-1 disease progression. Increased breadth of the Gag-specific $CD8^+$ T cell responses is found to be associated with decreasing viral loads. These data therefore support the Gag hypothesis, and suggest that targeting of certain regions of the HIV-1 genome may have a positive effect in disease outcome, even for individuals carrying "detrimental" alleles.

The extensive diversity of the HIV-1 genome and rapid viral adaptation are the main challenges to vaccine design. Previous studies have suggested that effective $CD8^+$ T cell responses drive selection of escape mutations that reduce viral replication capacity (VRC). There is also evidence that certain escape mutations can be transmitted from one host to another allowing for its accumulation in a population. The second study looked at the impact of HLA

driven evolution of HIV-1 in VRC at a population level. This study compared two ART-naïve HIV-1 B clade infected cohorts, in Mexico and Barbados, in which protective HLA alleles (HLA-B*27/57/58:01/81:01) are expressed at 10% and 35% respectively, to analyze differences in VRC at a population level. Viral loads (VL) were found to be significantly higher in Mexico compared to Barbados and median $CD4^+$ T cell counts significantly lower. Analysis of VRC in a subset of subjects in each cohort matched by $CD4^+$ T cell counts between 300-500 cells/μl revealed that VL and VRC was significantly higher in the Mexican subset. This VRC difference was associated with accumulations in Barbados of eight previously described Gag escape mutation where fitness cost has previously been implicated. Accumulation remained significant in mismatched subjects. These data suggest that VLs and disease progression rates may differ between distinct populations as a result of the frequency of protective alleles in the respective populations, and that $CD4^+$ T cell count-based guidelines to initiate antiretroviral therapy (ART) may need to be modified accordingly, to optimize the effectiveness of treatment-for-prevention strategies and reduce HIV-1 transmission rates to the absolute minimum.

The final project aimed to improve the HIV-1 replication fitness assays currently used in the context of C clade infection. In order to achieve this, we attempted to design a clade C infectious molecular clone for the testing of *gal-pol* gene regions. However, the clones produced were not replication competent. Sequence analysis showed a large quantity of stop codons, most located within *env* which may explain the lack of infectivity. Chapter 5 describes the methodology used in the construction of the clade C isolate and suggests future work. Although we were unsuccessful in producing a replication competent virus, the construction of a C clade backbone which replicates efficiently remain an aim due to its importance for research directed to the analysis of genetic determinants of C clade virus.

Data presented in this thesis suggest that vaccine-induced immune responses should aim to focus on vulnerable regions of the virus. These are conserved regions that can not escape without a high fitness cost and with a complex and difficult selection of compensatory mutations.

Although much work remains to be done to achieve an effective $CD8^+$ T cell based vaccine, hope remains that the induction of HIV control may be possible.

0.1 Acknowledgements

Firstly I would like to thank my supervisor, Professor Philip Goulder, for giving me the great opportunity of working in the group. This thesis would not have been possible without his supervision, advice and input, not forgetting the provision of generous funding.

Equally, I would like to thank everyone in the Goulder group that have directly contributed to this thesis and/or have provided me with support, help, advice and company for the past 4 years, including Dr Philippa Mathews, Dr Rebecca Payne, Dr Henrik Kloverpis, Dr Masahiko Mori, Hanna Podd, Emily Adland, Ellen Leitman, Jacqui Brener, Catherine Koophethile, Luisa Mori, Anna Csala, Nora Lavandier and Amna Malik.

A big thank you to Dr Julia Garcia Prado, your support and advice were invaluable. Thank you for inviting me to spend time with you and the Picardo group in IrsiCaixa, and sharing expertise and valuable samples.

Likewise a big thank you to everyone in the Reyes-Teran group for your great hospitality, friendship, generosity and sharing of valuable samples.

Additionally I would like to thank my colleges and friends in the Peter Medawar building for all their support and advice. Particularly Helen Brown and Nicola Robinson for sharing their wisdom with me; Nicole Yager, Isla Humphrey, Jenny Smith, James Williams, Peter Flanagan, Matthew Jones, Lyle Murray and Matthew Pace for making my journey much more enjoyable.

Thank you mum (Martha Molina) and dad (Xavier Gomez) for always being loving, supportive and encouraging, and more importantly for believing in me.

Above all, I would like to thank my fiancé Darren, you have been a constant source of strength. Your patience and love have been crucial in keeping me going.

Finally, a big thank you to the clinics, doctors, nurses and patients that allowed for this work to happen.

0.2 Statement of Authorship

Many individuals contributed to the work presented in this thesis. Unless otherwise specified, the data presented here were generated by myself, utilizing the methodology described.

Collaborators:

- Philip Goulder group, The Peter Medawar Building for Pathogen Research, South Parks Road, Oxford, UK

Professor Philip Goulder, Dr Philippa Matthews, Dr Rebecca Payne, Dr Henrik Kloverpris, Dr Masahiko Mori, Emily Adland, Ellen Leitman, Nora Lavandier, Jacqui Brener and Anna Csala.

- Gustavo Reyes-Teran group, Center for Research in Infectious Diseases, National Institute of Respiratory Diseases, Mexico City, Mexico

Gustavo Reyes-Teran, Dr Santiago Avila-Rios, Humberto Valenzuela-Ponce, Maribel Soto-Nava, Claudia Garcia-Morales, Daniela Garrido-Rodriguez.

- Clive Landis group, Ladymeade Reference Unit, Queen Elizabeth Hospital, Bridgetown, Barbados

Dr. Clive Landis, and Songee Branch

- Martinez-Picado Group, IrsiCaixa Foundation, Hospital Germans Trias I pujol, Badalona, Spain

Javier Martinez-Picado, Julia Garcia Prado, Judith Dalmau, Itziar Erkizia, Mari Carmen Puertas, Maria Teresa Rodriguez and Nuria Izquierdo

- Andrew Sewell Group, Cardiff University School of Medicine, Heath Park, Cardiff, UK

Professor Andrew Sewell, Garry Dolton and Anna Fuller

HLA typing was carried out at the ASHI* accredited HLA typing laboratory, University of Oklahoma Health Sciences Centre, USA, or by the Reyes- Teran group in Mexico.

For chapter 2, Humberto Valenzuela-Ponce, Maribel Soto-Nava, Claudia Garcia-Morales and Daniela Garrido-Rodriguez helped with the processing of samples and the process required for HLA typing of samples.

Dr Beatriz Mothe Pujadas and Judith Dalmau kindly provided us with HLA-B*35:08 samples from Spain.

In chapter 2, Dr Jonathan Carlson from the Microsoft Institute in Seattle, USA, conducted the mathematical analysis to identify HLA class I associated polymorphisms in the Mexican cohort.

In chapter 2, crystal structure and binding assays of B*35:08-NY10 E/D complexes were produced by Andrew Sewell and his group in Cardiff.

In chapter 3, the Gag-pro chimeric viruses for fitness assays from Mexico were produced by Maribel Zoto-Nava, Reyes-Teran group. Gag-pro chimeric virus and fitness assays from Barbados samples were carried out by Rebecca Payne, Goulder group.

Dr Julia Garcia Prado provided me with the technical training and support regarding the construction of the C clade backbone system described in chapter 4.

0.3 Glossary

℃	Degrees Celsius
AA	Amino acid
Ab	Antibody
ADCC	Antibody-Dependent cell-mediated cytotoxicity
AIDS	Acquired immunodeficiency syndrome
Amp	Ampicillin
APC	Antigen presenting cell
ART	Anti-retroviral therapy
BCL	B cell line
CO_2	Carbon dioxide
DC	Dendritic cell
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphates
DMSO	Dimethyl sulfoxide
E. coli	*Escherichia coli*
ELISpot	Enzyme-linked Immunospot assay
Env	Envelope Glycoproteins (gp120 and gp41)
ER	Endoplasmic Reticulum
FACS	Fluorescent activated cell sorting
FCS	Foetal calf serum
g	Gram
Gag	Capsid proteins
GWAS	Genome wide association study
HAART	Highly-active anti-retroviral therapy
HIV	Human Immunodeficiency virus
HLA	Human leukocyte antigen
IFN	Interferon
INT	Integrase

Kan	Kanamycin
LB-Broth	Luria-Bertani medium
log	Logarithm
LTNP	Long-term non-progressor
LTR	Long terminal repeat
µg	Microgram
mg	Milligram
µl	Microlitre
µM	Micromolar
MAb	Monoclonal antibody
MHC	Mayor histocompatibility complex
ML	Maximum likelihood
MOI	Multiplicity of infection
Mut	Mutants
N	Number
NAb	Neutralizing antibodies
NaCl	Sodium Chloride
Nef	Negative factor protein
ng	Nanogram
nM	Nanomolar
OLP	Over lapping peptide
PAMP	Pathogen associated molecular pattern
PBMC	Peripheral blood mononuclear cells
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
pH	Potential hydrogen
PHA	Phytohaemagglutinin
PI	Protease Inhibitor
Pol	Viral enzymes (protease, reverse transcriptase and integrase)
PR	Protease

PRR	Pattern recognition receptor
RE	Restriction enzyme
Rev	The second regulatory factor
RNA	Ribonucleic acid
RP	Rapid progressor
RT	Reverse transcriptase
SAP	Shrimp alkaline phosphatase
SFC	Spot forming cell
SIV	Simian immunodeficiency virus
SNP	Single nucleotide polymorphism
TAP	Transporter associated with antigen processing
Tat	Transactivator of HIV gene expression
TCR	T cell receptor
TLR	Toll-like receptor
Vif	Viral Infectivity Factor
VL	Viral load
vpr	Viral protein R
vpu	Viral protein U
VRC	Viral replication capacity
WT	Wild type

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 HIV/AIDS

1.1.1 The HIV pandemic

Although over 30 years have passed since the first description of the acquired immune deficiency (AIDS) in 1981 and the isolation of the agent provoking the disease in 1983 (Luc Montanier, Institute Pasteur), full understanding of the interactions between the host immune system and the Human Immunodeficiency Virus (HIV) still remains unclear. Most importantly, an effective vaccine is yet to be found.

The introduction of antiretroviral therapy in the mid 1990 has reduced the rate of mortality and morbidity caused by HIV-1 infection. However despite the strides on therapeutic regimens and despite the continued effort from the scientific community, HIV-1 persists to be a major health problem worldwide. It has been estimated that approximately 75 million (63-89 million) people have become infected with HIV-1 since the start of the epidemic (UNAIDS report on the global AIDS epidemic 2013) and about 35 million have died of AIDS (UNAIDS WHO epidemic update 2011).

According to the Joint United Nations Program on HIV/AIDS, the implementation of prevention programs and low cost testing for early detection have led to a decline in the number of new infection and the number of AIDS related deaths, however an estimated 35.3 million (32.2–38.8 million) of people were living with HIV-1 at the end of 2012 and an estimated 2.3 million (1.9-2.7 million) of new infections occurred worldwide in the same year.

Although the epidemic is widespread globally, sub-Saharan Africa remains the epicenter with 70% of all new infections (Figure 1.1.1). This is a region with considerable limitations in medical and social resources, thus, 50% of individuals eligible for treatment do not have access to it. Consequently the development of an effective, safe and affordable HIV-1 vaccine is imperative.

As yet a consistent correlate of immune protection has not been found, however long-term control of the disease by certain individuals gives hope for the development of an effective vaccine which would elicit the effective immune responses that lead to control. Hence allowing for reduction of the incidence of HIV-1 and its transmission.

To be able to design such a vaccine, further understanding of the HIV/immune system interaction is needed.

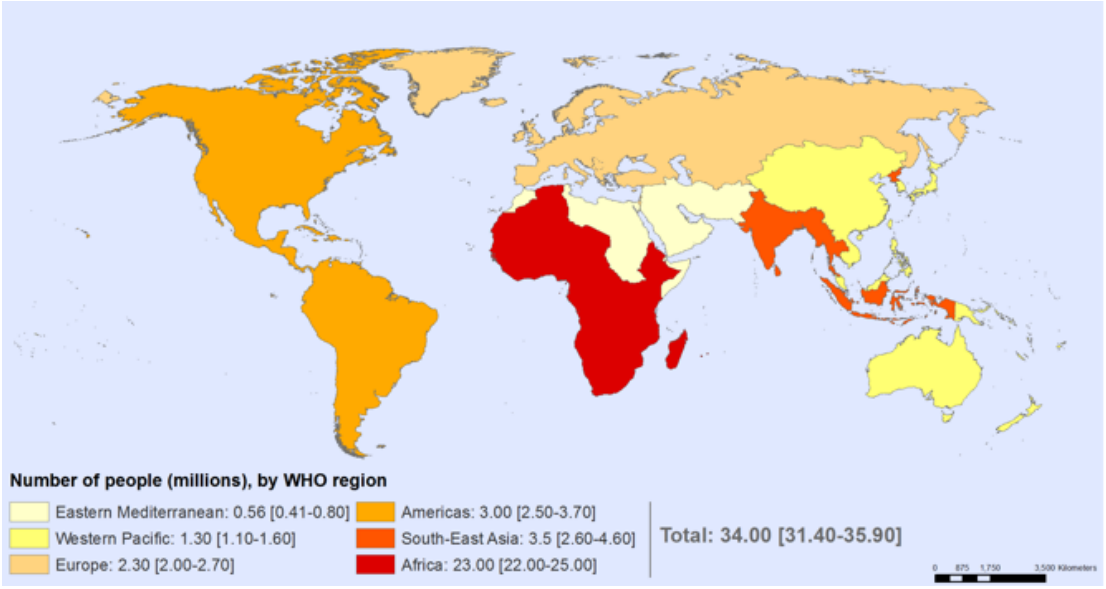


Figure 1.1.1: **Adults and children estimated to be living with HIV in 2011, WHO report.** It was estimated that 0.8% of adults were living with HIV in 2011. Sub-Saharan Africa continues to be the region most affected with approximately 1 in every 20 adult living with HIV in the same year.

1.1.2 Virology

Classification

HIV is a single-stranded enveloped RNA virus member of the Retroviridae family, genus Lentivirus.

Two types of HIV have been characterized, HIV-1 and HIV-2. HIV 1 and 2 have dis-

tinct zoonotic origins. HIV-1 is closely related to the SIV strain isolated from chimpanzees (*Pan troglodytes troglodytes*) in Central Africa, while HIV-2 is closer to the strain from sooty mangabeys (Cercocebus atys) of West Africa (Figure 1.1.2) [126] . HIV-1 is the most prevalent type worldwide and is generally more virulent than the HIV-2.

Phylogenetically, HIV-1 is divided into three major subgroups, M (major), O (outlier) and N (non M/non O). These groups are similar but can exhibit different phenotypes that may have an impact on fitness of the virus, drug resistance and vaccine design.

The O group is mainly confined to Africa, while the N group has only been found in Cameroon. The M group is substantially more widespread in the global epidemic and is formed by 11 genetic subtypes A1, A2, B, C, D, F1, F2, G, H, J and K (http://www.hiv.lanl.gov). Variation between subtypes or clades in amino acid sequence of the envelope protein may exceed 30%. Recombinant forms can also be found and they are formed of different genomic regions from distinct subtypes.

The present studies focus on B and C Clade HIV-1 that account for the majority of North American/European and Southern African infection respectively.

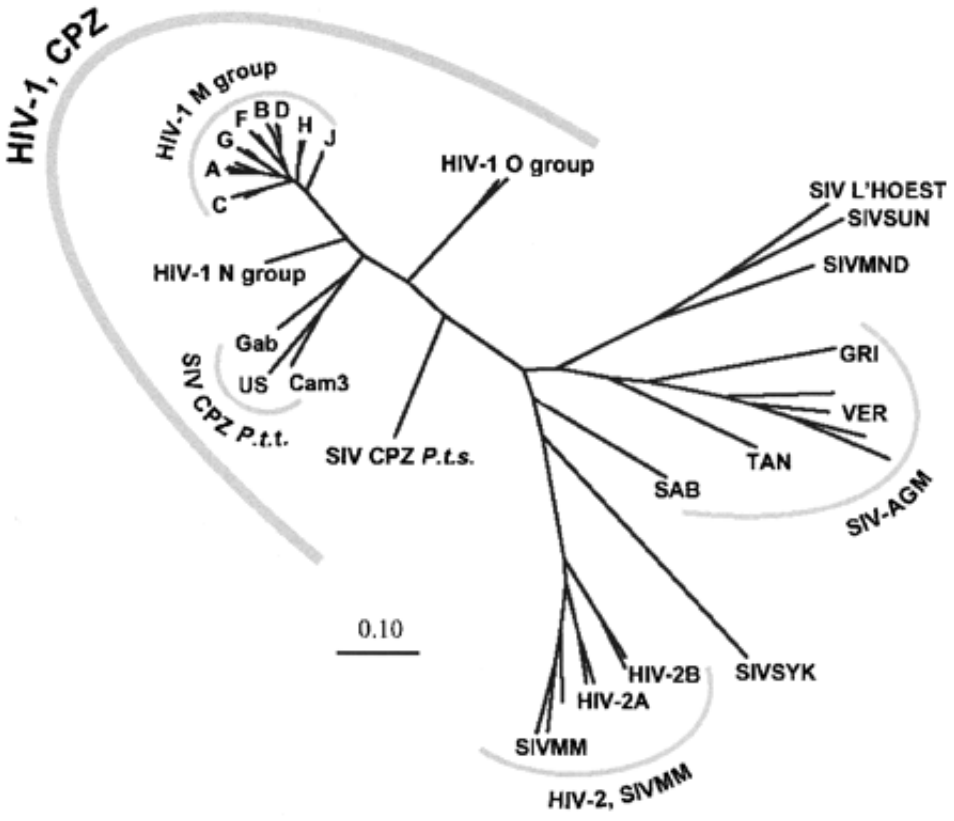


Figure 1.1.2: **HIV diversity.** Alignment of pol genes of HIV-1, HIV-2 and SIV strain SIVmm; SIV of sooty mangabeys or macaques experimentally infected with SIVmm. Reproduced from Human Retroviruses and AIDS (Los Alamos, New Mexico; Theoretical Biology and Biophysics Group, Los Alamos National Laboratory)

Virion structure

The HIV-1 genome, about 9.5 kb, is composed of nine genes; three encoding structural proteins (Gag, Pol and Env) and six encoding regulatory and accessory proteins (Tat, Rev, Vif, Vpr, Vpu and Nef) (Fig 1.1.3). The HIV genome is efficiently organized and uses all off the three reading frames that allows for overlapping of the gene coding regions [78].

The viral RNA is contained inside a cone-shaped capsid composed of approximately 2000 copies of the p24 protein. This is in turn surrounded by the viral envelope, which is formed of 71 viral envelope (Env) proteins embedded in a lipid bilayer from the membrane of a previously infected cell. A brief description of the function of each protein follows:

- *Envelope (env)* is the most variable region of the HIV-1 genome. It is formed of two non-

covalently linked glycoproteins (gp) 120 and the transmembrane gp41. gp120 is composed of a constant region and a highly variable region where changes alter co-receptor usage.

- *Gag* encodes the capsid proteins p17 (matrix), p24 (capsid), p7 (nucleocapsid) and p6. Its function involves association with the plasma membrane that allows for viral assembly.
- *Pol* encodes the viral protease, reverse transcriptase and integrase, required for the viral genome integration into the host.
- *Transactivator of HIV gene expression (tat)*, encodes two forms of the protein, Tat-1 and Tat-2, which bind to the Tar RNA element and activates initiation of transcription.
- *Rev* encodes a 19kD phosphoprotein, which function promotes nuclear export, stabilization and usage of the viral mRNAs.
- *The viral infectivity factor (vif)* encodes a cytoplasmic protein that is essential for subsequent infectivity of new viral particles and prevents the function of the APOBEC-3G protein, which determines DNA:RNA heteroduplexes.
- *Viral protein R (vpr)* is thought to target the nuclear import of pre-integration complexes, arrest of host cells at the G2/M phase of the cell cycle and transactivation of cellular genes.
- *Viral protein U (vpu)* acts in the degradation of CD4 in the endoplasmic reticulum (ER) and enhances virion release from the plasma membrane of infected cells.
- *Nef* encodes a myristoylated protein and promotes viral spread by downregulation of CD4 and MHC class I molecules.

Viral proteins act in various ways to facilitate viral replication and change normal cellular function, however these processes remain poorly understood.

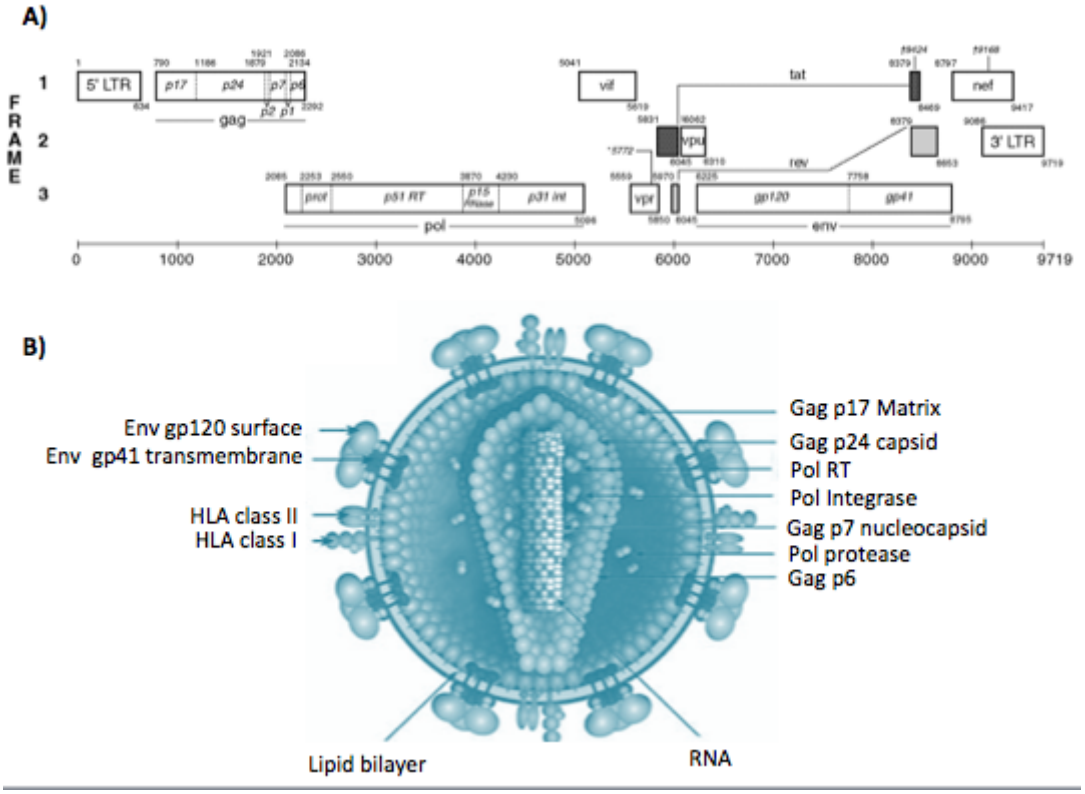


Figure 1.1.3: **Schematic representation of the HIV genome and virion.** Adapted from Los Alamos database, www.hiv.lanl.gov.

Life cycle

HIV binds to the cell membrane of host cells through the interaction of the gp120 and both the CD4 and a co-receptor, CC chemokine receptor type 5 (CCR5) or the CXC chemokine receptor type 4 (CXCR4). All HIV strains can recognize and bind CD4, however they can exhibit different affinities towards the chemokine receptors CCR5 or CXCR4, resulting in distinct tropisms. The affinity to either co-receptor is determined by the variable region of gp120. CXCR4 tropic viruses are generally characterized by high viral titters in comparison to CCR5 tropic viruses, which are thought to have a less aggressive growth. These co-receptors are differentially expressed on $CD4^+$ T cells. CCR5 is expressed by effector and effector memory T cells which are located in effector sites such as the *lamina propia* of the intestinal mucosa. HIV-1 preferentially utilizes CCR5 for transmission early in infection [121]. As disease progresses there is a switch on co-

receptor usage to the dual tropic or CXCR4 tropic virus [121]. Most peripheral $CD4^+$ T cells express CXCR4. It is thought that the preference of the virus for CCR5 allows for rapid spread of infection as the target cells have a high level of proliferation. Studies have shown preferential and marked depletion of $CD4^+$ T cells within the gastrointestinal tract during acute infection [91].

Interaction of both CD4 and the co-receptor with gp120 leads to a structural change on the N-terminus of the gp120, which allows for the fusion of the virus with the membrane. Allowing for the entry of the viral contents into the cytoplasm, including the RNA viral genome, tRNA primers, viral protease, RT and integrase. The viral RNA is then reverse transcribed into cDNA that is later transported to the nucleus by the interaction of the viral proteins including matrix, integrase and Vpr, and host proteins, such as the high-mobility group protein B1.

Once the viral DNA is in the nucleus, integrase leads to its integration into the host-cell genome through DNA splicing, or else the viral DNA circularizes. The integrated viral DNA is known as proviral DNA, and is replicated as a part of the normal host genome. Proviral DNA may persist for long periods and through several rounds of mitotic cell division.

The 5' LTR functions as a promoter and regulates the production of RNA transcripts. The 3' LTR acts as a polyadenylation and termination site. The transcribed HIV RNA molecules can be spliced for translation of viral proteins or be exported in an unspliced form for the production of a new virion [78].

Protease processes the viral polypeptides into their functional forms and they are assembled with the full-length viral RNA into nascent viral particles. Gag interaction at this stage with a host tumor suppressor gene is critical for the release of the new virion. Vpu also interacts with theterin, an endogenous membrane associated protein that can block release of viral particles, and facilitates viral release [78] (Figure 1.1.4).

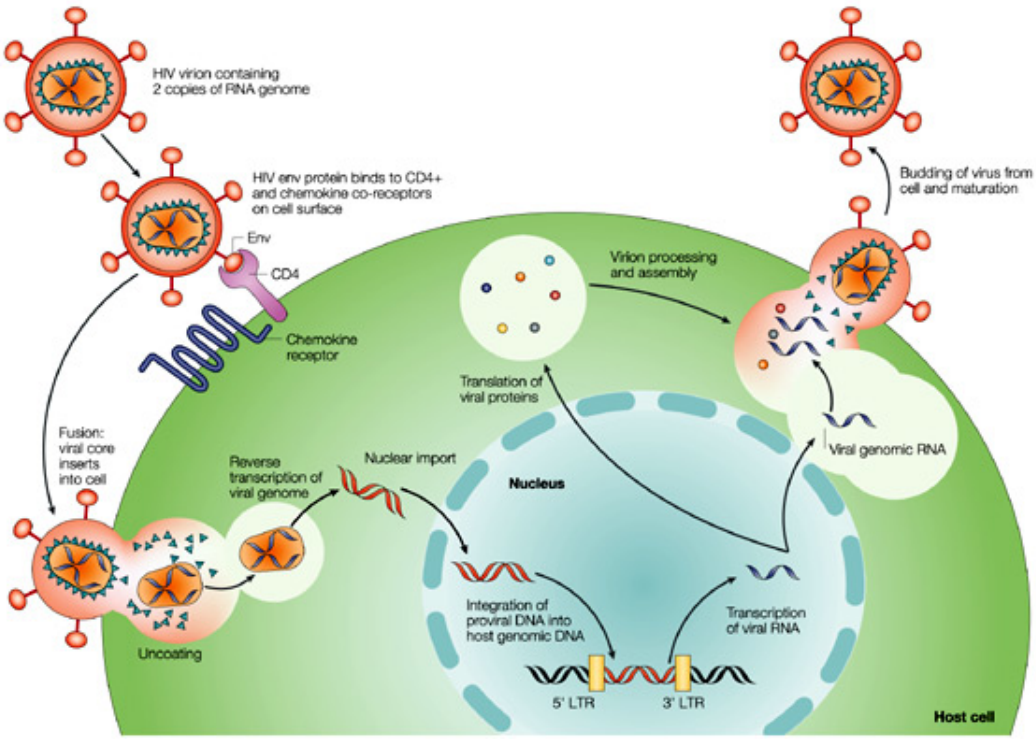


Figure 1.1.4: **Schematic representation of the HIV life cycle.** Reproduced from Andrew Rambaut et.al, Nature reviews Genetics, 2004.

1.2 HIV infection

HIV-1 is transmitted through the exchange of bodily fluids containing infectious viral particles. The primary route of HIV-1 transmissions worldwide occurs by sexual exposure through the genital tract or rectal mucosa. Non sexual transmission can occur through transfusion of contaminated blood, parenteral transmission through needle sharing amongst intravenous drug users, occupational exposure and mother-to-child transmission.

The mechanisms involved in early infection are not well understood. However single genome amplification of the first detectable virus have shown that in the majority of heterosexual cases (approximately 80%) a founder virus is responsible for the infection [65]. In the case of MSM this infection by a single founder virus is reduced to 60% [79].

1.2.1 Clinical course of disease

HIV-1 selectively infects $CD4^+$ T cells and to a lesser extent cells of the mononuclear phagocyte lineage (principally blood monocytes and tissue macrophages), B lymphocytes, natural killer (NK) cells, dendritic cells (DC) (Langerhans cells and follicular dendritic cells in lymph nodes), hematopoietic stem cells, endothelial cells, microglial cells in the brain and gastrointestinal epithelial cells.

HIV-1 infection is chronic and progressive. The clinical course of HIV-1 infection is commonly described in three stages; acute, chronic and AIDS [78]. Acute HIV-1 infection is the period of infection prior to the development of a detectable antibody response or seroconvertion. During this period plasma viral load (VL) reaches a peak with the range of 10^7-10^8 HIV-1 RNA copies/ml (Figure 1.2.1) [90]. Viral replication occurs in $CD4^+$ T cell subsets across the body, although only a modest drop in $CD4^+$ T cells is observed in the blood. Most of the destruction appears to be centered within the gastrointestinal associated lymphoid tissues (GALT), resulting in 80% destruction of the GALT during acute infection [17]. HIV-specific $CD8^+$ T cells undergo a marked clonal expansion which coincides with the reduction of viremia to a viral set point between 10^2-10^3 RNA copies/ml. The magnitude of the viral set point achieved is highly predictive of the disease progression [50]. The acute phase lasts approximately six months and patients experience flu-like symptoms.

The chronic stage in adults is slow, with a ten year median, and it is usually asymptomatic. HIV-1 infection triggers a potent host inflammatory and adaptive immune response. The virus utilizes multiple mechanisms to evade these responses, maintaining a level of viral replication that establishes and sustains the state of chronic persistent immune activation that continues throughout the course of infection in the absence of ART. High levels of viral replication and $CD4^+$ T cell turnover eventually lead to progressive immune dysfunction and the loss of cell-mediated immunity, which in turn causes progression to AIDS [78].

AIDS is the end stage of the disease when severe and progressive immunodeficiency renders the host susceptible to opportunistic infections and malignancies that eventually lead to death. The $CD4^+$ T cell count below 200 cells per mm^3 of blood marks the onset of AIDS.

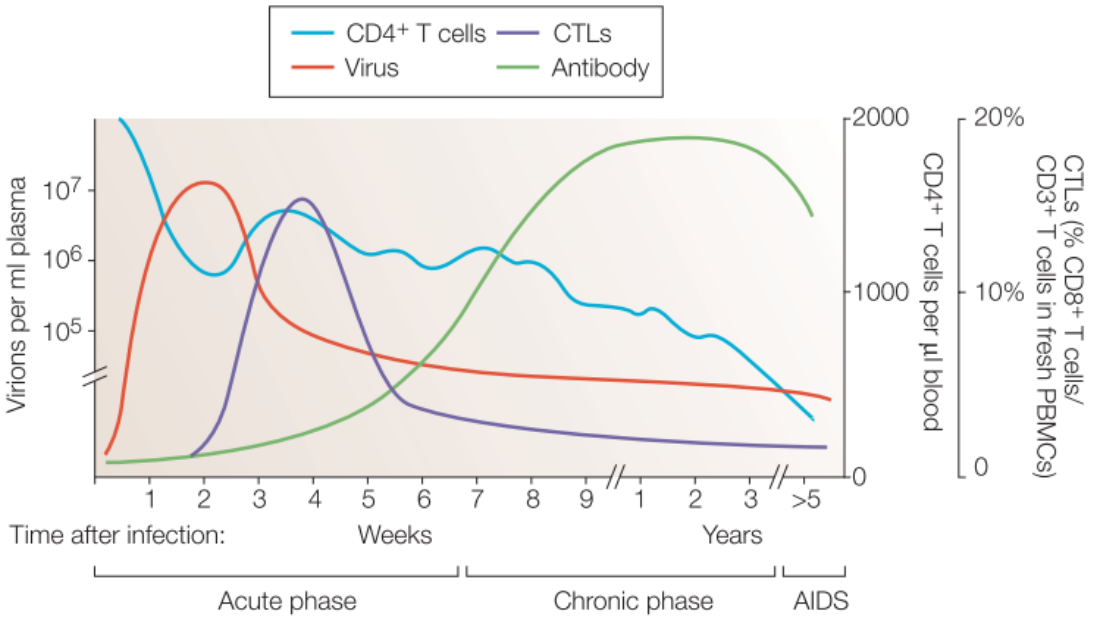


Figure 1.2.1: **Schematic representation of the typical course of HIV-1 infection**, image from Goulder and Watkins, Nature reviews Immunology, 2004.

1.2.2 HIV-1 reservoir

A major obstacle for the eradication of HIV-1 infection is the pool of latently infected, long-lived, resting $CD4^+$ memory T cells that form a cellular reservoirs for HIV-1 persistence in patients even when they are on ART [26]. Most effector T cells die during the immune response, but some re-enter a resting state and become memory T cells. The latent reservoir of HIV-1 infected $CD4^+$ T cells is established in primary infection and is composed of approximately 10^6-10^7 cells each with three to four copies of integrated provirus [26]. These cells are stable, have a long half-live and act as a constant source of the residual viremia that allows for the virus to persist after treatment. Identification of these infected resting cells and the elements that restrict viral replication and lead to latency are of great importance for drug development.

1.3 Immunity

1.3.1 Cellular immunity

$CD8^+$ T cells are effector cells capable of killing HIV-1 infected cells through T cell receptor (TCR) interaction and recognition of viral peptides associated with the human leukocyte antigen (HLA) class I molecules on the cell surface of infected cells. $CD8^+$ T cells are part of the adaptive arm of the immune system and work in the control of intracellular pathogens. HIV-1 specific $CD8^+$ T cells have been detected in early HIV-1 infection and their emergence correlates with temporal control of viremia as the viral set point is reached [15]. Moreover, high frequency of HIV-1 specific $CD8^+$ T cells during acute HIV-1 infection have been associated with lower viremia and slower decrease in $CD4^+$ T cell counts [100]. On the other hand, a number of HIV-1 infected individuals with advanced disease, have detectable HIV-1 specific $CD8^+$ T cell responses. Suggesting a 'quality' feature of responses capable of effective immune control. Indeed difference in polyfunctionality, proliferative capability and avidity of $CD8^+$ T cell responses have been suggested between long-term non-progressors (LTNPs) and normal progressors [119, 136].

HLA class I molecules also act as ligands for NK cell killer cell immuno-globulin-like receptors (KIRs). Certain KIRs, KIR3DL1 and KIR3DS1, have been shown to be associated with better control of HIV-1 in patients that also have HLA-B alleles with Bw4 epitope specificity (Bw4-80I) [136]. HLA-B alleles associated with HIV-1 control, namely HLA-B*57, B*27 and B*58, are members of the Bw4-80I group. Thus suggesting that the mechanisms involved in viral control may include the induction of an efficient NK cell response through their interaction with NK cell receptors [119].

NK cells are an important component of the innate system. They act in early responses against infected cells and do not require activation through cytokines. NK cells also produce cytokines and chemokines that attract and activate cells of the adaptive immune system.

HIV-1 downregulates HLA class I expression on the surface of infected cells to escape $CD8^+$ T cell responses, this leaves infected cells vulnerable to NK cell lysis. However Nef selectively downregulates HLA-A and HLA-B alleles but not HLA-C alleles, which are the dominant ligand for inhibitory KIR2D receptors [72]. Thus the virus is able to overcome both $CD8^+$ T cell and NK cell responses [83].

$CD4^+$ T cells play a critical role in host defense against infection, however they are also the main targets for HIV-1. $CD4^+$ T cells respond to infection through robust proliferation and differentiation, which gives rise to the effector cells required for the development and maintenance of specific $CD8^+$ T cell responses as well as helping to maintain many other branches of the immune response. $CD4^+$ T cells provide antigen-specific B cells with the necessary signals to produce and maintain the germinal center reaction, thus aiding the generation of long-lived antibody-secreting plasma cells. Therefore a key limitation of the HIV-1 specific immune responses is that all responses require 'help' by $CD4^+$ T cells, however $CD4^+$ T cells activation favors viral replication as it provides the virus with additional target cells, therefore leading to an impaired cellular immune response [108].

1.3.2 Humoral immune response

The most effective antiviral vaccines produced to date elicit neutralizing antibodies (Abs). HIV-1 entry is mediated by Env, therefore it has been suggested that antibodies against the conserved region of Env can block viral entry and therefore abrogate infection [128]. For this reason extensive research on antibodies against HIV-1 that may prevent infection has been carried out.

It has been demonstrated that the virus escapes Abs quite rapidly. Therefore, the high variability of the HIV-1 genome requires broadly neutralizing antibodies, which are hard to come by [85]. On the other hand the correlate analysis on samples from the RV144 trial in Thailand, which resulted in a 31% reduction of HIV-1 acquisition, showed no correlation between protection and $CD8^+$ T cell responses or with the level of broadly neutralizing Ab. The study showed that higher plasma levels of antibodies to the V1V2 loop were associated with a reduced risk of transmission. Additionally high envelope plasma IgA levels were associated with risk of infection. Moreover, in vaccinated individuals who became infected, no difference in markers of disease progression was observed [140].

In contrast to the important role of HIV-1-specific $CD8^+$ and $CD4^+$ T cells in elite controllers, several studies have shown that neutralizing antibodies are less frequently found in elite controllers than in viraemic progressors [136, 119]. Based on these data, Env-specific Abs are likely to be important for protection against HIV acquisition, however the role of Abs against

acquired HIV-1 infection may be less essential for immune control.

1.3.3 Other cellular defenses

DCs act as an intermediate link between the innate and adaptive immune responses, and as such they play an important role in the immune control against pathogens. DCs are located in all mucosal surfaces where they come into contact with pathogens such as HIV-1. Mature DCs travel to the lymph node, where they present processed antigen to T and B cells, triggering the adaptive immune arm of the immune system. On the other hand DCs have been shown to mediate cell to cell transmission of HIV-1 through close contact formation of a virological synapse (VS) or via the exosome secretion pathway after HIV-1 has been captured. DCs have been described as having an important role in the early transmission of HIV-1 to $CD4^+$ T cells. Further characterization of cellular and viral factors that regulates DC-mediated transmission of HIV to $CD4^+$ T cells is required, as they may be important targets for therapeutic strategies [28].

Pathogen associated molecular patterns

The innate immune system is the first line of host defence during HIV-1 infection [2]. This response relies on recognition of conserved structures on pathogens or pathogen associated molecular patterns (PAMPs) through patter recognition receptors (PRRs) such as the Toll-like receptors (TLRs) [2].

PAMPs are considered to be invariant among pathogens, as well as being essential for their survival and distinguishable from self [96].

Upon PAMP recognition, PRRs signal to the host the presence of pathogens and trigger pro-inflammatory and antimicrobial responses, which ultimately result on the activation of gene expression and synthesis of a broad range of pro-inflammatory molecules and immunosupressors [96].

The pro-inflammatory pathway induced by PRRs activates de immune response and also play a role in activation, maturation an shaping of the adaptive immune response. Overall PRRs regulate leukocyte recruitment to the site of infection by activation of several cell types

such as resident DC and macrophages. Recruitment is a complex process involving a series of interactions with integrins expressed in the vascular epithelium. Sequential adhesion, rolling and migration is highly regulated through expression of cell adhesion molecules and integrins in response to PRR.

1.3.4 Host Genetics

It is known that a small fraction of individuals appear to be resistant to HIV-1 infection, even after repeated exposure [83]. Additionally, a small fraction of HIV-1 infected individuals remain disease free for long periods of time (LTNP). The existence of these individuals have given hope to the design of an effective HIV-1 vaccine that may elicit immune responses capable of both provide an individual with protection against HIV-1 infection, or if infection does occur, elicit effective immune responses capable of controlling viral replication and therefore stop disease progression and limit transmission to healthy individuals. Nevertheless, studies of these populations of individuals have shown them to be heterogeneous, and not a single universal correlate of immune protection has being found.

Variation on disease progression arise from the complex interactions between viral and host genetics, as well as the environment. Defining these interactions may help to better understand HIV-1 pathogenicity and help design an efficient vaccine [83]. Comparison of elite controllers and progressors has concluded that there is not a single factor that can explain the variation in disease outcome observed in HIV-1 infected individuals. It is thought that the difference between these groups is the sum of multiple mechanisms involving the interaction between the host and viral genetics.

With regards to the host genetics, their importance has been tested by genome wide association studies (GWAS). Thus far, genes encoding the HLA class I loci have been shown to have the strongest effect on HIV disease outcome [41]. These results add to the mounting evidence that show the importance of HIV-1 specific $CD8^+$ T cell responses in disease control. Population studies show that protective HLA alleles are found in 67% of elite controllers and only 37% of progressors or HIV-1 negative populations [109]. However, the majority of individuals with protective HLA alleles still develop progressive disease [136].

On the other hand, HLA class I molecules also serve as ligands for the KIRs, which control NK cells activity. Beside their role in $CD8^+$ T cell and NK cell activation, HLA also interact with members of the Immunoglobulin like receptor family (LILR) that is expressed and regulates dendritic cells function [62]. Function of myeloid DCs appears to be better in elite controllers, while Toll Like receptor-dependant secretion of pro-inflammatory cytokines appears to be decreased [61].

Several other immune properties have being linked with variation in disease progression, this includes the activation receptor KIR3DS1 in conjunction with HLA BW4-801, which have being shown to have a positive effect on HIV-1 disease outcome. However it has not being shown that this combination is enriched in Elite controllers [107].

Up regulation of host protein p21 or cyclin-dependant kinase inhibitor (CDKNIA) has being suggested to inhibit HIV-1 replication by abrogation of effective HIV-1 transcriptional elongation [24, 118].

Macrophages from elite controllers have also being suggested to resist HIV-1 infection when compared to progressors [118].

The outcome in disease progression is therefore thought to be shappen by several factors. For example, HIV-1 infection with an unfit virus to an individual with HLA protective alleles that restrict potent $CD8^+$ T cell responses, clonotypically polyfunctional and less prone to exhaustion; with B cells producing broadly neutralizing antibodies; dendritic cells that lead to activation but with out extensive inflammatory responses and cytotoxic and polyfunctional NK cells are likely to have a better disease outcome than an individual infected with a fit virus and with defective immune responses. Understanding how to elicit these effective responses involving the several arms of the immune system is of importance for vaccine design. This thesis focuses on the role of $CD8^+$ T cells in the control of HIV-1 infection.

1.4 HLA and CD8$^+$T Cell role

1.4.1 Association

We focus on CD8$^+$ T cell responses as several lines of evidence point to the central role of HLA molecules and CD8$^+$ T cells in immune control of HIV-1 infection [56]. These include: the temporal association between emergence of HIV-specific CD8$^+$ T-cell responses and reduction in acute viraemia [71, 15]; macaque/SIV model studies which elegantly demonstrated that anti-CD8 monoclonal antibodies prevent the decline in viral load following peak viraemia [123]; viral escape from the dominant CD8$^+$ T-cell response precipitating loss of immune control [56, 39]; HLA class I homozygous disadvantage, suggesting a broad CD8$^+$ T-cell response is beneficial [22]; the association of certain HLA class I alleles (such as B*57) with slow disease progression; and of others (such as B*35) with more rapid disease progression [22, 106, 68]; that HLA-B alleles make the greatest contribution to the HIV-specific CD8$^+$ T-cell response, impose the strongest selection pressure on HIV and have greatest impact on outcome [68]; and recent GWAS of Caucasians [41] show that the major genetic variation affecting viral steeping and/or rate of disease progression is within the HLA.

Additionally, the protective role of major histocompatibility (MHC) driven CD8$^+$ T cell responses has been demonstrated by the vaccination of Indian rhesus macaques expressing the protective MHC Mamu-B*08 allele, with Mamu-B*08-restricted CD8$^+$ T cell epitopes [59]. The vaccinated animals controlled the replication of a pathogenic SIV clone after challenge.

On the other hand, HIV specific CD8$^+$ T cell responses are detected in individuals with AIDS and very few HLA class I alleles have a consistent effect on disease progression. Therefore we hypothesize that there are important differences in the CD8$^+$ T cell responses to epitopes based on HLA alleles presentation that are of importance for disease progression outcome among individuals [55].

1.4.2 HLA Class I

HLA class I molecules are encoded by the genes located within the MHC region on chromosome 6. This is the most diverse region within the human genome with >3000 HLA-B alleles, >2400

HLA-A alleles and >2000 HLA-C alleles described by October 2013 (http://www.ebi.ac.uk/imgt/hla). Most variation is located in the region encoding the peptide binding groove and therefore dictate the binding motif required for the peptides presented by each allele. The binding groove is formed of six pockets (A-F) that differ in size and electrostatic and hydrophobic potentials. Two of these pockets exhibit a particular preference to one or two amino acids and are known as primary anchor sites, the rest of the pockets are more flexible and act as secondary anchors fine tuning the peptide-MHC interactions [19]. The main function of the MHC molecules is to present endogenous antigen peptides to T cells to signal infection and lead to activation. Therefore the variety of HLA allotypes on a cell surface determines the range of peptides that can be presented. Hence individuals with an heterozygous locus will present a broader variety of peptides compared to an individual that is homozygous, thus a greater immune pressure is exerted on the pathogen. Indeed an advantage has been shown on HIV-1 disease progression in heterozygous individuals when compared to homozygous [22].

HIV-1 disease progression has been associated with HLA loci through GWAS and epidemiological studies. Studies have shown that the majority of the HIV-1 specific $CD8^+$ T cell responses are restricted by HLA-B alleles and that differences in median VL are associated with different HLA-B alleles [68]. One of the most consistent protective associations is to HLA-B*57, the mechanism involved in protection are not well understood, however it has been suggested the HLA-B*57 restricts broad $CD8^+$ T cell responses located in Gag, a highly conserved region of the HIV-1 genome, and that escape mutation of these peptides have a significant impact on viral fitness [74, 18]. Additionally HLA-B*57 is a ligand for KIR3DL1/S1 and thus regulates NK cell function in a way that may contribute to protection. HLA-B*27 is another allele that has consistently been associated with protection in the context of HIV-1 infection. Protection by HLA-B*27 is exclusively driven by a response to a single Gag epitope known as KK10 (KRWIILGLNK, Gag 263-272). Escape from KK10 requires a complex pattern of compensatory mutations, but once it is achieved the virus replicates efficiently, similar to the wild type virus and the mutations are stable upon transmission [56, 53, 66, 39, 124, 125]. Other alleles previously described as having a protective effect are B*51:01 [76], B*58:01 [68], and B*81:01 [137].

On the other hand, HLA-B*35 has consistently been associated with HIV disease progression [22, 47]. The effect by HLA-B*35 alleles has been observed to differ between subtypes.

HLA-B*35 alleles will be further discussed in chapter 3 where the impact of HLA-B*35 allele differences in disease outcome in a B clade infected cohort from Mexico is analyzed .

1.4.3 CD8$^+$ T cells

Clonotypic αβ T cell receptors (TCRs) recognize peptide bound to MHC class I molecules on the cell surface of infected cells. The specificity of the TCR repertoire is defined during thymic development through positive and negative selection [135]. Approximately 1x10^8 different TCRs remain. TCRs are inherently cross-reactive and exhibit a degree of flexibility that allows them to mold around a variety peptide-MHC (pMHC) complexes [135]. The degree of flexibility remains unpredictable and can vary among complexes [134].

TCR recognizes short peptides usually 8 to 11 amino acids in length derived from proteins degraded by multi catalytic proteosome in the cytoplasm, cleavage usually occurs after basic or hydrophobic amino acids. Fragments are then transported to the endoplasmic reticulum (ER) by transporter associated with antigen processing (TAP). Peptide transport is swayed by the side chains of the first three residues in the N-terminus, with particular amino acids either enhancing or reducing binding. Proline at the second aminoacid of the peptide (P2) nearly abolishes binding to TAP, yet approximately 20% of MHC class I molecules use proline at P2 as anchor. It has been suggested that presentation of such peptides require transport of larger precursors [57]. Finally the peptide is trimmed by ER amino peptidase (ERAAP). ERAAP is unable to cut the bond of the amino side of proline, and therefore there is an accumulation of XPX peptides.

The magnitude, breadth, specificity, polyfunctionality and avidity of HIV-1 specific CD8$^+$ T cell responses may all be important for the control of HIV-1 infection. A difference in frequency of CD8$^+$ T cell responses has been shown between LTNP and normal progressors [115], additionally an inverse correlation between CD8$^+$ T cell numbers and VL has also been observed [36]. Nevertheless not all CD8$^+$ T cell responses are effective and a distinction between quality over quantity has been suggested [63].

Priming of CD8$^+$ T cell responses and Mechanisms of CD8$^+$ T cell dysfunction

The generation of effective and durable HIV-1 specific CD8$^+$ T cell responses is a key challenge

to the design of an effective HIV-1 vaccine [141]. The functionality of $CD8^+$ T cell responses is formed by multiple factors including the duration and strength of antigenic stimulation, interaction with $CD4^+$ T cells and APCs, co-stimulation requirements, and cytokine and chemokine availability. These influence the transcription factors that regulates differentiation and phenotypic properties of $CD8^+$ T cells and collectively function to determine the development of effector T cells and establishing the memory pool [89].

Naïve $CD8^+$ T cells become activated through the interaction with APCs displaying cognate peptide complexed with MHC class I. The interaction between naïve $CD8^+$ T cells and APCs drives a TRC signal (signal I) as well as co-stimulation (signal II). These signals occur in the immunological synapse (IS) consisting of TCR-pMHC complexed with CD28/CD80-CD86 interactions. As the cells couple these signal molecules co-localize within the plasma membrane forming a central supramolecular activation cluster (CSMAC). Adhesion molecules then surround the CSMAC forming a peripheral supramolucular activation cluster (PSAMC) [89].

Stable formation of the IS allows strong signalling to the reacting $CD8^+$ T cell and keeps it in close proximity to the APC thus allowing for the cell to receive signal III in the form of cytokines such as IL-12 and other inflammatory mediators that influence the functional attributes of the developing response [122]. Signalling via IL-12 supports proliferation and development of citolytic activities in $CD8^+$ T cells in vitro.

$CD4^+$ T cells promote activation of APC via interaction with co-stimulatory molecules and by secretion of cytokines. Though deceivingly potent primary $CD8^+$ T cell responses may be induced in the absence of CD4 $^+$ T cell help, the subsequent establishment of long lived functionally robust memory $CD8^+$ T cells is compromised [46].

Acute infections often cause a massive antigen specific $CD8^+$ T cell response that contributes with successful clearance [32]. However, chronic responses may arise and are often associated with the development of phenotypically and functional inferior responses [46, 8]. T cell responses generated under this conditions are susceptible to exhaustion and are incapable of clearing the pathogen.

$CD8^+$ T cell exhaustion is characterized by a progressive, stepwise loss of function that can result in the production of defective $CD8^+$ T cells or in their physical deletion. $CD8^+$ T cell exhaustion often develops during chronic viral infection due to antigen persistence.

The ability to produce IL-2 and undertake robust proliferation appears to be highly sensible to exhaustion and are the first functions lost by $CD8^+$ T cells during chronic infection. Despite their loss of functionality, exhausted cells express patterns of surface receptors commonly associated with an activated effector state. Exhausted cells however also express a group of inhibitory receptors, such as PD-1 [46].

PD-1 is the best-studied inhibitory receptor associated with exhaustion. PD-1 functions to add peripheral tolerance and exhaustion, it inhibits proliferation and function of T cells. In HIV-1 infection, PD-1 expression positively correlates with VL and inversely correlates with $CD4^+$ T cell counts. In LTNP the levels of PD-1 are lower and these cells are more functional [113].

Furthermore, $CD8^+$ T cells from elite controllers appear to be more functionally competent. Variation in proliferative capability has also being shown in HIV-1 specific $CD8^+$ T cell responses from elite controllers [92]. Additionally, the ability to simultaneously carryout several effector functions (polyfunctionality) has also being linked with effective $CD8^+$ T cell responses in elite controllers [142, 13, 102].

Immunosuppressive cytokines, such as growth factor B (TGF-B) have being suggested to play a role in regulating the size of both the effector and memory $CD8^+$ T cell pools and therefore may have a role in cell exhaustion.

HIV-1 specific $CD8^+$ T cells restricted by protective HLA alleles, such as B*57, appear to be less susceptible to inhibition by T reg cells [136].

The selection of more cytolitic TCR clonotypes has also being suggested as a key feature of effective HIV-1 specific $CD8^+$ T cells. This suggests that structural interactions between TCR and pMHC may play an important role in the priming of effective immune responses to HIV-1 infection [25].

1.5 Viral evolution in host

1.5.1 HIV/HLA coevolution

Sequence variation of HIV-1 among infected populations is not random. It occurs as a consequence of Darwinian selection resulting from immune pressure of the host driving escape mutations. Stable transmission of escape mutations has previously been described and HIV-1 follows a predictable mutational pattern in response to specific HLA alleles, these are known as HLA footprints [53]. Upon transmission these changes may persist and the escape virus of an individual may become the wild type (WT) virus of the new recipient and thus allowing for accumulation in the population as the epidemic progresses.

The difference in genetic backgrounds in distinct populations may influence HIV-1 evolution in different ways. It has been shown that frequency of HLA driven HIV-1 polymorphisms is correlated with the distinct HLA frequency distributions in the study populations [64].

The role of reverting mutations has previously been highlighted as being important in immune control of HIV-1. Possibly due to the fitness cost imposed on the virus [88]. This in turn suggests that effective CTL responses drive escape mutation that cripple the virus.

1.5.2 Impact of HIV-1 escape within host

Presentation of viral peptides by HLA class I molecules on the cell surface allows for priming of $CD8^+$ T cells and the killing of infected cells. This creates a robust selection pressure for mutation that abrogates peptide recognition, the so-called escape mutations. These variants arise from amino acid substitutions in $CD8^+$ T cell epitopes, or proximal sites to epitopes, which abrogate viral peptide processing, binding and presentation or TCR recognition. $CD8^+$ T cells escape was first described in 1991 [111], however the consequences on HIV-1 evolution and disease progression remain to be fully understood. A clear example of progression by escape from an HIV-1 specific $CD8^+$ T cells response is the R264K change in the HLA-B*27 restricted Gag epitope KK10 [39, 53], which leads to the loss of immune control of HIV-1 infection. The majority of escape mutations, however, do not seem to affect viremia. Although it has been shown that a greater number of polymorphisms in Gag lead to better control of the disease [87].

The fate of escape mutations upon transmission to a new host depend on the interplay between host and viral genetic factors and the cost of escape on the viral replicative capacity (VRC). Mutations passed to a recipient that is HLA matched to the restricting allele have been shown to be detrimental [39, 53], because the virus is already adapted to the recipient's alleles. For HLA mismatched recipients, the transmission of escaped-virus does not affect epitope presentation. In fact studies have shown that if these escape variants have a fitness cost, their transmission can give an advantage to the new host, as the virus is crippled to an extent [52, 131].

Transmission of escape mutations with fitness cost are thought to revert to WT rapidly, despite reversion, transmission may be long term beneficial to a mismatched host by leading to lower viremia in acute infection.

1.5.3 Impact of HIV-1 escape within a population

Is has been suggested that HIV-1 adapts more readily to HLA alleles that are frequent in the population. To this end, a positive correlation between VL and the frequency of HLA class I alleles has been demonstrated [133]. Additionally the positive correlation between the frequency of the restricting allele and the frequency of the escape mutant has been shown [64].

Kawashima's study on adaptation of HIV-1 to HLA gave the most compelling evidence that the accumulation of escape mutations at a population level is occurring. The study comprised a large set of data including data from 8 diverse populations in 5 different continents. Data presented suggested that if an alleles is frequent in a population and produces robust immune pressure, the shared external force can lead to the shared accumulation of escape mutations on a population overtime and may lead to the establishment of new consensus sequences [64].

HLA associations with disease outcome may therefore change overtime, an example of this is the effect of HLA-B*51 in Japan. HLA-B*51 was suggested to be protective in 1983, however this observation was no longer true in individuals infected between 1997 and 2008. The loss of protection by B*51 is associated with the accumulation of the escape variant RT-I135V within the HLA-B*51 restricted epitope TAFTIPSI (RT 129-135) in the population. The high accumulation of the escape mutation is related to the high frequency of the restricting allele in the population [64].

The genetic diversity of HIV-1 in a population will also influenced by its descent (founder effect) [14].

1.6 Vaccine strategies

Diverse vaccine concepts and vaccination strategies have been studied in clinical trials without success. Strategies tested include DNA vaccines, subunit vaccines, live vectored recombinant vaccines and prime-boost combinations. Lack of success is due to the high genetic variability of HIV-1, the insufficient knowledge of immune correlates of protection, the difficulty of generating broadly neutralizing antibodies and the absence of a relevant animal model [54, 55]. Nevertheless, much can be learned from these studies, here I introduce some vaccine modalities previously and currently tested.

Vaccine modalities

Live attenuated virus are modified to be infectious but pathogenically attenuated, its use is not employed for HIV-1 vaccine design because of pathogenicity issues. Animal studies have shown that animals treated eventually developed AIDS [7].

Inactivated virus have being shown to be effective in animals that are challenged with the strain identical to the vaccine strain however is unsuccessful against other strains. These vaccines have being shown to fail in eliciting broad an vigorous protective responses [99].

DNA vaccine strategies involve the insertion of HIV-1 genetic material into known plasmids. These genes are expressed into proteins in patients treated, viral proteins are degraded into peptide and presented through MHC class I. No DNA vaccine has being approved for use in humans by the FDA due to safety issues, such as integration-induced mutagenesis.

Chemically synthesized HIV-1 peptides or proteins have being studied and shown modest protection only when challenged with the same vaccine virus, antibody responses were modest in titters and limited in spectrum of HIV-1 isolates they could neutralize. Addition of adjuvants in required to enhance immunogenicity [12].

Live recombinant vectors are constructed by insertion of HIV-1 genes into live, replication competent but not disease causing vectors. Upon infection, immunity occurs against both the

product inserted and the carrier vector. Recombinant vectors are particularly useful for eliciting cellular mediated responses, since the HIV-1 proteins are produced intracellulartly and therefore enter the MHC class I processing pathway [37].

Vaccinia, a replication competent virus used in the eradication of smallpox, have shown to elicit potent $CD8^+$ T cell responses. However it has being shown to cause disease in immuno-compromised patients, sometimes causing fatal encephalitis. Modified vaccinias have therefore being produced and being studied, an example is the MVA, which has deletions that leave the virus highly attenuated [127].

Canarypox, an avian virus that undergoes an abortive cycle of replication in human cells but initiates synthesis of viral proteins has being shown to be safe, immunogenic inducing both antibodies and cellular responses [38].

Adenovirus has being attenuated by deletions or inactivation of the E1 and E3 genes, they illicit high titter antibodies and high frequency $CD8^+$ T cell responses in animals. However pre-existing antibody responses to the adenovirus dampens in vivo expression and therefore immunogenicity. This can be improved by priming with plasmid DNA vaccines before immunization. Alternatively constructs developed using non-human primate isolates may be used, although immunogenicity in humans most be tested [129].

The development of an immunogen that elicits anti-HIV-1 broadly neutralizing Env responses is one of the mot difficult tasks for vaccine design.

It has previously being shown that neutralizing epitopes are shielded from the immune system during infection [73], efforts are therefore being made to develop subunit immunogens that mimic the fusion of the virus into the cell membrane and drive the conformational changes that allow epitope access by the immune system.

In addition, a common issue with vaccines tested is HIV-1 diversity. An attempt to overcome this problem includes the design of mosaic sequences for insertion into a viral vector and that covers potential T cell epitopes from world wide HIV-1 strains [43]. This approach may allow for vaccines to elicit immune responses capable of recognising virus from multiple clades. Studies in rhesus macaques have shown that mosaic sequences increase breadth and depth of T cell responses compared with consensus or natural sequence [10]. However, HIV-1 rapid adaptation may contribute to the eventual failure of these vaccines. It is suggested therefore that vaccine-

induced responses should target vulnerable HIV-1 genomic regions, where escape results in high fitness cost to the virus and with a complex and difficult selection of compensatory mutations. The analysis of such regions is therefore imperative.

T cell based vaccines have not been sufficiently immunogenic. The STEP trial (Merck) of a prophylactic trivalent adenovirus conjugate vaccine failed to find protection conferred by inducing small numbers of $CD8^+$ T cell responses (median one response per vaccine), and was consequently discontinued prematurely [98]. Results from the recently 'Thai Trial' RV144 showed a decrease in virus acquisition of 31.2% (p=0.04) based on a modified intention-to treat analysis [67].

Nonetheless, successful immune control of HIV-1 infection by some individuals, although rare, is well-described and possible to achieve [54]. The understanding of which immune responses are responsible for controlling HIV-1 replication is therefore the aim for the development of a $CD8^+$ T cell-based HIV-1 vaccine.

The ultimate configuration of an effective HIV-1 vaccine remains uncertain. Since HIV-1 is able to be transmitted both as cell bound or cell free virus, an effective HIV-1 vaccine may need to elicit several arms of the immune system, including antibody responses capable of neutralizing free virions as well as cell-mediated immune responses capable of eliminating infected cells [77]. It is therefore likely that a combination of vaccine modalities will be required. A live recombinant vector or plasmid DNA could be used to elicit $CD8^+$ T cell responses while subunit immunogens may be used to elicit antibody responses.

1.6.1 Correlates of immune protection

To date no therapeutic regimen has been shown to eradicated HIV-1 infection fully or induced a complete recovery of the host immune system. However recent studies have highlighted the potential for a 'functional cure' through early initiation of ART. One of the most high profile reports is that of a child in Mississippi (Mississippi child) that started ART treatment 31 hours after birth to an HIV infected mother. She was found to be HIV positive by two different blood tests (20, 000 RNA copies/ml) through standard HIV testing before treatment was initiated. VL test was repeated and was positive at days 7, 12 and 20. Subsequent tests at day 29 and

up to 26 months the VL remained undetectable. Interestingly treatment was stopped after 18 months due to loss of follow up. Upon re-admission of the child, VL test was found negative. A year since treatment was stopped, and although low levels of HIV-1 DNA has been detected in the child's cells, there is no evidence of viral replication [31].

Although it is still early days in the case to draw any conclusions, the Mississippi child has given hope to the field that a cure may in fact be possible. It has been suggested that the child's early treatment, may have limited the formation of a latent reservoir in long-lived central memory $CD4^+$ T cells and hence may explain treatment success, although there is no evidence of this. Regardless the report of the Mississippi child has had important consequences on paediatric care.

With regards to adult infection, a great variability in HIV-1 disease outcome is observed among individuals. 10% of HIV-1 infected patients progress to AIDS with in 5 years of infection (rapid disease progressors (RP)) [110], while 5% maintain normal $CD4^+$ T cell counts and undetectable VL for over 10 years LTNP [60].

The mechanisms involved in differences in disease outcome remain unclear. However certain factors have been described to influence HIV-1 disease progression. Variation in the CCR5 coreceptor of the host is one of them, a 32-bp deletion (CCR5$\triangle$32) encodes a truncated version of the receptor, that is not expressed on the cell surface, and thus homozygous individual exhibit a strong resistance to HIV-1 infection [80]. Immune activation appears to play a critical role on disease outcome, data on apathogenic SIV infection on Sooty Mangabeys and African Green Monkeys, suggest that protection against disease occurs mainly due of their capacity to rapidly downregulate innate immune responses, which include type I interferon production, after acute infection. The lack of chronic immune activation avoids disruption and loss of functionality of the lymphoid tissue, including the memory T-cell pool [33, 130]. Other factors that influence disease outcome include differences in HLA type molecules and viral fitness. Viral fitness refers to pathogenicity of a virus. The viral genetic background will influence its ability to produce infectious progeny in a specified environment. Viral replicative capacity (VRC) is studied to measure a given virus ability to replicate [119].

Results from GWAS have lead to a better understanding of the impact of HLA class I alleles on control of HIV-1 infection. Single-nucleotide polymorphisms (SNPs) in the genes encoding

for the HLA-B and HLA-C alleles are the strongest host genetic factors that are associated with immune control of HIV-1 [41, 40]. Four HLA class I polymorphisms in Caucasians have been shown to be associated with control. These are located in the region encoding the HLA-B binding groove and the association may therefore be related to the presentation of peptides by MHC molecules and their interaction with TCR. However only a 20% difference between VL in HIV infected individuals can be explained by host genetic factors [136].

The incredible diversity of HIV-1 is the main hurdle for immune control. Based on the positive results from the RV144 Thai trial, Env Abs may be good candidates to block HIV-1 acquisition. However studies carried on acquisition protection did not address HIV-1 diversity. The immunogen utilized in the RV144 trial matched the local Thai consensus. The level of protection conferred by Env-specific Abs in the RV144 trial may therefore not be as substantial against different HIV-1 clades [114]. Because of the genetic diversity of HIV-1 and likely escape from the Ab response, the incorporation of immune responses capable of controlling viral replication remains essential. As described before there is mounting data supporting the role of HIV-1 specific $CD8^+$ T cell responses in the control of viraemia.

Despite the supporting data on different arms of the immune system in control of HIV-1 infection, a vaccine is yet to be found. Thus far no significant impact on viraemia has been observed in vaccinated individuals in any clinical trial. Nevertheless, HIV-1 infected elite controllers give evidence that effective and durable control of HIV-1 infection is achievable in humans. Although this group of individuals (LTNP) is a heterogeneous group, data suggest that $CD8^+$ T cell responses are an important component of the immune control of HIV-1. Better understanding of what $CD8^+$ T cell responses a vaccine needs to induce to achieve durable control of HIV-1 infection is imperative and the ultimate goal of this thesis.

1.7 Aims of the study

The aim of these studies is to help define the $CD8^+$ T cell responses a vaccine needs to induce to achieve durable immune control of HIV-1 infection. The present projects focus on the following aims:

1. Through study of a new cohort in Mexico, to develop insights into immune control of B and C clade HIV-1 infection;

2. To investigate the impact of HLA-mediated selection pressure on population-level viral fitness;

3. To develop a C-clade backbone as a research tool, to determine the impact of clade-specificity on viral fitness effects of specific HLA-mediated escape mutants.

Chapter 2

Materials and Methods

2.1 Study cohorts

Mexican cohort (chapter 3 and 4)

B-clade HIV-1 infected treatment-naïve adults from Mexico were recruited from the Centre for Research in Infectious Diseases, Mexico city, Mexico (n= 926). VL and $CD4^+$ T cell counts were determined by real-time PCR (Abbott, UK) and multi-parametric flow cytometry (BD, US), respectively. VL and $CD4^+$ T cell counts were available for 926 and 919 study subjects respectively. VL and absolute $CD4^+$ T cell count measurements were obtained at study entry (baseline) for all individuals. The median VL of this cohort was 40,878 HIV RNA copies/ml plasma (inter quartile range (IQR) 10,334- 135,043), and median $CD4^+$ T cell count 380 cells/mm^3 (IQR 151-592). All study subjects provided written informed consent and ethical approval was obtained from the relevant institutional boards.

Barbados cohort (chapter 4)

190 Antiretroviral therapy-naïve study subjects with B clade infection were recruited from Bridgetown, Barbados. All study subjects provided written informed consent and ethical approval was obtained by the Barbados Ministry of Health, and by the Oxford Research Ethics Committee. VL and $CD4^+$ T cell counts were determined by Roche Amplicor version 1.5 assay and multi-parametric flow cytometry (BD, US), respectively. VL and absolute $CD4^+$ T cell count measurements were obtained at study entry (baseline) for all individuals. VL and $CD4^+$

T cell counts were available for 176 and 171 study subjects respectively. The median VL of this cohort was 14,200 copies/ml plasma (IQR 2,775- 59,275), and the median $CD4^+$ T cell count was 403 cells/mm^3 (IQR 325-569).

Badalona patients

Three HIV-1 B clade infected individuals with HLA-B*35:08, were recruited from a cohort in Badalona, Spain. Individuals were recruited from the Hospital Germans Trias i Pujol. All three subjects gave their written informed consent to participate. Ethics approval was obtained from the Oxford Research Ethics Committee.

2.2 Processing of samples

2.2.1 Separation of plasma and PBMCs

Blood samples were collected in EDTA and centrifuged for 10 minutes at 894.6 G-force for the separation and the removal of plasma. Plasma was stored at -80℃ in 1 ml aliquots for future RNA extraction. The remaining blood was diluted 1:1 with phosphate-buffered saline (PBS, Invitrogen, UK) and layered onto Lymphoprep (Axis-Shield, Norway) at a ratio of 2:1. Samples were centrifuged for 30 minutes at 894.6 G-force, with the break off on the centrifuge. The resulting layer of peripheral blood mononuclear cells (PBMCs) was harvested and washed three times with RPMI (RPMI 1640, Sigma-Aldrich, UK) followed by resuspension in 20 ml R10 culture media. PBMCs were counted by hemocytometer and used in cellular assays immediately or were cryopreserved.

2.2.2 Cryopreservation and reconstitution

Cells to be frozen were pelleted in a 4℃ chilled centrifuge for 5 minutes at 894.6 G-force. Supernatant was discarded and cells were resuspended at 10 million cells/ml of freezing media (10% Dimethyl sulfoxide (DMSO) in filtered heat inactivated Fetal Calf Serum (FCS)). Vials were placed in a 'Mr Frosty' stratacooler box, and placed at -80℃ overnight. Cells were then transferred to liquid nitrogen for storage.

Cells were thawed rapidly in a water bath at 37 ℃. 20 ml of pre-warmed R10 media was added slowly to wash diluted cells. Cells were centrifuged for 5 minutes at 894.6 G-force. Two more washes followed and cells were resuspended in growth media and left to rest at 37 ℃.

2.2.3 Cell counts

Cell counts were carried out by hemocytometer in a 1:1 ratio with Trypan blue.

2.2.4 DNA extraction

Chromosomal DNA was extracted for HLA typing and pro-viral sequencing from whole blood utilizing the Puregene DNA kit (Gentra, UK). In short, 3 ml of whole blood were incubated in 9 ml of Red Blood Cell Lysis for 10 minutes at room temperature, followed by centrifugation for 10 minutes at 894.6 G-force. Supernatant was discarded and the cell pellet was resuspended in 3 ml of cell lysis buffer. The suspension was left at room temperature for at least 3 days. 1 ml of protein precipitation solution was then added followed by centrifugation at 894.6 G-force for 10 minutes. Supernatant was harvested and mixed with 3 ml isopropanol to precipitate DNA. After centrifugation at 894.6 G-force for 3 minutes the DNA was washed with 70% ethanol and dried. DNA was then resuspended in 250 µl RNase/DNase-free H_2O (Sigma-Aldrich, UK) and stored at -20 °C.

2.2.5 RNA extraction

Viral RNA was extracted from plasma or viral stocks using a RNA extraction kit from Qiagen and following manufacturer's instructions. RNA was eluted in RNase/DNase-free H_2O (Sigma-Aldrich, UK) and stored at -80 °C.

2.2.6 HLA typing

Four-digit HLA typing of the class I locus for the samples was performed from genomic DNA as previously described [35] by sequence-based typing. Samples from Barbados and Barcelona where typed at the ASHI* accredited HLA typing laboratory, University of Oklahoma Health Sciences Centre, USA [20]. Samples from Mexico were typed by the Reyes-Teran lab, Mexico

City, Mexico with the same method. Exons 2 and 3 of HLA class I were amplified by locus-specific PCR and then sequenced. Resolution of ambiguities was undertaken according to the ASHI committee recommendations.

2.3 Amplification of HIV-1 genes by PCR

2.3.1 DNA quantitation

Quantitation of DNA was carried out using a ND-1000 spectophotometer (Labtech International). 2 µl of DNA was utilized for quantitation.

2.3.2 Primers

Primer sets (Table 2.1) were designed to bind conserved regions of the HIV genome and were checked for specificity and melting temperatures (Tm) utilizing the Los Alamos data base QUICKAlign tool. Primers were purchased from Eurofins MWG Operon. Specific regions were amplified from both proviral DNA and cDNA using the nested PCR method.

Fragment	Primer name	Sequence	Aligment site to HXB2	Bases
1.Gag	Frag1_2R	CTGCAGAATGGGATAGTTACAT	1415-1437	22
	Frag1_3F	GACACCAAGGAAGCCTTAG	1075-1093	19
	Frag1_4R	CTCCCACTGGAAtAGGTG	1550-1567	18
	Frag1_5F	GGAACAAATAGCATGGATGAC	1521-1541	21
2.Pol	Frag2_2R	ATCATCTGCTCCTGTGTCTA	2323-2342	20
	Frag2_3F	AGGCCAGGGAATTTCCTTCA	2119-2138	20
	Frag2_4R	CATTGTTTAACCTTTGGGCCATC	2598-2620	23
	Frag2_5F	CCCATTGAAACTGTACCAG	2559-2577	19
3.Pol	Frag3_2R	TGGGTCATAATATACTCCATGTA	3490-3512	23
	Frag3_3F	CAGACATAGTACCACTAACTGAA	3418-3440	23
	Frag3_4R	CATAGAAAGTTTCTGCTCCT	3854-3873	20
	Frag3_5F	CTCCCCTAGTAAAATTATGGTA	3808-3829	22
4.Pol	Frag4_2R	GATTCACTCTTATCTGGTTGTG	4072-4093	22
	Frag4_3F	CAGACTCACAGTATGCATTAGG	4039-4060	22
	Frag4_4R	ACTGGCCATCTTCCTGCTA	4540-4558	19
	Frag4_5F	AGTGGCTACATAGAAGCAGAGGT	4470-4492	23
5. Vif/Vpr/ Vpu	Frag5_2R	TGTCCTGCTTGATAGTCACA	5437-5456	20
	Frag5_3F	TTTGATTGTTTTGCAGACTCTGC	5374-5396	23
	Frag5_4R	GATGGTTCCAGGGCTCTA	5853-5870	18
	Frag5_5F	ATGGAGCCAGTAGATCCTA	5831-5849	19
6. Env	Frag6_2R	TAAAGTGACACAGAGTGGGG	6592-6611	20
	Frag6_3F	TCAGTTTATGGGATCAAAGCCT	6550-6571	22
	Frag6_4R	GTCTTGTACACGTAATTTCTACAG	7096-7119	24
	Frag6_5F	ACAATGCCAAAACAATAATAGTACA	7060-7084	25
7. Env	Frag7_2R	GTACCGTCAGCGTTATTG	7825-7842	18
	Frag7_3F	TCCTTGGGTTCTTGGGAGC	7780-7798	19
	Frag7_4R	ACCAATTCCACAGATTTTCC	8222-8241	20
	Frag7_5R	TGGGATAAAGAAATTAGTAATTA	8115-8137	23
	Frag7_6R	TCTTTTTTAGTTCCAGACCCCAA	8630-8652	23
8. Nef	Frag8_2R	ACCTGAGGTCTGACTGGAAAG	8997-9017	21
	Frag8_3F	AGATAAACATGGGGCACTTAC	8907-8927	21
	Frag8_4R	AGCAGTCTTTGTAATACTCCG	9395-9415	21
	Frag8_5F	ACCTCAGGTACCTTTAAGAC	9009-9028	20

Table 2.1: **PCR primer list**

2.3.3 RT- PCR

RNA was reverse-transcribed using Superscript Reverse Transcriptase III (Invitrogen, UK), according to manufactures instructions. In brief, specific primers (10 µM), template RNA (0.01 pg-1 µg), 2X reaction Mix and SuperScript III RT/PlatinumTaq were mixed and placed in a

preheated thermal cycler. The following program was utilized: 45 °C for 30 minutes, 94 °C for 2 minutes, 40 cycles of 94 °C of 15 seconds, 55 °C for 30 seconds, 68 °C for 1 minute per kilobase and a final 68 °C for 5 minutes. cDNA was stored at -20 °C.

2.3.4 PCR reaction

Platinum Taq DNA polymerase (Invitrogen, UK) was utilized to amplify different HIV fragments for sequencing. The error rate of the enzyme has been calculated to be between $2x10^{-5}$- $1.1x10^{-4}$ errors/base pair (bp) [9, 81, 132], therefore we expect errors to occur at a frequency of 0.24-1.3 per 1000 bp. 50 µl reaction mix was used for HIV gene amplification:

5 µl	10x reaction buffer
1 µl	5' primer (10 µM)
1 µl	3' primer (10 µM)
1.5 µl	MgCl2 (50 mM)
1 µl	dNTP mix (10 mM)
0.2 µl	Platinum Taq DNA Polimerase
38.3 µl	RNase/DNase-free H_2O
2 µl	Template DNA

The following two step program was performed in a Peltier Thermal Cycler PTC-200 (MJ Research, UK). Specific annealing temperatures (Table2.1) were used for each set of specific primers.

1. Denaturation 94 °C 2 minutes
2. Denaturation 94 °C 15 seconds
3. Annealing T°C + 3 °C 30 seconds
4. Extension 72 °C 1 minute
5. Repeat steps 2-4 19 times
6. Denaturation 94 °C 15 seconds
7. Annealing T °C 30 seconds

8. Extension 72 °C 1 minute
9. Repeat steps 6-9 19 times
10. Final extension 72 °C 10 minutes
11. Cooling 4 °C

2.3.5 PCR purification

PCR products were purified to remove excess dNTPs and primers using one of two methods:

- QIAquick PCR purification Kit (QIAGEN, UK) following manufacturer's instructions.
- ExoSAP-IT (GE Lifesciences, UK, a mixture of Exonuclease I and Shrimp Alkaline Phosphatase), following manufacturer's instructions. In short 2 µl of the enzyme was added per 5 µl of PCR reaction and incubated at 37 °C for 15 minutes followed by 15 minutes at 80 °C.

2.3.6 Sequencing

Purified PCR products and plasmids were sequenced as described previously [75] with BigDye Terminator v3 (Applied Biosciences, UK) and utilizing 6 overlapping primers (Table 2.2). In brief a mixture of 1.875 µl H_2O, 1.875 µl BigDye buffer, 2 µl primer (0.0625 uM), 0.25 µl BigDye terminator v3 reaction mix and 2 µl DNA template was produced. The program in the thermocycler was: 96 °C for 10 seconds, 50 °C for 5 seconds, 60 °C for 2 minutes, for 30 cycles. The sequencing product was then purified by sodium acetate precipitation. Briefly, 10 µL of ddH2O, 2 µL of NaOAC (3 M, pH 5.2, Sigma-Aldrich, UK) and 50 µL of 100% ethanol were added to the sequencing reaction, mixed and incubated at RT for 30 minutes. The mixture was then centrifuged at 3578 G-force for 90 minutes at 4 °C. The supernatant was removed and the pellet was washed twice with 150 µL of 70% ethanol. The pellet was then dried at 94 °C for 1 minute.

Sequences were analyzed on the ABI 3700 DNA analyzer (Applied Biosciences, UK) or alternatively, purified products were sent to Macrogen Ltd, South Korea.

Sequences were aligned using Sequencher version 5 (Gene Codes Corporation) and Se-Al (version 2.0a11, A. Rambaut, Department of Zoology, University of Oxford) software.

Fragment	Primer name	Sequence	Aligment site to HXB2	Bases
1.Gag	Frag1_2R	CTGCAGAATGGGATAGTTACAT	1415-1437	22
	Frag1_3F	GACACCAAGGAAGCCTTAG	1075-1093	19
	Frag1_4R	CTCCCACTGGAAtAGGTG	1550-1567	18
	Frag1_5F	GGAACAAATAGCATGGATGAC	1521-1541	21
2.Pol	Frag2_2R	ATCATCTGCTCCTGTGTCTA	2323-2342	20
	Frag2_3F	AGGCCAGGGAATTTCCTTCA	2119-2138	20
	Frag2_4R	CATTGTTTAACCTTTGGGCCATC	2598-2620	23
	Frag2_5F	CCCATTGAAACTGTACCAG	2559-2577	19
3.Pol	Frag3_2R	TGGGTCATAATATACTCCATGTA	3490-3512	23
	Frag3_3F	CAGACATAGTACCACTAACTGAA	3418-3440	23
	Frag3_4R	CATAGAAAGTTTCTGCTCCT	3854-3873	20
	Frag3_5F	CTCCCCTAGTAAAATTATGGTA	3808-3829	22
4.Pol	Frag4_2R	GATTCACTCTTATCTGGTTGTG	4072-4093	22
	Frag4_3F	CAGACTCACAGTATGCATTAGG	4039-4060	22
	Frag4_4R	ACTGGCCATCTTCCTGCTA	4540-4558	19
	Frag4_5F	AGTGGCTACATAGAAGCAGAGGT	4470-4492	23
5. Vif/Vpr/ Vpu	Frag5_2R	TGTCCTGCTTGATAGTCACA	5437-5456	20
	Frag5_3F	TTTGATTGTTTTGCAGACTCTGC	5374-5396	23
	Frag5_4R	GATGGTTCCAGGGCTCTA	5853-5870	18
	Frag5_5F	ATGGAGCCAGTAGATCCTA	5831-5849	19
6. Env	Frag6_2R	TAAAGTGACACAGAGTGGGG	6592-6611	20
	Frag6_3F	TCAGTTTATGGGATCAAAGCCT	6550-6571	22
	Frag6_4R	GTCTTGTACACGTAATTTCTACAG	7096-7119	24
	Frag6_5F	ACAATGCCAAAACAATAATAGTACA	7060-7084	25
7. Env	Frag7_2R	GTACCGTCAGCGTTATTG	7825-7842	18
	Frag7_3F	TCCTTGGGTTCTTGGGAGC	7780-7798	19
	Frag7_4R	ACCAATTCCACAGATTTTCC	8222-8241	20
	Frag7_5R	TGGGATAAAGAAATTAGTAATTA	8115-8137	23
	Frag7_6R	TCTTTTTTAGTTCCAGACCCCAA	8630-8652	23
8. Nef	Frag8_2R	ACCTGAGGTCTGACTGGAAAG	8997-9017	21
	Frag8_3F	AGATAAACATGGGGCACTTAC	8907-8927	21
	Frag8_4R	AGCAGTCTTTGTAATACTCCG	9395-9415	21
	Frag8_5F	ACCTCAGGTACCTTTAAGAC	9009-9028	20

Table 2.2: List of sequencing primers

2.4 Statistics and Phylogenetics

All statistical calculations were carried out using Prism 5 (Graphpad, US), unless other wise stated. Student t-tests (two-tailed) and Mann-Whitney U tests (two-tailed) were used to compare various clinical parameters and HLA allele or VRC measurements. q-value was utilized to

adjust for multiple comparisons by calculation of a false discovery rate (FDR) for each p-value [21].

2.4.1 Identification of HLA class I associations with HIV amino acid polymorphisms

HLA population frequencies were obtained through the HLA Frequency Analysis tool of Los Alamos database (www.hiv.lanl.gov/content/immunology/tools-links.html). Mathematical analysis was undertaken by collaborator Jonathan Carlson (Microsoft, USA), based on a previously described lineage-correlation approach [14]. The analysis was carried out using a multivariate analysis correcting for phylogenetic relatedness of sequence, for multiple tests, and for linkage disequilibrium between HLA class I alleles, with 5% expected false positives HLA associated viral amino acid polymorphisms. The analysis showed statistical associations between HLA class I expression and HIV sequence polymorphisms, which were used to infer the biological processes of escape and reversion.

- 'Escape' refers to amino acid selection away from consensus in the presence of a specific HLA allele.
- 'Reversion' describes amino acid selection towards consensus in the absence of a specific HLA allele.

2.4.2 HLA linkage analysis

Fisher's exact test was employed to test the association between HLA allele expression and changes within specific residues.

2.4.3 Phylogenetic analysis

Phylogenetic trees were used to check for contamination of samples, to confirm clade of HIV infection and to construct ancestral sequences of the Mexican and Barbados cohorts. Sequences were aligned using Se-Al (tree.bio.ed.ac.uk/software/seal).

Phylogenetic trees were constructed either:

1. By utilizing ClustalX version 2.1 to construct Neighbor-joining (NJ) phylogenetic trees from amino acids with distances calculated by the multiple alignment parameters method and excluding positions with gaps. Bootstrap values are the result of 1000 resampling and are given as percentage.

2. by utilizing GARLI software (Genetic Algorithm for Rapid Likelihood Inference, available at www.nescent.org) to build maximum-likelihood (ML) trees from nucleotides and read bootstrap values from 1000 bootstrap replicates in Bootscore (http://sourceforge.net/projects /bootscore).

3. or by utilising FastML web server to countruct Maximum Likelihood trees using the Generalised Time-Reversible model of substitution, with Gamma distribution and a 0.5 probability cutoff to prefer ancestral indel over character [6].

Trees constructed were visualized by Figtree.

2.5 Molecular biology

2.5.1 Restriction enzyme digestion

Restriction sites in targeted DNA templates were found using the online tool NEBcutter V2.0. Restriction enzymes were utilized to cut plasmid DNA at specific sites. Restriction enzymes were purchased from New England Biolabs and following the manufacturer's instructions. Incubations times raged from 1 to 16 hours at 37 °C depending on efficiency. The restriction enzymes used in this study include MluI, NotI and EcoRI. Digestion was evaluated by 1% agarose gel electrophoresis followed by gel extraction of the correct band when required.

To prevent re-ligation of the cut vector, we employed Shrimp Alkaline Phosphatase (SAP, Ab gene UK) according to manufacturer's instructions.

The vector DNA was then purified by QIAquick PCR purification Kit (QIAGEN, UK) following the manufacturer's instructions and eluted in 30 µl RNase/DNase-free H_2O (Sigma-Aldrich, UK).

2.5.2 DNA ligation

DNA ligation of the target insert with the vector was performed utilizing the T4 Ligase (New England Biolabs) according to the manufacturer's instructions. In brief ligation was carried out at a vector to insert ratio of 1:3 based on DNA concentration, following an incubation of 1-8 hours with T4 DNA ligase and T4 DNA ligase buffer at 25 °C

2.5.3 Agarose gel electrophoresis

PCR products, cut insert DNA and cut vector DNA were visualized by agarose gel electrophoresis. DNA was mixed with gel loading buffer (Bioline, UK) at a 1:5 ratio and loaded onto a 1% agarose (Sigma-Aldrich, UK) gel containing 10mg/ml of Ethidium Bromide in TBE buffer (Tris-Borate-EDTA, Promega, UK).

2.5.4 Bacterial cloning

5 µl of gel purified PCR product was cloned for sequencing using the TOPO TA cloning kit for sequencing from Invitrogen (Invitrogen, UK) following manufacturer's instructions. Plasmids were transformed into chemically competent *E. coli* and plated on ampicillin agar plates (100 µg/ml) and grown overnight at 37 °C. Colonies were picked and plasmid DNA extracted with the miniprep kit (QIAGEN, UK) according to manufacturer's instructions. Plasmids were linearized by EcoRI digestion and the size was confirmed by gel electrophoration.

2.6 Cellular assays

2.6.1 Media and reagents

R10 Medium

R10 medium was utilized to process patients' PBMCs and to culture CEM-GXR cell lines. RPMI 1640 medium (Sigma-Aldrich, UK) was supplemented 10% heat inactivated FCS, 2 µM L-glutamine and 50 U/ml penicillin/streptomycin.

300 µg/ml geneticin (Sigma-Aldrich, US) and 1 µg/ml puromycin (Sigma-Aldrich, US) was the added to the R10 for the culture of Hut-R5 cell line.

H10 Medium

H10 medium was used for the culture of T cell lines (chapter 3). RPMI 1640 medium (Sigma-Aldrich, UK) was supplemented with L- glutamine (146 mg/L), penicillin G (100 IU/ml), streptomycin sulphate (100 µg/mL) and heat inactivated (56 °C for 1hr) human serum to a final concentration of 10% by volume. 10% natural T-cell growth factor (TCGF; Helvetica Healthcare) was added to supplement the H10 and this media was called H10-10.

Luria-Bertani Medium (LB broth)

The LB broth was utilized for the culture of *E. coli* bacteria in liquid medium. The LB broth tablets (Sigma-Aldrich, UK) were diluted in required amounts of deionised water and autoclaved according to manufacturer's instructions. LB broth media is comprised of 10 g enzymatic digest of casein, 5 g yeast extract, 5 g of NaCl and 2 g of inert agent necessary for tableting process per liter. Ampicillin or kanamycin was added to the final concentration of 100 µg/mL (LB/Amp broth) as required.

LB/Agar bacterial culture plates

Ampicillin or kanamycin LB agar plates were used for the culture of *E. Coli* bacteria successfully transformed with anti-ampicillin or anti-kanamycin resistant plasmid (carrying insert of interest) on solid medium on Petri dishes. The agar comprises 9.14 g enzymatic digest of casein, 4.57 g yeast extract, 4.57 g of NaCl, and 13.72 g bacteriological agar per liter. Tablets were dissolved in the required amount of deionised water and autoclaved. Ahead of plating, agar was dissolved in a microwave, and cooled. Ampicillin or kanamycin was then added to the final concentration of 100 µg/mL, and approximately 20 ml of medium poured into 100 mm culture dishes, which were stored at 4 °C.

2.6.2 IFNγ- Enzyme-linked immunospot (ELISpot) assay

A panel of 410 overlapping 18mer peptides spanning the HIV-1 proteome were used for IFNγ-ELISpot screening of HIV-1-specific $CD8^+$ T cells. Optimal peptides specific to the patient's HLA type were also tested. Peptides were supplied by the Massachusetts General Hospital Peptide core facility or by the National Institute of Heath (NIH). In brief, 96-well polyvinylidene difluoride-backed plates (Millipore, GE) were coated with100 µl of anti-IFNγ MAb (Mabtech, SE) overnight at 4 °C. Plates were then washed 6 times with blocking buffer (PBS+ 1%FCS) and peptides were added at the final concentration of 10^{-5}M. Each plate included 4 negative controls (no peptide) and 2 positive controls (10 µl PHA). 50, 000 or 100, 000 PBMCs were added per well, and incubated overnight at 37 °C 5%CO_2. Plates were then washed 6 times with 200 µl/well of PBS; 100 µl of biotinylated ant-INFγ MAb (1µg/ml; Mabtech, SE) was added per well and incubated at RT for 90 minutes. Plates were washed 6 times with PBS and 100 µl of streptavidin-alkaline phosphatase conjugate (1 µg/ml; Mabtech, SE) was added to each well and incubated at RT for 40min. After washing 6 times, IFNγ-producing cells were detected utilizing chromogenic alkaline phosphatase substrate (Bio-Rad, UK).

The AID version 3 Elispot plate reader (AID, Germany) was used to measure the number of spot forming cells (SFC) per well. The background for each assay was calculated by the mean SFC in the negative control wells plus 3 standard deviations (STD) of that mean. Background was then subtracted from the number of spots counted in each peptide stimulated well, responses

of greater than 100 SFC/million PBMC were considered.

2.6.3 HLA class I tetramer assay

For NY10-Gag dual-tetramer staining, PBMCs from subject 4888607 were stained after expansion for 5 days using 10 μg/ml of NY10-260D or NY10-260E in H10. Cells were costained with HLA-B*3508-NPPIPVGDIY (PE conjugated) and HLA-B*3508-NPPIPVGEIY (allophycocyanin (APC) conjugated) tetramers (ex vivo PBMCs) for 30 minutes, and subsequently stained with pretitrated extracellular antibodies CD8-Pacific Blue (BD Pharmingen) and CD3-Pa cific Orange (Invitrogen). Dead cells were excluded using the Vivid LIVE/DEAD marker (Invitrogen, UK) [86]. Samples were acquired on a FACSCalibur instrument within an hour of staining and data analyzed using FlowJo version 8.8.6.

2.6.4 Crystallization and structure determination

HLA-B*35:08 in complex with NPPIPVGDIY or NPPIPVGEIY (Gag-NY10) was refolded by our collaborators in Cardiff University as previously described [27].

2.6.5 pMHC I stability assays

pMHC stability assays were carried out by our collaborators in Cardiff University as previously described [27].

2.7 Viral culture and assays

2.7.1 Generation of Gag-protease NL4-3 recombinant viruses (chapter 4)

Gag-protease NL4-3 recombinant virus stocks were generated as previously described [94] by co-transfection of Tat-inducible, GXR-CEM cells with gag-protease-deleted NL4-3 plasmid and clinically derived gag-protease amplicons from study subjects. Gag-protease-deleted pNL4-3 (pNL4-3Δgag-protease) plasmid was kindly provided by Mark Brokeman, Canada. In brief, viral RNA was isolated from plasma samples using a QIAamp Viral RNA Mini kit from (Qiagen, GE).

RT- PCR was performed using a Superscript III One-Step RT-PCR kit (Invitrogen) using the following primer set: 623Fi (5'-AAATCTCTAGCAGTGGCGCCCGAACAG-3') and 2cRx (5-' TAACCCTGCGGGATGTGGTATTCC-3'). A second round of PCR was performed with 100-mer forward and reverse primers complementary to NL4-3 on either side of Gag-protease, using a HighFidelity Platinum Taq polymerase (Invitrogen). The primer set used was: GagPro-chimeric-recomb-F-Toshi (5'-GACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGG GCGGC GACTGGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGA GATGGG-3') and GagPro-chimeric-recomb-R (5'-GGCCCAATTTTTGAAATTTTTCCTTCCTTTTCCA TTTCTGTACAAATTTCTACTAATGCTTTTATTTTTTCTTCTGTCAATGGCCATTGTTT AACTTTTG-3'). For each sample 10 μg of the pNL4-3Δgag-protease was linearized by BstE II digestion and co- transfected with 5μg of purified PCR product of autologous Gag-protease from patients into 2×10^6 GXR-CEM reporter cells in 300 μl of R10 media. Electroporation was carried out at 250v, 950μF for 30 mse. The cells were left to rest for 5 minutes, and transferred to a T25 flask containing 1×10^6 GXR cells in 10 ml R10 media. Cells were then incubated at 37 °C. GFP expression was measured every 2 days using a FACSCalibur flow cytometer (BD Biosciences, San Jose, CA) as per manufactures instructions. Chimeric viruses were harvested when >15% of GXR-CEM cells where GFP+ and stored at -80 °C.

2.7.2 Viral replication capacity (VCR) assay (chapter 4)

Titration of virus stocks and replication assays were performed as previously described [18, 137] using a multiplicity of infection (MOI) of 0.003. All assays were carried out in Oxford using aliquots of the same NL4-3 viral stocks as control to normalize measurements of VRC. In brief, titration was carried out by the addition of 400 μl of supernatant to $1x10^6$ GXR cells in 1 ml R10. GFP% was measured after 48 hours to calculate the MOI. The VCR assay was then carried out based on titration by infecting $1x10^6$ GXR cell in the same way described. The mean slope of exponential growth from days 2 to 7 was calculated using the semi-log method in Excel. This was divided by the slope of growth of the wild-type NL4-3 control included in each assay to generate a normalized measure of replication capacity. Replication assays were performed at least in duplicate and results were averaged.

2.7.3 Enzyme-linked immunosorbent assay (ELISA) (chapter 4 and 5)

An HIV p24 ELISA kit (PerkinElmer, Life sciences, inc.) was utilized to monitor viral growth on Hut-R5 cells, according to the manufacturer's instructions.

Chapter 3

Impact of HLA-B*35 subtype differences on HIV-1 outcome in Mexico

3.1 Introduction

The specific focus of this study is on HLA-B alleles that are highly prevalent in Mexico and that have previously been associated with significant susceptibility to rapid HIV-1 disease progression in populations with B clade infection.

UNAIDS estimated that 170,000 (150,000-210,000) people were living with HIV-1 at the end of 2012 in Mexico, giving a prevalence rate of approximately 0.2% (0.2-0.3%), one of the lowest in the world. The National Centre for HIV/AIDS prevention and control (CENSIDA) in Mexico estimates that the epidemic is primarily concentrated among men who have sex with men (17% prevalence), male sex workers (18% prevalence) and their clients and people who inject drugs (6%). Relevant clinical data of our study cohort can be found in Table 3.1.

Clinical parameter	Median	Standard Error	Standard Deviation
VL (RNA copies/ml) *	40878 (10374-135043)	11949	360266
CD4+ T cell count (cells/µl) *	380 (151- 592)	9.587	290.6
	Gender [n(%)]		
Male	721 (78.5%)		
Female	197 (21.5%)		

*Data from 7 patients missing

Table 3.1: **Relevant clinical parameters for the study cohort of 926 Mexican adults.** Patients were recruited from the Centre for Research in Infectious Diseases, Mexico city, Mexico.

HLA-B*35 alleles are found in most populations with varying frequency, with 5-10% in most Asian and African populations to 20% in American populations [82]. Our study cohort has an approximate 30% population frequency of HLA-B*35 alleles, to our knowledge the highest frequency recorded in a population. HLA-B*35 is the most prevalent HLA-B allele in Mexico [11], making this cohort the ideal population which to study HLA-B*35-mediated disease control (Figure 3.1.1).

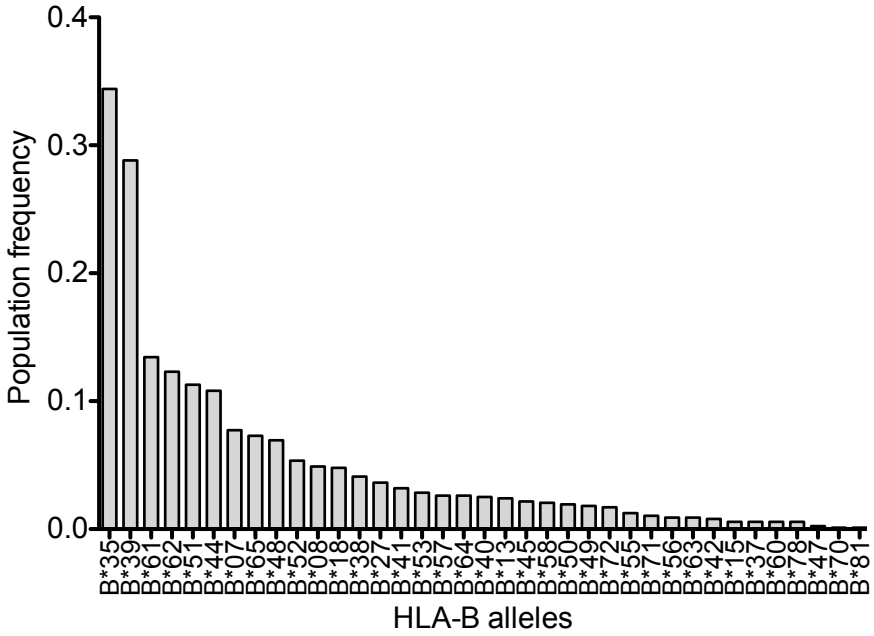


Figure 3.1.1: **Population frequency of HLA-B alleles** in a cohort of 926 HIV-1 B clade chronically infected individuals from Mexico. Two digit HLA types shown. Figure shows frequency of alleles occurring at a population frequency of 1% and over. HLA typing was carried out by the Reyes-Teran group.

B*35 alleles

Previous studies have consistently associated HLA-B*35 with rapid disease progression in the context of B clade HIV-1 infection [22, 47, 23, 49]. Differences on disease outcome between HLA-B*35 subtypes and between clades of infection have also been shown. The mechanisms involved in this observation are not well understood. However characterization of the $CD8^+$ T cell response in the Durban population showed that increasing breadth of the Gag-specific response is associated with decreasing viral loads [69]. Additionally, HLA alleles associate with lower median VL tend to be members of the Bw4 group, which act as KIR ligands and therefore incorporate the NK cell responses against the virus; B*35 alleles are members of the Bw6 group, and therefore lack interaction with KIR. It has also been shown that certain alleles, including B*35:03, bind with higher affinity to the immunoglobulin-like transcript 4 (ILT4) which is an inhibitory MHC class I receptor expressed by dendritic cells that is involved in dendritic cell dysfunction and therefore may contribute to loss of immune control [62]. Finally it has been suggested that HLA alleles associated with disease susceptibility have a peptide binding motif which allows for a significant number of self peptides to bind and hence result in a smaller TCR repertoire after negative selection by the thymus; in contrast protective alleles have a more restrictive peptide biding motif with fewer auto reactive T cells and therefore a larger TCR repertoire [86].

It is possible that the sum of some or all the features mentioned may have an additive effect on disease outcome. However in this project we analyze the role of different epitopes presented by HLA-B*35 alleles, as it has previously been shown that even one effective $CD8^+$ T cell response can mediate long term immune control of HIV-1 infection, such as KK10 presented by HLA-B*27 [56]. It has also been shown that HLA-B*35:01 association with susceptibility or control of HIV-1 infection is related to immunogenicity to a single Gag epitope (NY10-Gag 253-262) [86].

B*35 subtypes

Historically HLA-B*35 alleles have been divided into two groups based on their amino acid preference at the anchor position at the C terminus of the peptides they present. The PY

group bind proline at P2 and tyrosine (Y) at the C terminus (hence known as PY group); while members of the PX group also bind proline at P2 but do not bind tyrosine at C terminus (hence PX), instead they prefer smaller hydrophobic residues such as leucine, methionine or valine [47] (Table 3.2 and 3.3).

Allotype	Residue									
	63	94	95	97	103	109	114	116	152	156
B*3501	N	I	I	R	L	L	D	S	V	L
B*3508	-	-	-	-	-	-	-	-	-	R
B*3514	-	-	-	-	-	-	-	-	E	W
B*3516	E	-	-	S	V	-	-	-	-	-
B*3517	-	-	-	S	V	-	-	-	-	-
B*3543	-	T	L	-	V	-	-	-	E	W
B*3502	-	-	-	-	-	F	N	Y	-	-
B*3503	-	-	-	-	-	-	-	F	-	-
B*3512	-	-	-	-	V	-	N	Y	-	-

Table 3.2: **Site variation of HLA-B*35 subtypes found in the study cohort with 1% and over population frequency.** Members of the PY group are shaded in grey, the PX group is not shaded[82].

	HLA alleles	Position 2	C terminus
B*35 PY	B*35:01 B*35:08 B*35:14 B*35:16 B*35:17 B*35:43	**P** (A,V)	**Y** (F,M,L)
B*35 PX	B*35:02 B*35:03 B*35:12	**P** (A)	**M, L** (F)

Table 3.3: **Classification and binding motifs of HLA-B*35 alleles in our study cohort**. Preferred amino acids for binding at HLA anchor positions are shown in bold, other options in brackets (Marsh et al. 2000; Gao et al. 2001; Honeyborne et al. 2006; Escobar et al. 2008).

Subtypes of the PX group, such as HLA-B*35:02 and HLA-B*35:03 have been particularly associated with worse disease outcomes, as subjects with these alleles exhibit early decline in $CD4^+$ T cell counts [106] and have worse outcomes than homozygotes [63]. HLA-B*35 alleles are structurally related to HLA-B*53:01 and share similar architecture in the B pocket of the peptide binding groove, and hence prefer small hydrophobic residues at P2 [82]. HLA-B*53:01 has also been associated with susceptibility to HIV-1 disease and therefore has been included as a member of the PX group in previous studies of HLA-B*35 alleles [47]. Conversely, subjects with HLA-B*35:01 may have better outcomes, particularly in the context of C clade infection

[86].

The mechanisms underlying the difference in disease outcome between these closely related alleles are not clear, however it has been suggested that the difference in affinity for tyrosine at the C terminus of the peptide is a critical distinction between HLA-B*35 alleles that have either a beneficial effect or no effect (PY) or a negative effect (PX) on HIV-1 disease outcome. We therefore investigate whether this difference in outcome is based on the variability of HIV-1 specific peptides that can be presented [47].

B*35 micropolymorphisms: B*35:01 and B*35:08 differences

HLA polymorphism plays a central role in protective immunity, disease susceptibility, autoimmunity, and drug hypersensitivity. Previous studies of closely related HLA-B allele pairs B*57:01/03 and B*44:02/03 [19, 5, 70] have shown that small differences, including a single amino acid change at position 156 between HLA subtypes, can impact the dynamics of HLA/peptide interaction and can therefore alter disease outcome. Changes at position 156 have been shown to affect both the MHC heavy chain and the bound peptide conformation through changes in the hydrophobic and electrostatic properties of the antigen binding cleft [5]. HLA-B*35:01 and HLA-B*35:08 differ only at position 156 (L156 and R156 respectively), a position buried within the antigen binding cleft.

Although previous studies have suggested the HLA-B*35:01 and B*35:08 both have a null or positive effect in the context of B clade HIV-1 infection [47, 23], we observe a difference in disease outcome in our study cohort. We therefore investigate the possible mechanisms of HIV-1 control by these two closely related alleles.

Previous studies have looked into differences in peptide presentation by these closely related alleles and have found that in some instances both alleles are capable of presenting and mounting an immunogenic response against the same epitope but restrict different TCRα chains [93]. It has also been shown that B*35:01 and B*35:08 direct CTL responses against epitopes with different lengths [57]. Finally, both alleles may be able to present the same peptide on the cell surface, with the same degree of effectivity and availability, but only one elicits an immunogenic response [134]. These differences are thought to occur due to difference in TCR-pMHC interactions rather than due to a peptide processing or presentation problem. It is thought that the different

conformation of pMHC related to the change at position 156 leads to these different TCR interactions.

Previous studies have suggested that HLA-B*35:01 control of HIV-1 in the context of C clade infection is due to a single Gag response to the peptide known as NY10 (Gag 253-262, NPPIPVGDIY); progression in the context of B clade infection occurs as the B clade consensus sequence has an escape already in place for this epitope at D260E (NPPIPVGEIY). $CD8^+$ T cells specific to NY10-260D can cross-react and recognize NY10-260E, however no CTLs specific only to NY10-260E were found [86]. It is therefore possible that the amino acid change at position 156 may allow for $CD8^+$ T cell recognition of the NY10-260E when presented by HLA-B*35:08 while there is a blind spot when presented by HLA-B*35:01, hence the difference in disease outcome observed.

Specific aims

To further elucidate these observations we studied the difference in HIV-1 disease outcome of HLA-B*35 subtypes in a cohort of 976 individuals from Mexico with chronic HIV-1 B clade infection.

The specific aims of the present study are:

1. To compare viral set point and absolute $CD4^+$ T cell count in association with different HLA-B*35 subtypes;

2. To characterize HLA-B*35-restricted epitopes, with a particular focus on immune control effected through Gag epitopes;

3. To identify sites of escape mutation selected by HLA-B*35 subtypes, and to analyze the effects of these polymorphisms in Gag epitopes on disease outcome;

4. To identify the impact of HLA-B*35 micropolymorphisms on immune control of HIV-1 through the study of HLA-B*35:01 and HLA-B*35:08.

3.2 Methods

Study cohort

Subjects recruited from Mexico and Badalona as described in chapter 2.

ELISpot assays

IFNγ CD8$^+$ T cell responses from 46 patients carrying at least one HLA-B*35 allele were studied using 16 optimal epitopes identified for HLA-B*35 alleles (Table 3.4). The assays were carried out as described in chapter 2.

Epitope designation	Protein	C clade	B clade	HXB2 location	Published
WF9	p17 Gag	WASRELERF	---------	Gag 36-44	Goulder, et.al. 1997
NY9	p17 Gag	n/a	NSSQVSQNY	Gag 124-132	Rowland-Jones, et.al. 1995
HA9	p24 Gag	HPVHAGPIA	---------	Gag 216-224	Gudmundsdotter, et.al. 2008
NY10	p24 Gag	NPPIPVGDIY	-------E--	Gag 254-262	Rowland-Jones, et.al. 1995
TY9	RT	TVLDVGDAY	---------	Pol 262-270	Wilson, et.al. 1999; Wilkes, et.al. 1999
VY10	RT	VPLDEDFRKY	----K-----	Pol 273-282	Shiga, et.al. 1996; Sipsas, et.al. 1997
NQY9	RT	NPEIVIYQY	--D------	Pol 330-338	Shiga, et.al. 1996; Sipsas, et.al. 1997
EY10	RT	EPIAGAETFY	---V------	Pol 587-596	Shiga, et.al. 1996; Tomiyama, et.al. 2000
IY11	RT	IPAETGQETAY	----------	Pol 804-814	Mathews, et.al. 2012
K-QY10	Rev	QAVRIIKILY	KT--L--F--	Rev 14-23	Mathews, et.al. 2012
VL11	gp120	VPVWKEAKTTL	-------T---	Env 42-52	Wilkes, et.al. 1999
DL9	gp120	DPNPQEMVL	------V--	Env 78-86	Shiga, et.al. 1996
TW9	gp41	TAVPWNSSW	------A--	Env 606-614	Johnson, et.al. 1994
RY11	Nef	RPQVPLRPMTF	----------Y	Nef 70-81	Shiga, et.al. 1996
VF8	Nef	VPLRPMTF	-------Y	Nef 73-81	Culmann, et.al. 1991; Culmann-Penciolelli, et.al. 1994
YF9	Nef	YPLTFGWCF	---------	Nef 135-143	Kiepiela, et.al 2007

Table 3.4: **List of sixteen HLA-B*35-restricted epitopes in HIV-1 studied.** B and C-clade consensus sequence of each epitope are listed. n/a = not applicable (no C-clade equivalent).

Analysis of Gag sequence data

574 full-length Gag population sequences from Mexico were analyzed according to the lineage correction method as described in chapter 2. Analysis was carried out by Jonathan Carlson [21].

Statistics and terminology

Box plots show the median and interquartile percentiles. Whiskres show minimum and maximum values.

Throughout this chapter the term HLA-B*35 is used to represent B*35:01, B*35:02, B*35:03, B*35:08, B*35:12, B*35:14, B*35:16, B*35:17 and B*35:43 alleles collectively.

3.3 Results

3.3.1 Comparison on markers of disease progression in association with different HLA-B*35 subtypes in Mexico

Seventeen HLA-B*35 subtypes were found in the study cohort of which B*35:01/02/03//08/12/14/16/17/43 are found at ≥1% population frequency (Figure 3.3.1, C).

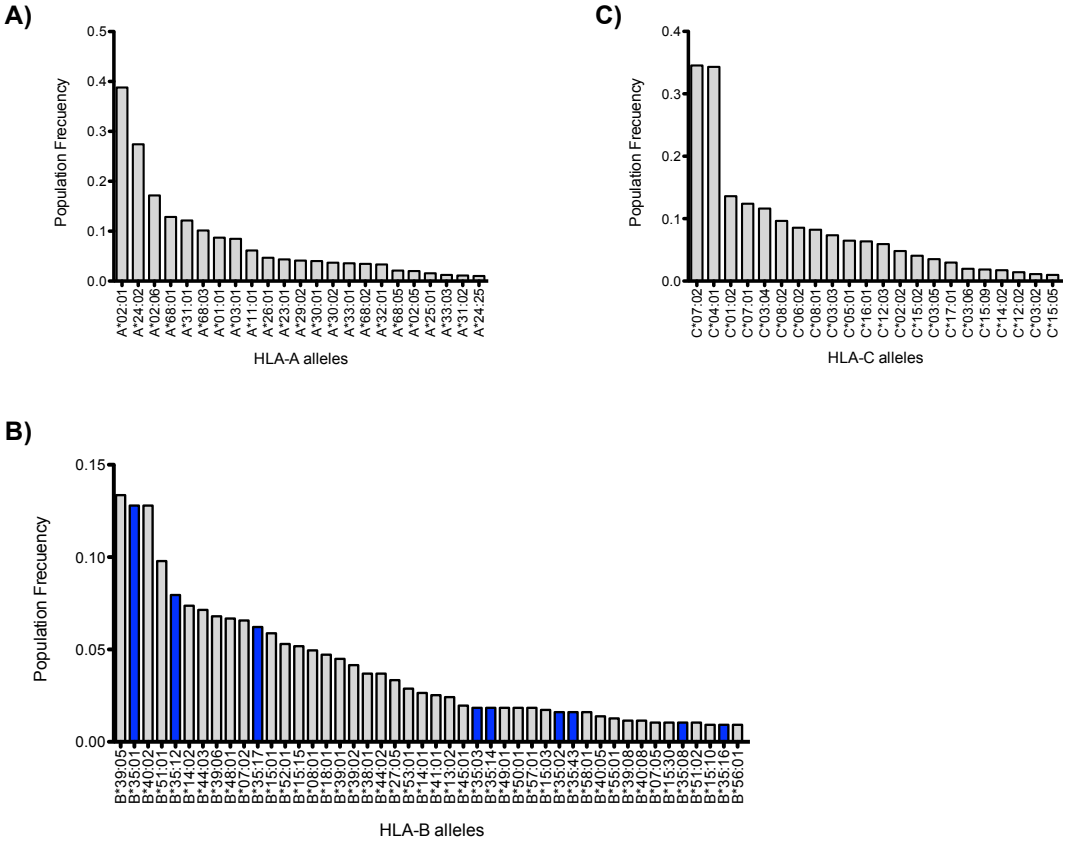


Figure 3.3.1: **Population frequency for class I HLA alleles in a cohort of 926 HIV-1 B clade chronically infected individuals from Mexico.** Figure shows frequency of alleles occurring at a population frequency of ≥1%. HLA-A alleles are shown in A), HLA- B alleles in B) and HLA-C alleles in C). HLA-B*35 subtypes are highlighted in blue. HLA typing was carried out by the Teran group in Mexico.

We compared the median VL and median $CD4^+$ T cell counts of all HLA-B alleles with ≥1% population frequency. Consistent with previous studies of B clade infected Caucasians, HLA-B*57 and HLA-B*27 were the most protective alleles (p=<0.0001, q=0.002 and p=0.0021, q=0.029 respectively) (Fig 3.3.2). Diverse ranking of HLA-B*35 alleles according to median VL

was observed. HLA-B*35:08 was 17th of 46 HLA-B alleles, while B*35:16 was 46th. HLA-B*35:01, a PY allele not universally considered a risk allele, was significantly associated with higher VL (p= 0.0154, q=0.065).

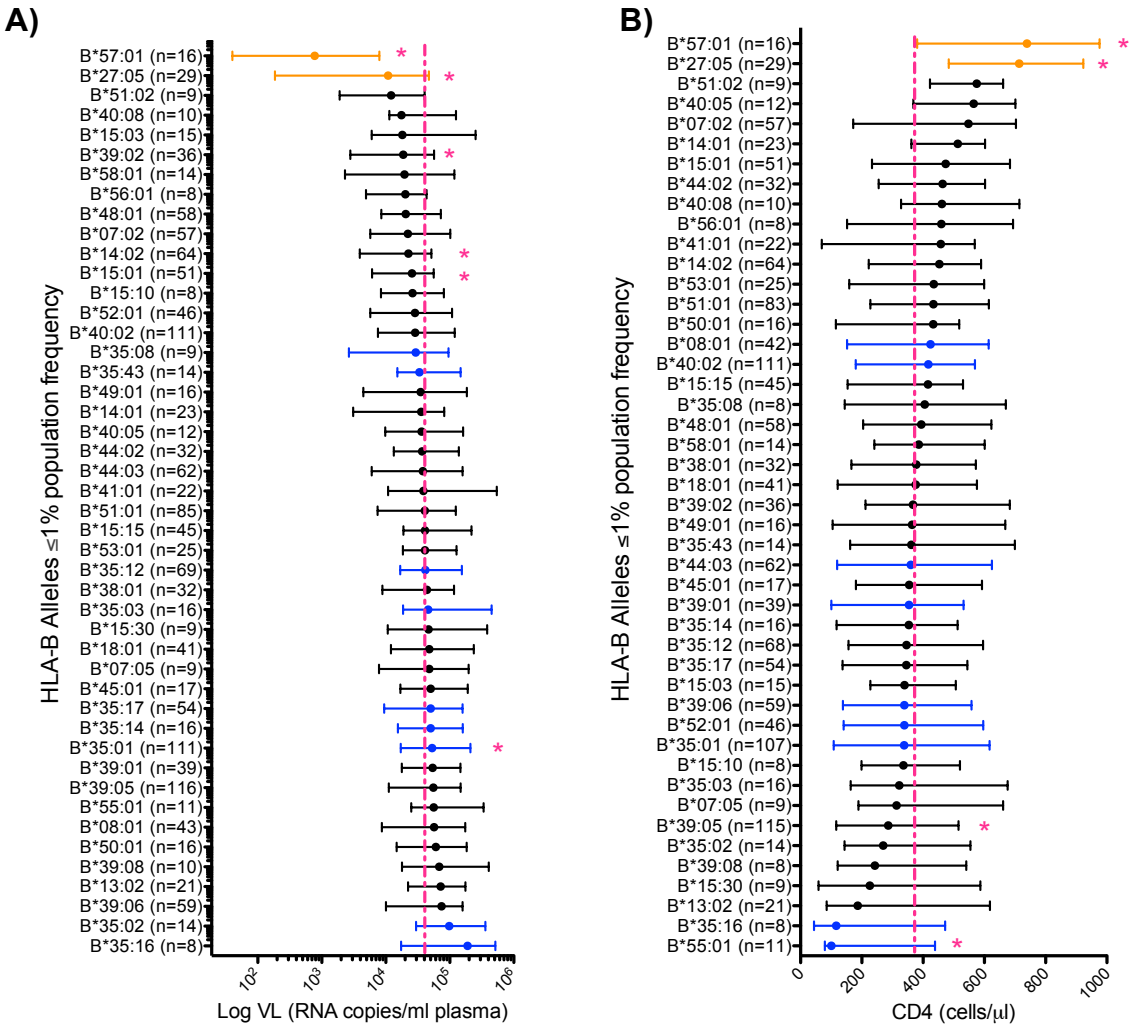


Figure 3.3.2: **Association of HLA-B alleles occurring at a population frequency of ≥1% (n=926) with viral load (A) and absolute CD4$^+$ T cell count (B).** HLA-B*35 subtypes are highlighted in blue. Associations with P value of <0.05 (as measured by a Mann-Whitney U test) and q of <0.2 are marked by * and those retaining significance after Bonferroni correction for multiple comparisons are highlighted in orange. X-axis shows the association to viral load or CD4$^+$ T cell count expressed as median with interquartile ranges.

When HLA-B*35 alleles were pooled, in accordance with previous studies of B clade cohorts [106, 63], a significant overall association with disease progression was found. Subjects with HLA-B*35 alleles exhibit a significantly higher median VL and significantly lower median

absolute CD4$^+$ T cells count (p=0.002 and p=0.01 respectively) (Figure 3.3.3).

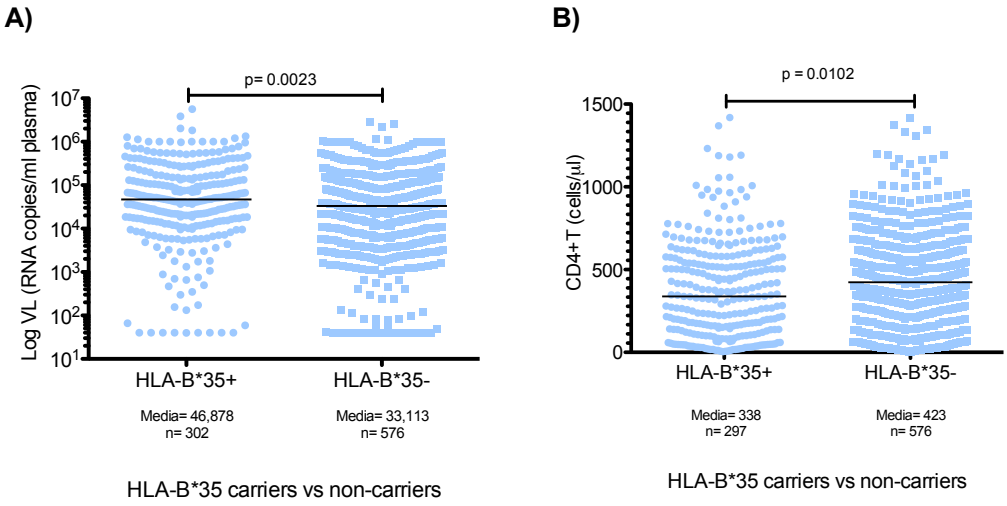


Figure 3.3.3: **Comparison of markers of disease progression between individuals carrying B*35 alleles and individuals without B*35 alleles.** A) Association with viral load and B) absolute CD4$^+$T cell count. P values generated by a Mann-Whitney U test. Horizontal lines indicate median.

To investigate whether the advantage of the HLA-B*35 PY group over the HLA-B*35 PX group of alleles is true in this study cohort, we compared markers of disease progression in patients in each group of alleles. In contrast with previous work [22, 47], no significant impact on median VL or median CD4$^+$ T cell count was observed between the HLA- B*35 PY versus PX groups (p=0.93 and p=0.58 respectively) (Figure 3.3.4, A-B), even when B*5301 was included in the analysis (Figure 3.3.4, C-D).

3.3.2 Characterization of HLA-B*35-restricted epitope responses, with a particular focus on immune control effected through Gag epitopes.

Detailed comparative studies of HLA-B*35 subtypes and the CD8$^+$ T cell responses restricted by these alleles have not been done; however, a small study was previously undertaken which showed an inverse correlation between viral load and the magnitude of the Gag-specific CD8$^+$ T cell response in HLA-B*35:01-positive subjects compared to HLA-B*35:02/03-positive subjects [63]. Therefore we aimed to compare the pattern of CD8$^+$ T cell responses to previously

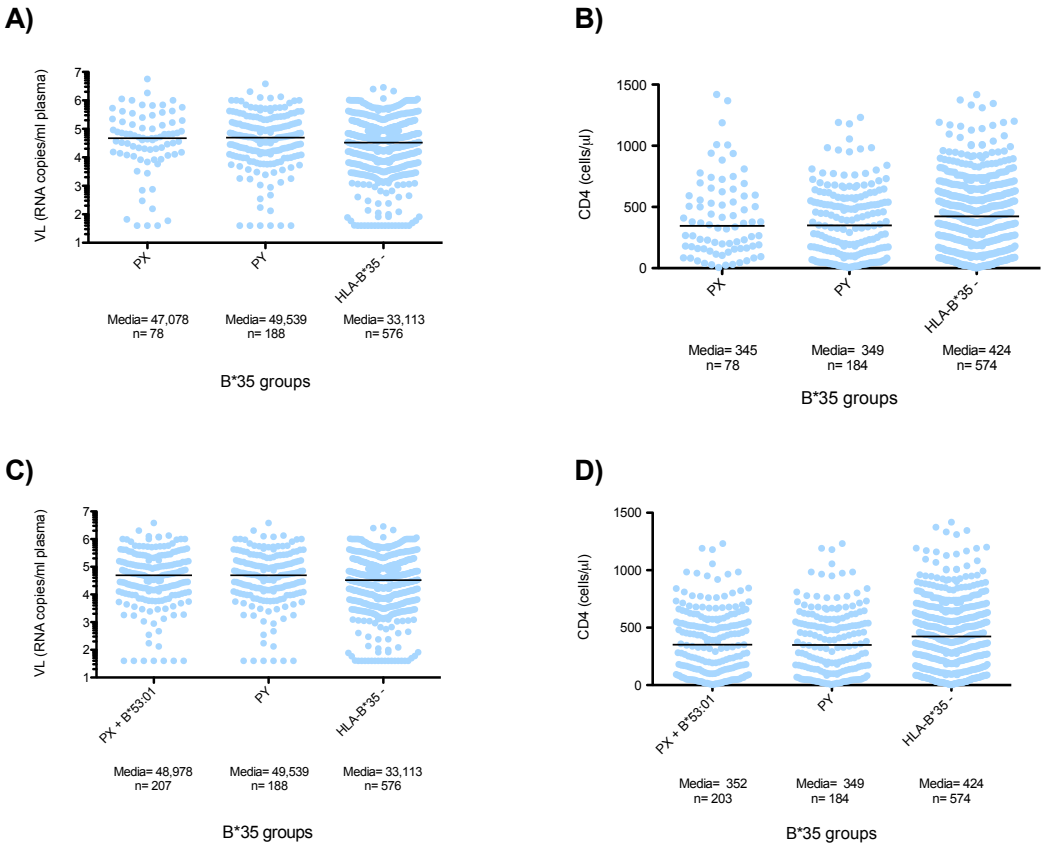


Figure 3.3.4: **Comparison on the effect of expression of HLA-B*35 PY alleles, HLA-B*35 PX alleles and non-B*35 carriers**. A) Association with viral load and B) absolute CD4+ T cell count. C) and D) show associations were B*53:01 is included in the B*35 PX group. Analysis was carried out by 1way ANOVA (Krustal-Wallis- non parametric) analysis with Dunn's Multiple comparison test. No significant difference between groups was found.

described HLA-B*35 epitopes between the subtypes found in the study cohort. However, limited availability of PBMCs did not allow for this kind of comparison between subtypes to take place; instead we looked at differences in VL based on the specificity of $CD8^+$ T cell response in individuals carrying HLA-B*35 alleles. The association of individual B*35 restricted epitopes previously identified in Gag, Pol, Env and the accessory and regulatory proteins Rev, Tat, Vif, Vpr, Vpu and Nef ('Acc/Reg') were tested by IFN-γ ELISpot assays and are shown in Table 3.4. Due to limited cell numbers we were unable to test the full length HIV-1 proteome, and therfore the effect of additional novel responses driven by HLA-B*35 alleles or indeed responses restricted by other HLAs were not measured.

We carried out assays using PBMCs from 46 individuals with HLA-B*35 alleles for which there were sufficient cells available. Data shows a trend towards an association between Gag responses and lower VL, although this is not significant. Response to the Gag-NY10 epitope, shows a trend towards lower viraemia (p=0.06), but again significance is not achieved, probably due to small numbers. Significant association between response to the Int-IY11 epitope and higher viremia was found (p=0.04) (Figure 3.3.5).

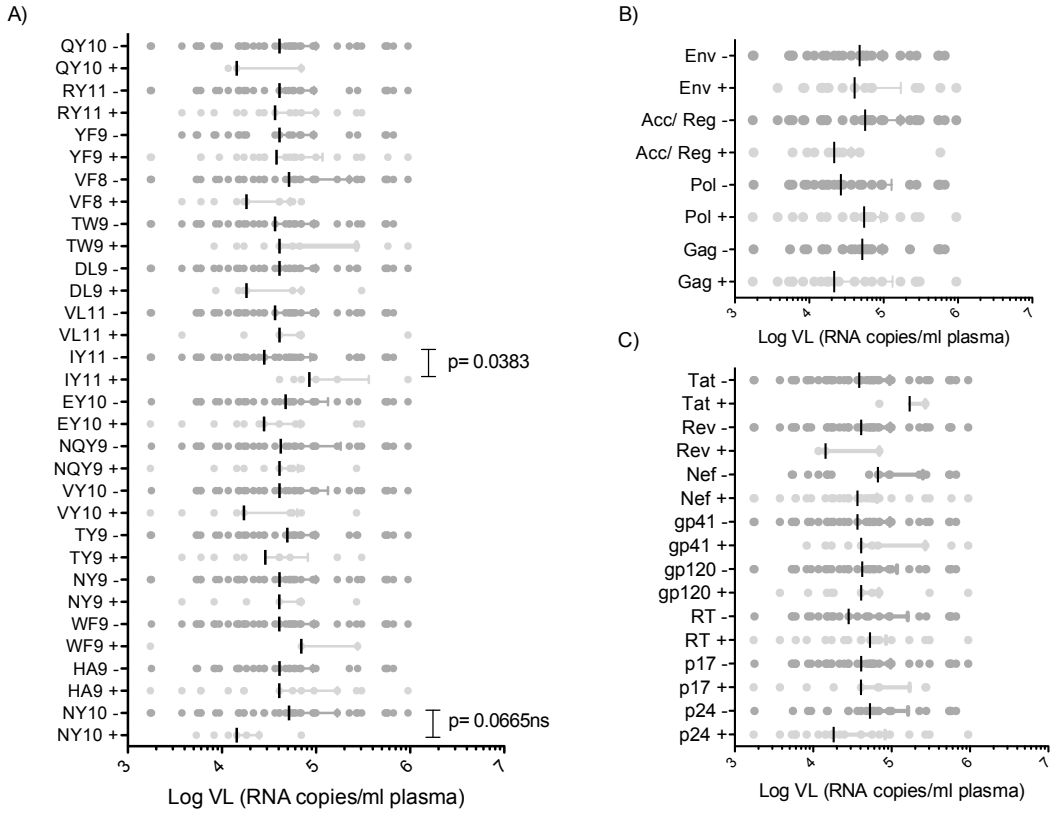


Figure 3.3.5: **Association on viral load of individuals with HLA-B*35 alleles that either produce or do not produce a response to the peptides tested.** 46 subjects carrying at least one HLA-B*35 allele were analyzed, only previosly decrided epitopes restricted by HLA-B*35 alleles were tested. A) shows association of Individual epitopes identified in Gag, Pol, Env and the accessory and regulatory proteins Rev, Tat, Vif, Vpr, Vpu and Nef ('Acc/Reg'), and (B-C) of individual proteins. Horizontal lines indicate median viral loads. P values generated by Mann-Whitney U test.

When looking at the breadth of protein-specific $CD8^+$ T cell responses, VL was significantly lower only when there was two Gag-specific $CD8^+$ T cell responses compared to non-Gag specific responses (p=0.0499) (Figure 3.3.6).

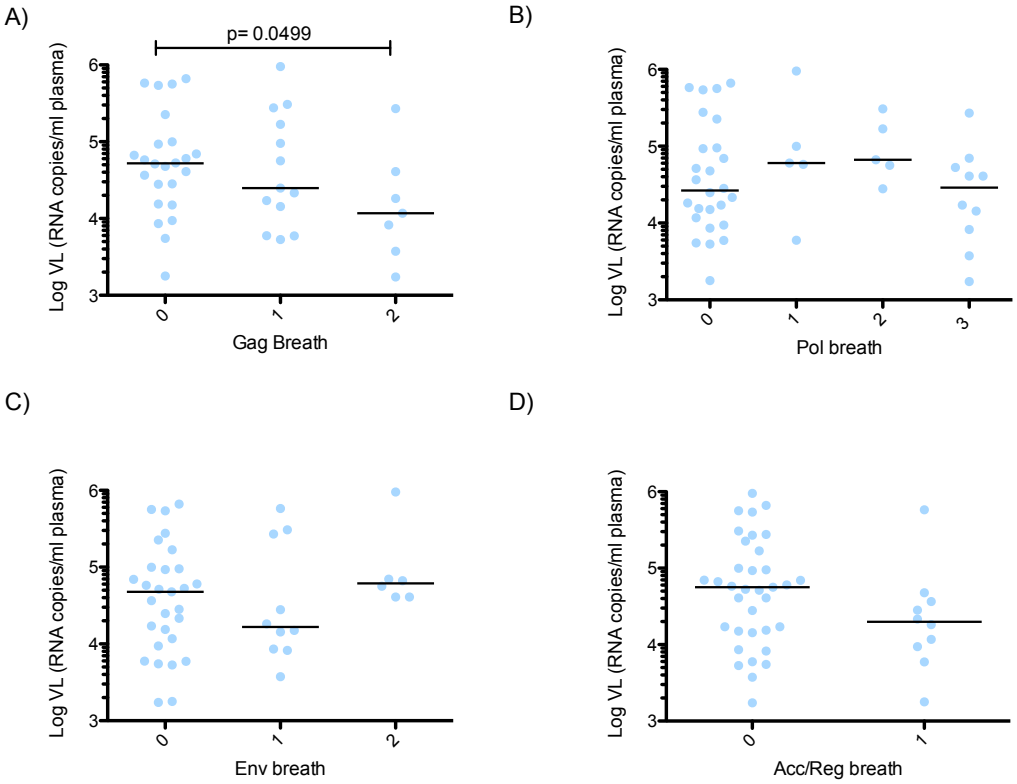


Figure 3.3.6: **Breadth of protein-specific CD8$^+$T-cell responses in relation to viral load**. (A-D) Viral load of individuals in terms of number of Gag-, Env-, Pol-, and Acc/Reg- specific responses detected in 46 patients carrying at list one HLA-B*35 allele. Horizontal lines indicate median viral load. P values generated by a Mann-Whitney U test.

3.3.3 Identification of sites of escape mutation selected by HLA-B*35 subtypes, and analysis of the effects of these polymorphisms in Gag epitopes on disease outcome.

Previous studies have shown that the number and protein location (Gag vs non-Gag) of HLA-associated HIV-1 polymorphisms is linked to HIV-1 disease outcome [88]. We carried out a polymorphism analysis to find mutations driven by HLA-B*35 alleles in our study cohort. Phylogenetic analysis of the Gag sequences studied can be seen in Figure 3.3.7. Polymorphisms in Gag associated with HLA-B alleles are shown in Table 3.5. However no statistically significant associations between HLA-B*35 alleles and HIV-1 polymorphisms were found in this cohort. As previously described, the presence of HLA-B associated polymorphisms in Gag was associated

Protein	HXB2 position	Target variant	HLA	Polymorphism location	P value	q value	Reversion infered
p17 Gag	26	K	B*1501	W E K I R L R P G G **K** K K Y K L K H I V W	2.35451E-09	2.93113E-05	R
	26	R	B*1501	W E K I R L R P G G **K** K K Y K L K H I V W	8.2178E-15	2.55759E-10	
	28	R	B*1302	K I R L R P G G K K **K** K L K H I V W A S	3.15E-05	0.084861358	
	35	V	B*4801	G K K K Y K L K H I **V** W A S R E L E R F A	6.90E-05	0.121224715	R
	54	S	B*15	F A V N P G L L E T **S** E G C R Q I L G Q L	0.000111543	0.151194772	
	81	T	B*1503	G S E E L R S L Y N **T** V A T L Y C V H Q R	0.000103267	0.146086907	R
	104	V	B*4901	V K D T K E A L E K **I** E E E Q N K S K K K	4.86E-05	0.092774522	
	116	Q	B*0801	E E Q N K S K K K A **Q** Q A A A D T G H S N	1.20496E-05	0.030001201	R
p24 Gag	147	I	B*13	V Q N I Q G Q M V H Q A **I** S P R T L N A W V K	4.23382E-06	0.014640785	R
	147	L	B*13	V Q N L Q G Q M V H Q A **I** S P R T L N A W V K	8.72E-06	0.023555195	
	215	M	B*5101	I N E E A A E W D R **L** H P V H A G P I A P	7.81E-05	0.125671215	
	242	T	B*57	R G S D I A G T T S **T** L Q E Q I G W M T N	2.18474E-10	3.39973E-06	R
	242	N	B*57	R G S D I A G T T S **T** L Q E Q I G W M T N	7.41E-11	1.54E-06	
	242	T	B*58	R G S D I A G T T S **T** L Q E Q I G W M T N	2.80E-06	0.010248888	R
	242	N	B*58	R G S D I A G T T S **T** L Q E Q I G W M T N	7.26E-06	0.020535709	
	264	R	B*27	P P I P V G E I Y K **R** W I I L G L N K I V	6.69327E-05	0.134189927	R
	280	T	B*52	L N K I V R M Y S P **T** S I L D I R Q G P K	3.07746E-07	0.002128402	R
	280	S	B*52	L N K I V R M Y S P **T** S I L D I R Q G P K	5.67E-08	0.000441009	
	280	T	B*15(01)	G L N K I V R M Y S P **T** S I L D I R Q G P K	9.79189E-05	0.145118166	
	302	K	B*1401	P F R D Y V D R F Y **K** T L R A E Q A S Q E	0.000114164	0.151194772	R
	302	R	B*1401	P F R D Y V D R F Y **K** T L R A E Q A S Q E	0.000114164	0.151194772	
	357	G	B*0702	M T A C Q V G G P **G** H K A R V L A E A M	9.08225E-06	0.023555195	
	357	S	B*0702	M T A C Q G V G G P **G** H K A R V L A E A M	1.35E-06	0.005608834	R
p2 Gag	374	N	B*27	A E A M S Q V T N S **A** T I M M Q R G N F R	1.38E-05	0.031705689	
p7 Gag	390	I	B*2705	R G N F R N Q R K T **V** K C F N C G K E G H	0.000141988	0.180368657	R
	398	E	B*4002	K I V K C F N C G K **E** G H I A K N C R A P	5.96857E-07	0.003377396	R

Table 3.5: **Sites of sequence polymorphism in HIV-1 Gag associated with host HLA class I alleles ($p<0.05$ and $q<0.2$)** 576 Gag sequences analyzed. Sites of polymorphism are marked in black, ten amino acids shown at either side according to cohort consensus. Epitopes published restricted by the defined allele are highlighted in grey. Sites where reversion is inferred are indicated by R. Analysis was carried out according to the lineage correlation method by Jonathan Carlson.

with a lower VL (p=0.004) (Figure 3.3.8, (A)). Association remained between lower median VL and the number of reverting HLA-B associated polymorphisms in Gag (p=0.002) (Figure 3.3.8, B) but not with non-reverting HLA-B associated polymorphisms (p=0.0886) (Figure 3.3.8, C).

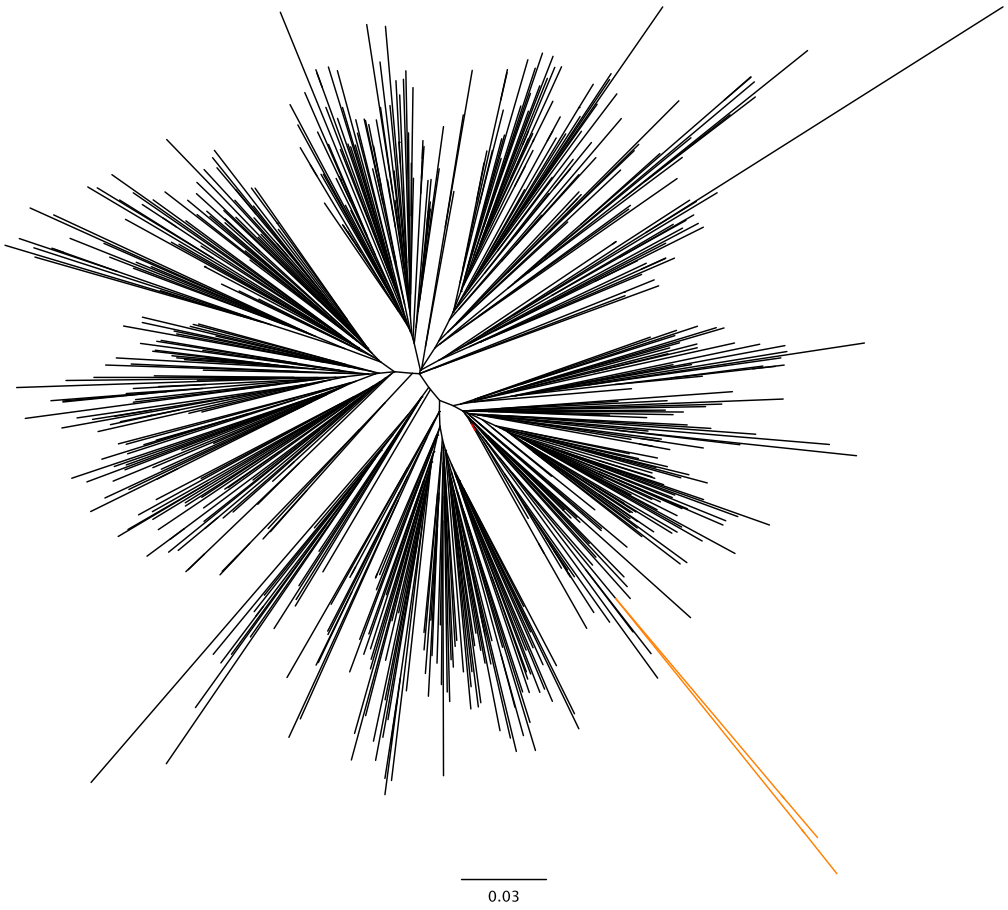


Figure 3.3.7: **Maximum likelihood phylogenetic tree of the full Gag** region of HIV-1 nucleotide sequences from 574 HIV-1 infected subjects from Mexico (black), B clade consensus (red), C clade and A clade consensus (orange). Consensus reference sequences were downloaded from Los Alamos sequence database. The tree was created using GARLI software and edited in Figtree v1.2.1

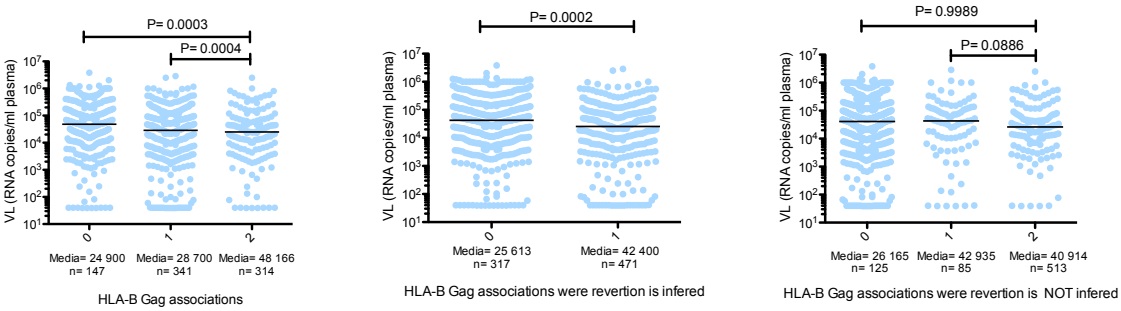


Figure 3.3.8: **Association of individuals carrying alleles that drive a number of HIV-1 polymorphisms in Gag with VL.** A) shows associations of all alleles collectively, B) Associations of alleles were reversion is inferred and C) association of alleles were reversion is not inferred. P values generated by a Mann-Whitney U test. Horizontal lines indicate median.

As the initial analysis did not identify any Gag polymorphisms associated with HLA-B*35 alleles, we proceeded to establish whether potential selection might be hidden by the large number of data comparisons in the analysis. To do this we reduced the data by looking only at the sequence of previously determined HLA-B*35 epitopes (Table 3.4). That is with q values computed to only consider sites in optimal epitopes, all alleles that are in the same supertype as at least one allele that is published as restricting that epitope were included in the analysis. Again no associations were found with $q<0.2$.

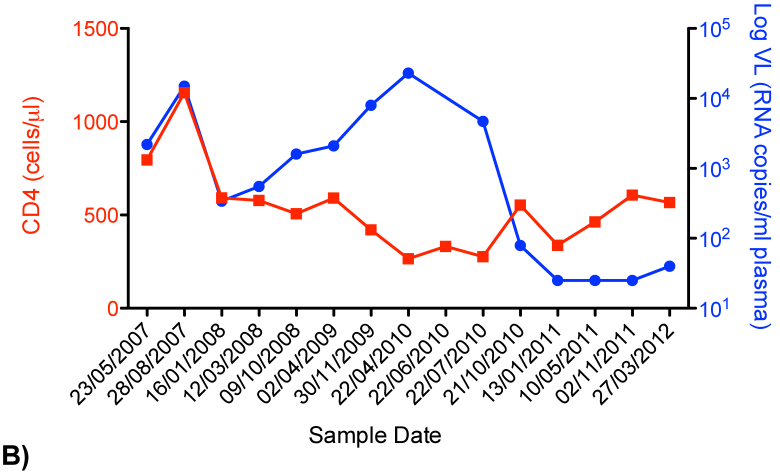
3.3.4 Impact of HLA-B*35 micropolymorphisms on immune control of HIV-1: B*35:01 vs B*35:08

HLA-B*35:01 and HLA-B*35:08 are both members of the PY group and are thought to present the same repertoire of HIV-1 peptides. In this cohort however HLA-B*35:08 appears to have a null effect on markers of disease progression while HLA-B*35:01 is associated with progression. The Gag-NY10 (Gag 253-262) $CD8^+$ T cell response has previously been linked with HLA-B*35 control of HIV-1 infection [86]. We therefore hypothesized that although B*35:01-NY10-260E (B clade consensus) is not immunogenic [86]; B*35:08-NY10-260E is. To test our hypotheses, and because no PBMCs were available for individuals carrying the alleles of interest in our study cohort, we recruited three HLA-B*35:08 positive patients from Barcelona which were B clade infected. The $CD8^+$ T cell responses from these patients was analyzed as previously described.

Although conservative, we found one patient (ID 4888607; HLA-A*0101/1101, B*35:08/52:01,

C*04:01/12:02) who responded to the NY10-260E peptide and did not have a response to NY10-260D. Longitudinal ELISpot data showed that as the NY10 response was lost, loss of control of viraemia was also happening. The patient (ID 4888607) started ART treatment on the 22nd of April 2010 (Figure 3.3.9) .

A)



B)

EPITOPE	Date of sample 30/11/2009	21/10/2010	13/01/2011	02/11/2011
	Spot forming cells/ million PBMCs			
B35-NY10-E	64	17	0	0
B35-NY10-E	80	25	0	0
B35-NY10-E	64	100	0	0
B35-NY10-D	0	0	0	0
B35-NY10-D	0	0	0	0
B35-NY10-D	0	0	0	0
B35-HA9	0	0	0	0
B35-WF9	72	210	116	65
B35-NY9	0	0	0	0
B35-TY9	104	50	0	0
B35-VY10	56	59	0	0
B35-NQY9	248	302	317	145
B35-EY10	272	269	172	115
B35-VL11	0	0	72	105
B35-DL9	264	185	0	85
B35-TW9	0	0	0	0
B35-VY8	0	42	0	0
B35-RY11	0	1445	1038	425
B35-EW10	0	0	0	0

Figure 3.3.9: **Clinical and ELISpot data from patient 4888670.** A) shows viral load and absolute $CD4^+$ T cell counts displayed overtime. B) Shows IFN-γ Elispot responses to previously described B*35 epitopes over time.

We sequenced Gag from plasma RNA from later dates, and found E260D variant, which may indicate escape and explain the loss of NY10 response in this individual. A second HLA-B*35:08

patient, for which DNA was available, was sequenced from pro-viral DNA and from plasma RNA and we found that the pro-viral sequence had the consensus E260, while the plasma had the E260D variant. Which indicates possible escape caused by B*35:08 immune pressure.

We followed by comparing the percentage of individuals with B*35:08 carrying the E260D change (E260D in 30%) and the percentage of individuals without HLA-B*35:08 alleles carrying the change (E260D in10%), although no significance was achieved (p=0.077) a trend toward accumulation of 260D was observed (Figure 3.3.10) .

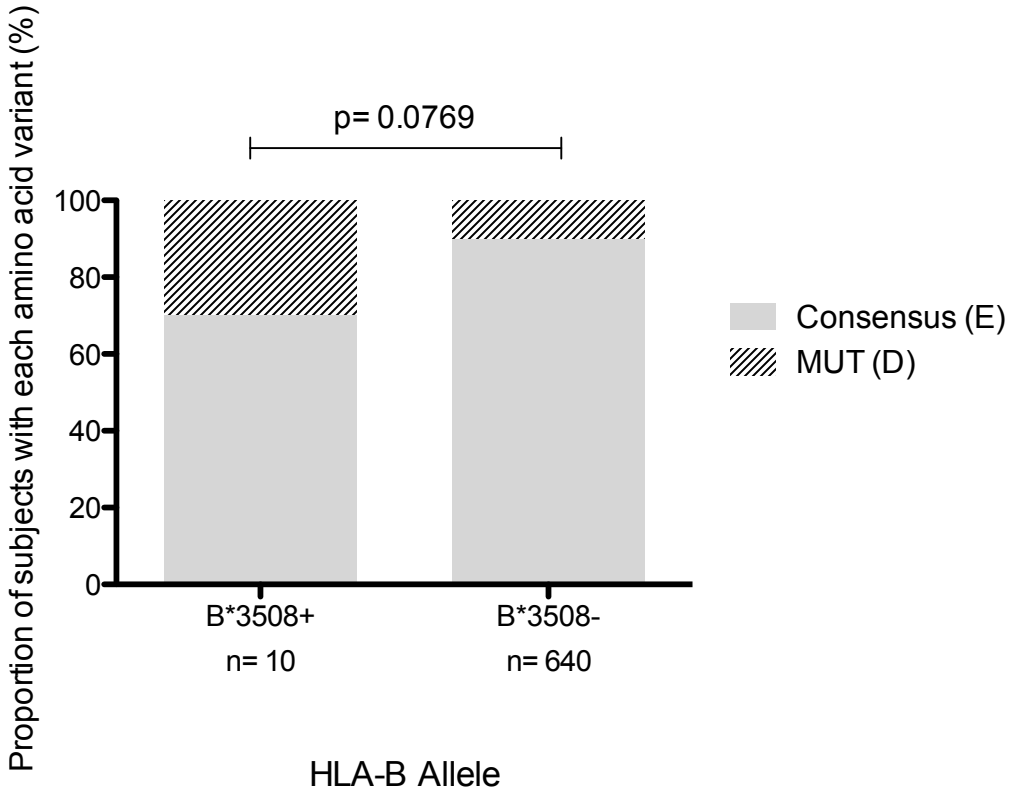


Figure 3.3.10: **Relationship between B*35:08 and accumulation of E260D.** Pooled data of subjects from the Mexican cohort and the B*35:08 subjects from Barcelona were analyzed. P value was generated by Fisher's exact test.

To confirm $CD8^+$ T cells specific for NY10-260E, we performed tetramer staining of PBMCs from subject 4888607 after expansion for 5 days using 10 μg/ml of NY10-260D or NY10-260E peptides in H10. Cells were incubated with HLA-B*35:08-NPPIPVGDIY and HLA-B*35:08-NPPIPVGEIY tetramers. We found cells specific for the NY10-260E peptide which have not

been found in previous studies of HLA-B*35:01 individuals. A small number of $CD8^+$ T cells specific for the NY10-260D were also observed (Figure 3.3.11).

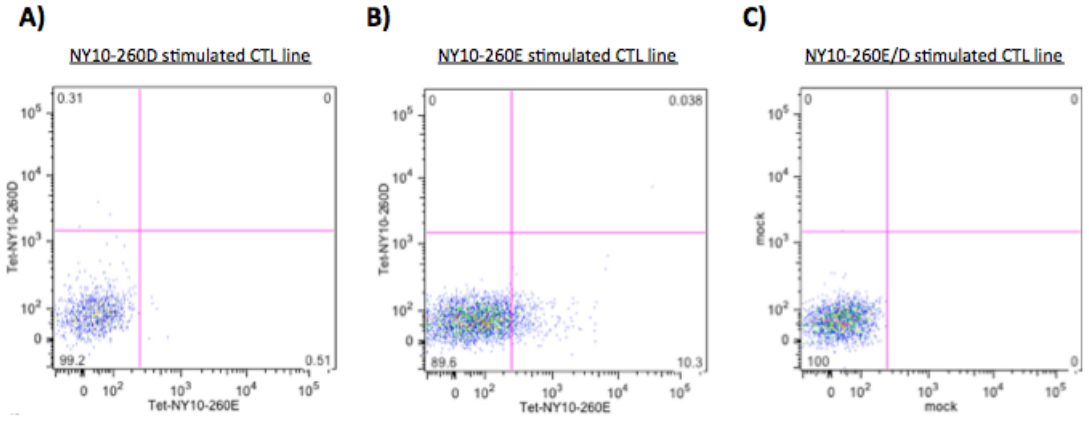
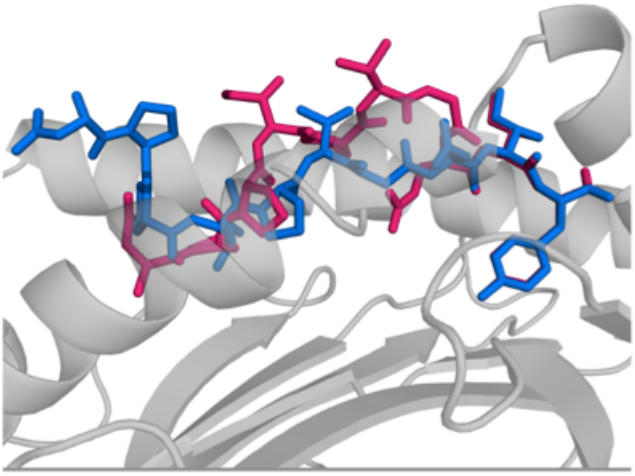


Figure 3.3.11: **CD8+ T cell recognition of B*3508-NY10- 260E/D**.In vitro-expanded PBMCs from subject 4888760 using NY10-260D (A) and NY10-260E (B) peptides and stained with dual HLA-B*3508 tetramers (260D/260E) gated on CD8+T cells (dot plots) and expressed as CD3+/Tet+ positive cells controlled by HLA-B*4201 mismatch tetramers (C).

Finally, to elucidate differences in peptide/MHC structures of the NY10 epitope, our collaborators Dr David Cole and Dr John Miles from the Department of Medical Biochemistry and Immunology, University of Cardiff, carried out crystal structural studies. The crystal structures are yet to be refined, however the preliminary data show that HLA-B*35:08-NY10-260E form a normal pMHC conformation, similar to that of HLA-B*35:01-NY10-260D (Figure 3.3.12 red). The negatively charged E at P8 appears to interact with the positively charged 156 R. On the other hand the HLA-B*35:08-NY10-260D peptide adopts a novel pMHC conformation not seen before in class I presentation. The P1 and P2 of the peptide extends outside the N-terminus binding grove (Figure 3.3.12, blue).

A)



B)

Sequence	Half life (h)
B*35:08/NPPIPVGEIY	0,3
B*35:08/NPPIPVGEIY	0,2
Positive control	11,8

Figure 3.3.12: **B*3508-NY10 260E/D crystal structure and stability assay.** A) Shows crystal structure, HLA-B*3508 is shown in grey, the NY10-260E peptide (NPPIPVGEIY) is shown in red sticks. Position P8E appears to be looking down at the MHC binding grove. NY10-260D (NPPIPVGDIY) is shown in blue. B) Shows B*35:08-NY10 E/D stability. The pMHC structure and stability assays were determined by the Sewell Group.

3.4 Discussion

The data presented in this chapter support the previous suggestion that the HLA-B*35 effect on HIV-1 disease outcome differs between subtypes. However, these data challenge previous studies where HLA-B*35 subtypes have been divided into the PY and PX groups to describe the alleles that have a null effect on disease progression, from those that have a negative effect. The data instead suggest that the HLA-B*35-mediated HIV-1 control may be associated with an individual Gag response (NY10, Gag 253-262), and highlights the importance of micropolymorphisms in determining disease outcome.

PY and PX classification

The Mexican cohort makes an interesting data set as HLA-B*35 is the most prevalent B allele in the Mexican population. Nine HLA-B*35 subtypes are found at $\geq$1% population frequency in the study cohort, providing us with the ideal setting to study the mechanisms involved in the

association to disease progression inferred by these alleles.

Previous studies of HLA-B*35 alleles have been done in populations where mainly three subtypes exist. These studies showed that HLA-B*35:02 and HLA-B*35:03 had a more rapid disease progression than HLA-B*35:01, and concluded that the HLA-B*35 subtypes can be divided into two groups, the PX group associated with more rapid disease progression that members of the PY group [22, 47, 63].

This observation is not true in our study cohort. We do not observe a difference in markers of disease progression between these groups (Figure 3.3.4).

Additionally HLA-B*35:01 and HLA-B*35:08 are both members of PY group. In our study HLA-B*35:08 has a null effect on markers of disease progression while HLA-B*35:01 is associated with progression. Our first assumption was that as both alleles have been grouped together in previous studies, B*35:08 might have been masking the negative effect of HLA-B*35:01, giving a possible explanation to the differences in our findings. However this is unlikely as the frequency of HLA-B*35:08 was even lower in the study carried out by Gao et.al 2001 [47]. The difference on population frequency of the studied alleles in each cohort (B*35:01- 4.3% vs 12.79% and B*35:08- 0.4% vs 1%) may explain the observed differences. Additionally, we carried out a cross-sectional study while previous studies were longitudinal [47].

Never the less, our data highlights the importance of micropolymorphisms in disease outcome. It also highlights that alleles may operate differently in different populations. Therefore caution should be exercised in making generalizations about the effect of alleles.

Immune control effected through Gag epitopes

Several studies have shown that targeting of Gag and the breadth of Gag-specific responses are important in HIV-1 disease control. Although HLA-B*35 alleles are associated either with a null effect or a negative effect on disease outcome in the study cohort, we analyzed whether the Gag effect remained. We compared the specificity of responses made by individuals with HLA-B*35 alleles by IFNγ ELISpot assays in association with VL. We found a trend towards association of responses to Gag with lower VL (Figure 3.3.5, B) although not significant. The NY10 epitope, previously associated with HIV-1 control by B*35:01 alleles, also shows a trend towards lower VL without been significant (Figure 3.3.5, A). It is possible that the constant

encounter of the virus with HLA-B*35 alleles, as they are highly frequent in the population, has led to viral adaptation and hence explaining why this 'Gag effect' is not observed. Additionally the consistent association of HLA-B*35 alleles with disease progression, suggest a predisposition to the absence of Gag responses. It must be noted however, that we only analyzed responses to previously described HLA-B*35 epitopes due to limited cell numbers. Novel HLA-B*35 restricted responces or responses restricted by other HLA alleles were not studied. A more comprehensive analysis on responses of the full length HIV-1 proteome is required to obtain conclusive results.

When looking at the breadth of protein-specific $CD8^+$ T cell responses, VL was significantly lower when there were two Gag-specific $CD8^+$ T cell responses compared to non Gag specific responses (p=0.0499, Fig 3.3.6), supporting the 'Gag hypothesis', where even the addition of a single Gag response appears to significantly alter the impact on HIV-1 disease outcome.

Surprisingly, associations of HLA-B*35 alleles with HIV-1 polymorphisms were not found in this cohort (Table 3.5). This suggest that the HLA-B*35 may not create enough immune pressure to drive viral escape, or it may be possible that the viral strain found in Mexico is already adapted to this alleles due to a founders effect, or potentially because of adaptation over time in proportion to the phenotypic frequency of HLA-B*35 alleles , it is a possibility that B*35 alleles were protective early in the epidemic.

Nevertheless, consistent with previous studies, we observe an association of the number of HLA-B-associated polymorphisms in Gag with lower VL (p=0.0003) (Figure 3.3.8). Association was found between lower median VL and the number of reverting HLA-B-associated polymorphisms in Gag (p=0.0002) (Fig 3.3.8, B) but not with non-reverting HLA-B polymorphisms (p=0.0885) (Figure 3.3.8, C). These data support the hypothesis that the escape mutations that reduce viral fitness are those most strongly associated with control of viraemia.

B*35 156 micropolymorphism effect

B*35 alleles may share the repertoire of peptides that they are able to present, however different patterns of immunodominance may occur and can explain difference in markers of disease progression. Additionally differences in the pMHC structure driven by these micropolymorphisms can affect TCR repertoire and $CD8^+$ T cell priming [58, 134, 93].

HLA-B*35:01 immunogenicity to the Gag NY10-260D peptide has been shown to be associated with control of HIV-1 infection in C clade [86]. However the change at D260E is selected and the response is lost. B clade HIV-1 already has the E at 260 as consensus and it may explain why individuals with C clade infection fare better than when infected with B clade virus in which the escape variant is already in place. Because individuals with HLA-B*35:08 exhibit a better disease outcome in our cohort, we hypothesize that the HLA-B*35:08-NY10-260E is immunogenic, unlike HLA-B*35:01-NY10-260E.

HLA-B*35:01 and HLA-B*35:08 differ only by one amino acid at residue 156. The arginine at 156 in HLA-B*35:08 leads to a more positively charged D pocket and hence charge complementarity for negatively charged amino acids; for example it has been shown that R usually interacts with glutamate on the peptides presented by building a salt bridge which stabilizes the pMHC structure [134, 93]. Therefore an interaction between the 260E and 156R is possible and the change in pMHC conformation may lead to altered immunogenicity.

We show that a response is indeed possible to NY10-260E by an individual carrying HLA-B*35:08 (patient ID 4888670). Longitudinal ELISpot data shows that as the patient progresses, the NY10-260E response is lost (Figure 3.3.9). This may indicate that indeed the NY10 response is important in HIV-1 control. Further more Gag sequence from proviral DNA suggest that a HLA-B*35:08 patient may have been infected with the WT viral strain E260, and HLA-B*35:08 pressure may have driven escape leading to the E260D found in the plasma sequence of the patient. Although the percentage of E260D in HLA-B*35:08 individuals is not significant, a trend can be observed ($p=0.077$, Figure 3.3.10).

Additionally we confirm the $CD8^+$ T cell response to NY10-260E by tetramer staining. The opposite pattern of immunogenisity found in the HLA-B*35:01 study of C clade infection, were $CD8^+$ T cells specific for NY10-D but not for NY10-E were found, is shown for HLA-B*35:08. Our data show $CD8^+$ T cells specific for NY10-E and not NY10-D when presented by HLA-B*35:08 (Figure 3.3.11). Therefore suggesting that NY10 immunogenicity is associated with either HLA-B*35:01 or HLA-B*35:08 depending on the residue at P8.

Preliminary datum on the crystal structure of the HLA-B*35:08-NY10-260E shows that the pMHC conformation is effective. Binding of the peptide is similar to that of previously described models of the HLA-B*35:01-NY10-260D that has been shown to be immunogenic [86]. In the

context of HLA-B*35:08, both NY10-260D and NY10-260E peptides overlap at PC and P9 when bound. However the pMHC structures divert dramatically from this point onwards and NY10-260E adopts a normal conformation with P2 in the B pocket, P3 in the D pocket and a middle bulged region (Figure 3.3.12, A red). The side chain of the P8-E appears to be pointing down in the middle of the MHC binding grove and appears to interact with the positively charged 156R. On the other hand the NY10-260D peptide takes a flattened conformation with P5 in the D pocket, P4 in the B pocket and P3 in the A pocket, leaving P1 and P2 sticking out of the N terminus of the binding cleft (Figure 3.3.12, A blue). The interaction between P8 and 156R is lost with the E260D change. The conformation of the NY10-D may explain the reduction in immunogenecity observed. We suggest that the difference in immunogenicity is due to pMHC interaction with TCR, rather than presentation differences of the peptide. Although peptide-MHC binding assays would be needed for confirmation of the affinity and crystal structure studies.

These preliminary data support the hypothesis that the HLA-B*35 restricted NY10 response may be important for a control of HIV-1 replication. Nevertheless only one HLA-B*35:08 patient was found to make the response and the study of a larger number of individuals carring the allele is required to confirm our findings.

Marked differences on peptide presentation between closely related HLA alleles is observed, this suggests that $CD8^+$ T cell responses may play an important role on HLA-B*35 control of HIV-1 infection. Nevertheless there are other factors that need to be taken into consideration, the D260E change between HIV-1 clades has been shown to increase the VRC by 30% [16]. Therefore the virus is inherently fitter in the context of B clade infection and this may contribute to the poor performance from HLA-B*35:01 individuals.

3.4.1 Summary comments

In conclusion, we observe substantial differences in markers of HIV-1 disease outcome associated with small differences between HLA-B*35 alleles. Our data challenge the PY/PX division of HLA-B*35 alleles to separate those with a null effect from those with a negative effect on disease progression. These data support the notion that differences in disease outcome of HLA-B*35

alleles may be associated with the specificity of $CD8^+$ T cell responses with a particular focus to the importance of preserved responses to Gag epitopes.

There are however limitations to the study that need to be taken under consideration. For example the IFNγ ELISpot assays and tetrameter staining utilized does not confirm an effective response or the ability to suppress virus. The ability of $CD8^+$ T cells to kill infected cells need to be further assessed. The number of studied subjects is also a limitation.

3.4.2 Future work

Fine tuning of the HLA-B*35:08-NY10-260 E and D crystal structures and pMHC binding studies are in process by our collaborators. We will also carry out clonal sequence of subject 4888670 and of the individual carrying the E260D change from pro-viral DNA to plasma RNA to try to confirm that they were infected with the WT E260 and virus escaped to E260D as they progressed. Further recruitment of B clade HIV-1 infected individuals with HLA-B*35:08 allele is needed to obtain statistical significance.

Chapter 4

Impact of HLA-selection pressure on HIV-1 fitness at a population level

4.1 Introduction

It has been established that HLA exert selection pressure on HIV-1 to escape effective $CD8^+$ T cell immune responses [53, 97, 75]. Studies on HIV-1 adaptation to HLA have suggested that protective HLA alleles restrict $CD8^+$ T cell responses that impose a strong selection pressure on particular Gag epitopes [45] for which fitness costs has been demonstrated either through fitness assays [74, 18, 29, 137] or through in vivo reversion following transmission to HLA-mismatched individuals [88, 137]. It has also been shown that the frequency of escape variants in a given population is correlated with the prevalence of the relevant HLA allele in that population [64]. Although the escape mutants enable HIV-1 to evade $CD8^+$ T cells, the accumulative cost of multiple Gag escape mutants on viral replicative capacity facilitates immune control via remaining $CD8^+$ T cell responses.

It is well established that HIV-1 adaptation at a population level is already occurring, what is less clear is the speed, and the extent to which it is occurring, and the practical consequences of this phenomenon. Nevertheless, HIV-1 evolution in a population may be relevant to vaccine design as established associations with disease outcome may be altered over time [64], and these may differ from one population to another.

To better understand the effect that protective HLA alleles have on HIV-1 adaptation over-time at a population level, we compared two ART-naïve B clade HIV-1 infected cohorts that

have a difference in frequency of protective alleles. Low prevalence of protective alleles (HLA-B*57/58:01/81:01) was observed in the Mexican cohort, in contrast the Barbados cohort exhibit high prevalence of protective alleles (Figure 4.1.1). As protective alleles are thought to drive escape mutations inferring viral fitness costs, we hypothesized that these types of mutations have accumulated over time in Barbados and that the population viral fitness might therefore be lower in this population when compared to the Mexican cohort, where the immune pressure for escape mutation is lower. To further test our hypothesis we also compared the accumulation of escape mutation with fitness cost driven by HLA-B*27/57/58:01/81:01 between cohorts.

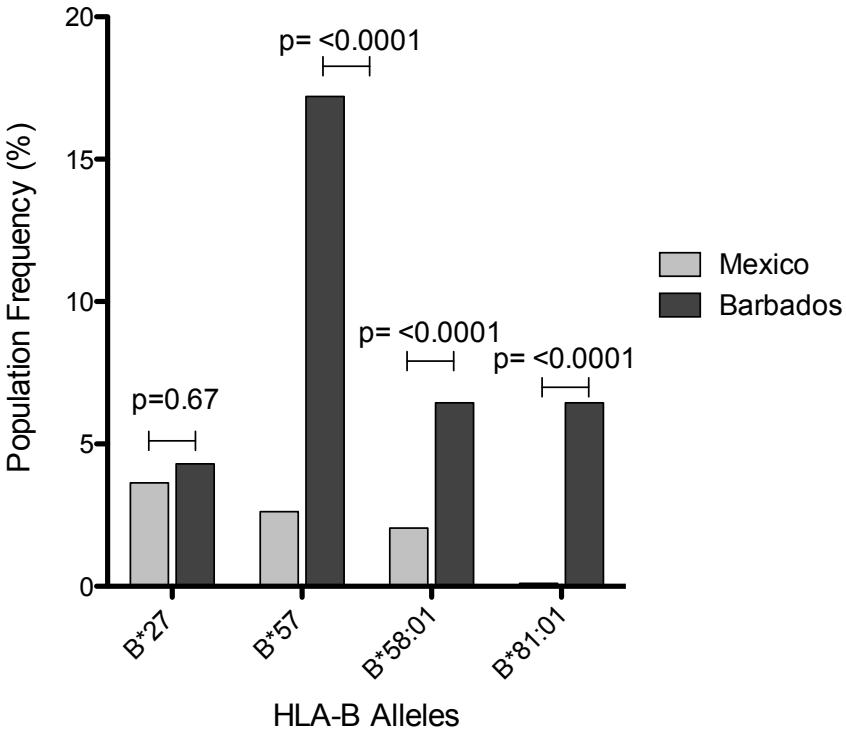


Figure 4.1.1: **Population frequency of protective class I HLA-B alleles** for HIV-1 B clade chronically infected individuals from Mexico (n=976) and Barbados (n=190). Marked differences in frequency of HLA-B*57, B*58:01 and B*81:01 can be observed between Mexico (blue) and Barbados (red).

Specific aims

The present study focuses on two geographically distinct but closely related cohorts, one in Mexico described in the previous chapter (chapter 2) and one in Barbados. The specific aims are:

1. To determine whether there is a difference in markers of disease progression between our study cohorts;
2. To determine whether population level viral fitness differ from one population to another;
3. To determine whether viral escape mutants driven by previously described protective alleles have accumulated to a higher degree in the Barbados cohort, where the prevalence of protective alleles is higher.

4.2 Methods

Study cohorts

Subjects recruited as described in chapter 2.

Fitness assays

The assay is described in chapter 2. Viral replicative capacity was measured through a recombinant assay in which PCR amplicons of HIV-1-Gag protease from patients was inserted into a pNL4-3 backbone where the Gag-protease genes were absent. A disadvantage of this method is that it is done in parallel rather than competitive format, and therefore has a low sensitivity for small VRC differences. The use of recombinant clones on the other hand, allow for a consistent background in which to analyze mutations. Data should be interpreted based on the control viruses used for normalization rather that as absolute measure of viral fitness [94, 137]. We utilized a WT NL4-3 for normalization. We used aliquots from the same NL4-3 stocks so each assay was normalized with the same virus.

4.3 Results

4.3.1 Relationship between viral replicative capacity and markers of HIV-1 infection

VL and absolute $CD4^+$ T cell counts have previously been associated with in vitro RC [105]. We therefore first examined the relationship of markers of disease progression with the RC of

the chimeric viruses constructed from our study cohorts (Figure 4.3.1, n=261). No significant correlation between RC and VL was observed (R=0.07; p=0.24) (Figure 4.3.2, A). Consistent with previous studies, a significant inverse correlation between in vitro RC and CD4+ T cell counts was observed (R=0.25; p=<0.0001) (Figure 4.3.1, B).

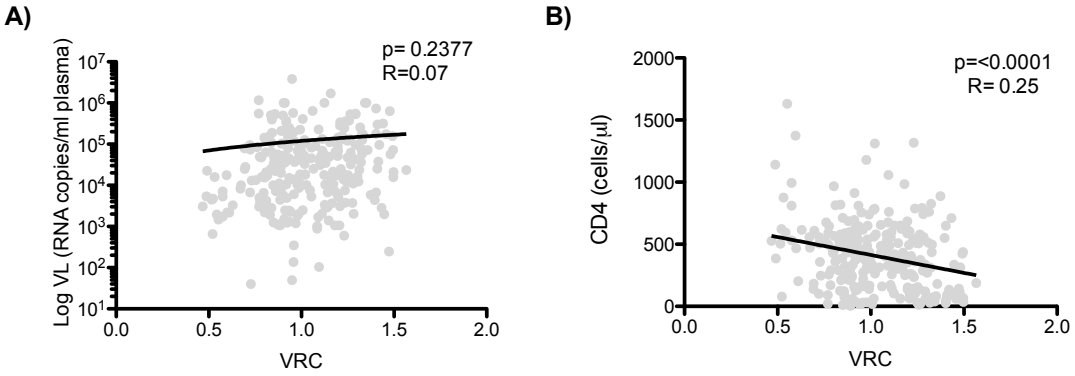


Figure 4.3.1: **Correlation between replication capacity of chimeric NL4-3 and clinical markers of HIV-1 infection.** The correlation of replication capacity of chimeric NL4-3 with VL A) and $CD4^+$ T cell counts B) (n = 261) is shown. Chimeric viruses shown are from individuals from both the Mexican and Barbados cohorts. A statistically significant negative correlation between RC and $CD4^+$ T cell counts was observed (R= 0.25; p=<0.0001).

4.3.2 Barbados has a lower median VL and higher median absolute $CD4^+$ T cell count

It has been previously shown that HLA class I alleles can have a significant effect on HIV-1 immune control. In order to investigate if the difference in frequency of protective alleles between out study cohorts has led to a difference in markers of disease progression, we compared the median VL and median $CD4^+$ T cell counts from subjects from the Barbados cohort and subjects from the Mexican cohort, all were antiretroviral therapy-naïve B-clade infected. We observed a significant difference in the median VL and median $CD4^+$ T cell counts between cohorts. Median VL was significantly higher (median 40,774 versus 14,200 RNA copies/ml plasma; p=<0.0001) in the Mexican cohort where there is a decrease frequency of HLA-B*57/58:01/81:01 alleles. In addition the median $CD4^+$ T cell counts is also significantly lower in the Mexican cohort (median 380 vs 403 cells/µl, p=0.0007) (Fig 4.3.2).

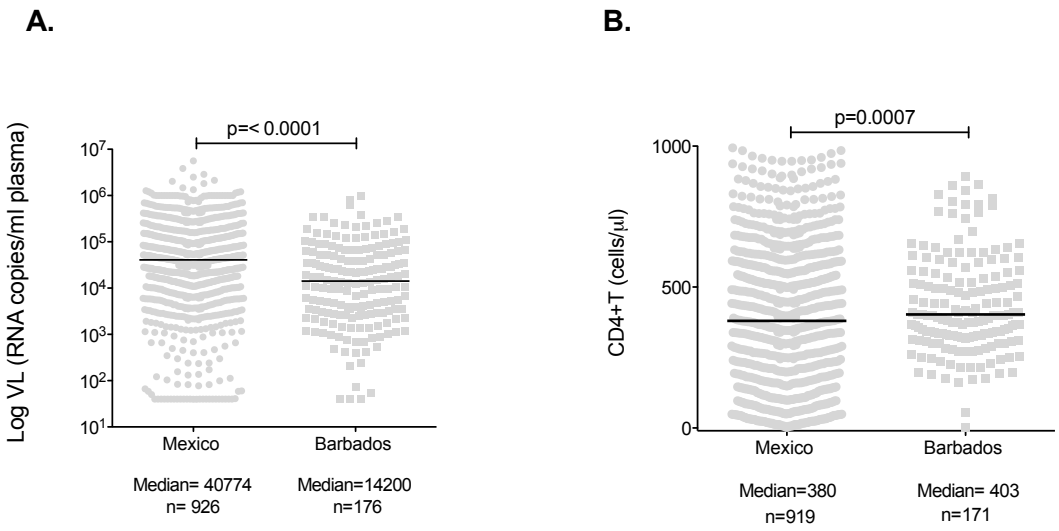


Figure 4.3.2: **Comparison of Median VL and CD4$^+$T cell counts for the Mexican versus the Barbados cohort**. VL are shown in A) and CD4$^+$T cell counts in B). P values were generated by a Mann-Whitney U test. Horizontal lines indicate median viral loads or CD4$^+$ T cell counts.

4.3.3 Gag-protease chimeric virus from Barbados have a lower RC that those from Mexico

VRC analysis between cohorts was carried out to investigate if the difference in prevalence of protective alleles in our study cohorts have led to accumulation of escape mutations with fitness cost in Barbados and therefore to a difference in VRC at a population level. Analysis was done on a subset of individuals with a CD4$^+$ T cell count between 300-500, that did not differ between the two cohorts, to avoid for bias for disease progression. Consistent with our hypothesis, the median VRC in Mexico was found to be significantly higher (1.27 vs 1.034, p=0.003) (Figure 4.3.3, A). Although the comparison groups were CD4$^+$ T cell matched, a significantly higher median VL was still observed in the Mexican subset (p=0.04) (Figure 4.3.3, B-C).

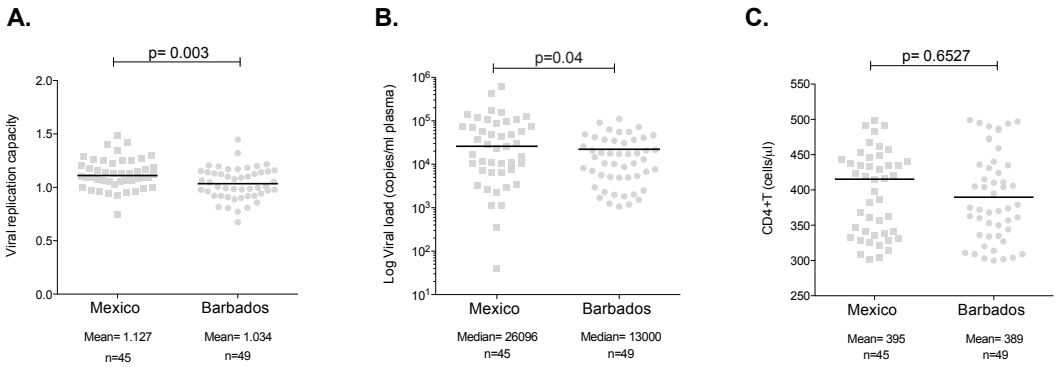


Figure 4.3.3: **Comparison of VRC between Mexico and Barbados matched by CD4$^+$T cell counts.** VRC of chimeric gag-protease-NL4-3 was compared between individuals from Mexico and Barbados, with a CD4$^+$T cell count between 300-500 cells/µl (A). Comparison of VL (B) and CD4$^+$T cell counts (C) show that in these CD4$^+$T cell matched groups a significant difference in VL is still observed. P values were generated either by a Mann-Whitney U test for non-normally distributed data or using the Student's t-test for normally distributed data, horizontal lines indicate median or mean respectively.

To establish that the difference observed is not solely reflected by the high frequency of individuals with protective alleles in Barbados and is indeed at a population level. We repeated the analysis the CD4$^+$ T cell matched individuals in the absence of individuals carrying protective alleles. VRC remained significantly higher in the Mexican cohort (p=0.04) (Figure 4.3.4, A). Furthermore the median VL remained significantly higher in this cohort (p=0.03) (Figure 4.3.4, B).

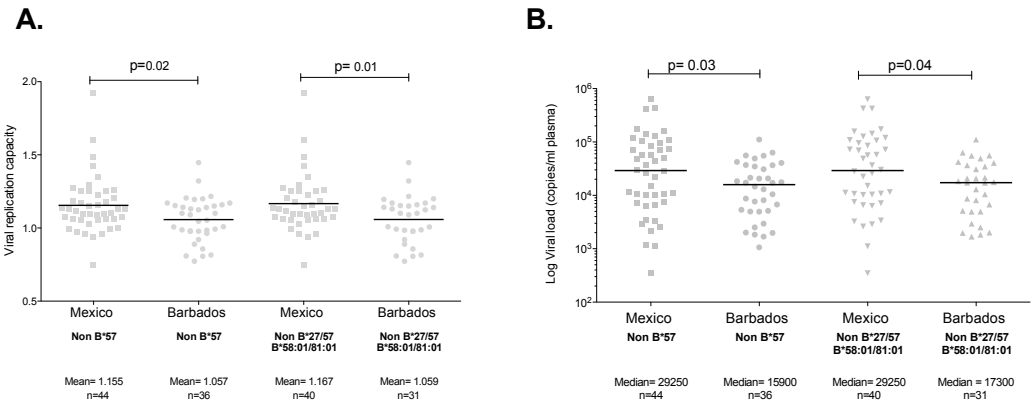


Figure 4.3.4: **Analysis of VRC and VL in the CD4$^+$T cell matched groups from both cohorts, after removal of individuals with protective alleles.** VRC of chimeric gag-protease-NL4-3 was compared between individuals with out protective alleles from Mexico and Barbados, with a CD4$^+$T cell count between 300-500 cells/µl. Comparison of RC shown in panel A) and VL B). P values were generated either by a Mann-Whitney U test for non-normally distributed data or using the Student's t-test for normally distributed data, horizontal lines indicate median or mean respectively.

4.3.4 Accumulation of 3 Gag escape mutation with viral fitness is found in Barbados in comparison to Mexico

To further investigate if the difference on markers of disease progression and VRC observed between our study cohorts is associated with the accumulation of escape mutations with fitness cost in Barbados, where a higher frequency of protective alleles is observed. We analyzed eight escape mutation were fitness cost have previously been inferred and that were restricted by the studied alleles, namely A146X, I147X, A163X, T186X, T242X, G248X, R264X, T310X [74, 39, 88, 29, 16] (Table 4.1). A phylogenetic tree of Gag sequences studied from both cohorts can be seen in Figure 4.3.5.

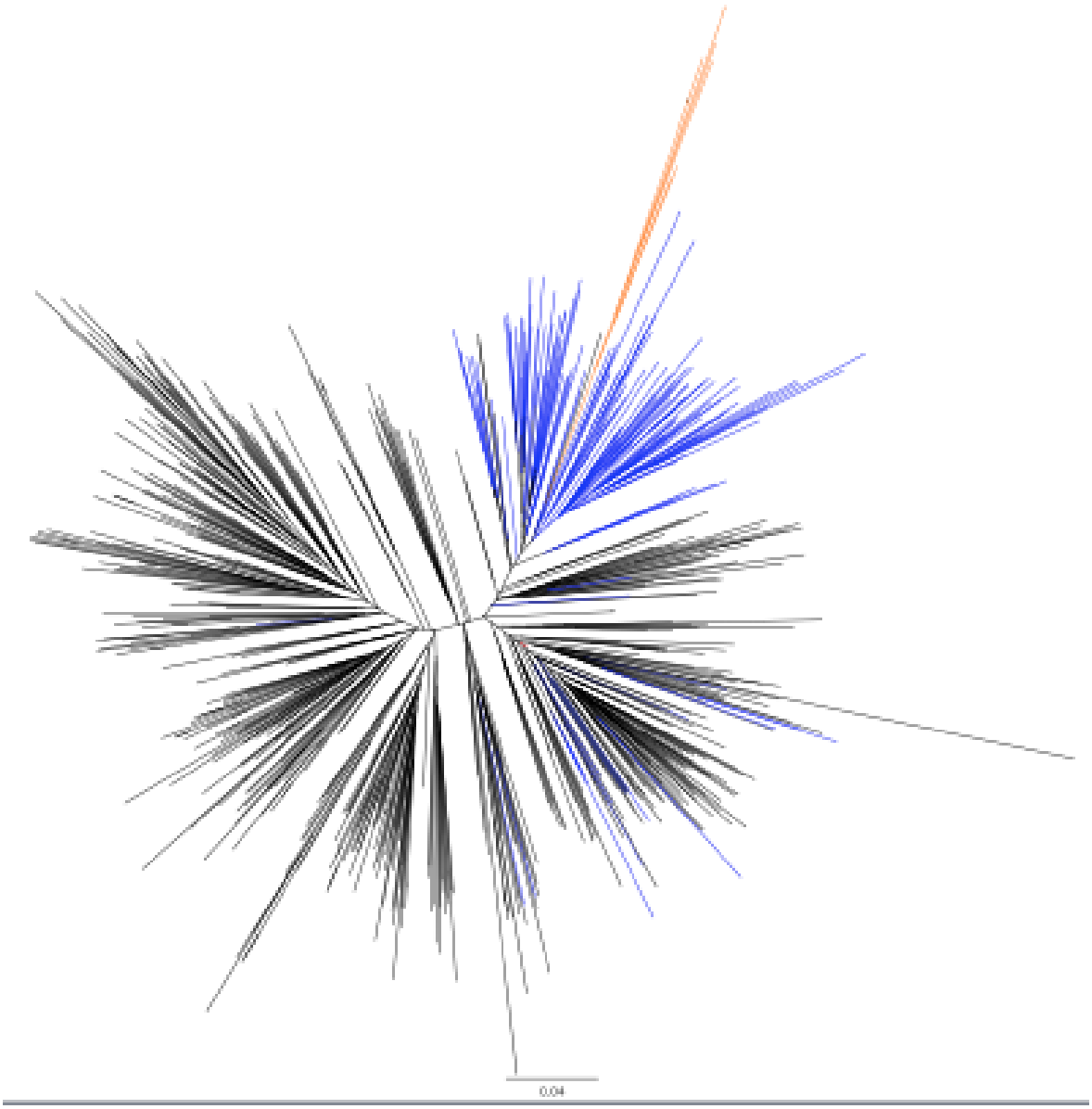


Figure 4.3.5: **Maximum likelihood phylogenetic tree of the full Gag region** of HIV-1 nucleotide sequences from 574 HIV-1 infected subjects from Mexico (black), 181 sequences from HIV-1 infected individuals from Barbados (blue), B clade consensus (red), C clade, A clade, D clade and H clade consensus sequences (orange). Consensus reference sequences were downloaded from Los Alamos sequence database. The tree was created using GARLI software and edited in Figtree v1.2.1

We analysed the Gag sequences of the viruses included in the VRC measurements. As anticipated, the frequency of these HLA-B*27/57/58:01/81:01-associated escape mutations was significantly higher in the viruses generated from the Barbados cohort (p=0.004, Fig 4.3.6. A),

where the prevalence of these HLA alleles is substantially higher (Fig 4.1.1). To test if the accumulation of such mutations was present in individuals without the restricting alleles, we reanalyzed the data in the absence of the individuals with the restricting HLA alleles for each mutation and found that all remain significantly higher in Barbados (p=0.01, Figure 4.3.6, B).

Escape mutation	Epitope name	Epitope sequence	Reference	Study clade
B*57-I147X	ISW9	ISPRTLNAW	Crawford, et.al. 2009	C
B*57-G248A	TW10	TSTLQEQIGW	Boutwell, et.al. 2013	B
B*58:01-T310X	QW9	QATQDVKNWM	Matthews, et.al. 2008	C
B*57/58-T242N	TW10	TSTLQEQIGW	Boutwell, et.al. 2013	B
B*57-A163G	KF11	KAFSPEVIPMF	Boutwell, et.al. 2013	B
B*27-R264K	KK10	KRWIILGLNK	Boutwell, et.al. 2013	B
B*27-L268M	KK10	KRWIILGLNK	Boutwell, et.al. 2013	B
B*81-T186S	TL9	TPQDLNTML	Wright, et.al. 2012	C

Table 4.1: **Escape mutations where fitness cost has previously been inferred.** Table shows nine mutations included in the study.

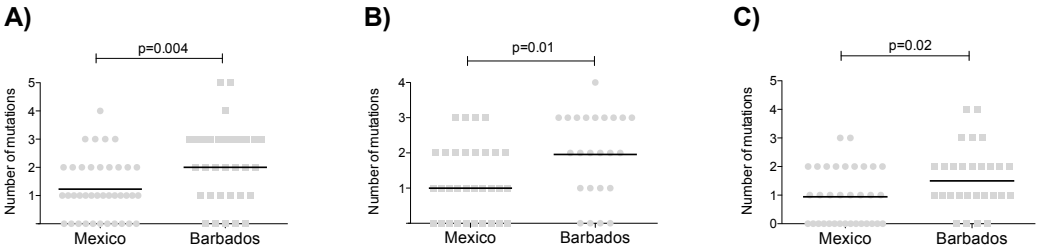


Figure 4.3.6: **Frequency of Gag escape mutants associated with fitness cost in Mexico and Barbados study subjects.** A. Frequency of selected Gag escape variants in Mexican and Barbados populations (see table 4.1). B. Frequency of the same variants in the HLA-mismatched subjects. C. Frecuency of selected Gag variants excluding the I147L mutant.

To determine whether the differences we observed between the two cohorts in median viral loads and mean viral replication capacity might be related in part to differences in the founder viruses in Barbados and Mexico, maximum likelihood phylogenetic trees were constructed for each cohort with the aim of estimating the founder virus in each case (Table 4.2). They show that the estimated ancestral sequence for the Barbados and Mexico sequences differed by 7 amino acids within Gag, including I147L which is one of the HLA-B*57-associated mutants. We therefore excluded the I147L mutant from the analysis to determine whether the differences in viral replicative capacity observed in Barbados and Mexico could be explained by this estimated founder effect. We observed that there is still a significantly greater number of escape

mutants associated with the protective HLA alleles HLA-B*27/57/58:01/81:01 in the Barbados sequences analyzed for viral replicative capacity, even in subjects not expressing these particular protective HLA alleles (mean 1.5 versus mean 1.0, p=0.02, Fig 4.3.6. C). Thus, the differential viral replicative capacities observed in Barbados and Mexico are not explained by HLA-B*27/57/58:01/81:01-associated mutation differences in founder viruses.

Differences from HXB2 of estimated ancestral Gag amino acid sequence in Barbados and in Mexico												
HXB2 Gag	30	76	84	94	124	126	138	147	159	215	280	312
HXB2 Sequence	K	R	T	I	H	N	I	I	V	V	T	E
Barbados	-	K	V	V	N	S	L	L	I	L	V	D
Mexico	R	V	V	V	N	-	L	-	-	L	-	-

Table 4.2: Estimation of founder sequences for the Mexico and Barbados cohorts was undertaken by the reconstruction of maximum likelihood phylogenetic trees using the FastML web server [6].

To further investigate differences in HIV-1 adaptation of HLA-B*57 in Mexico and Barbados we compared markers of disease progression in individuals with HLA-B*57 from both cohorts. Although not significant, a trend towards lower median VL and higher median $CD4^+$ T cell counts in individuals with HLA-B*57 in Mexico was observed, in comparison to individuals with HLA-B*57 from Barbados (Figure 4.3.7).

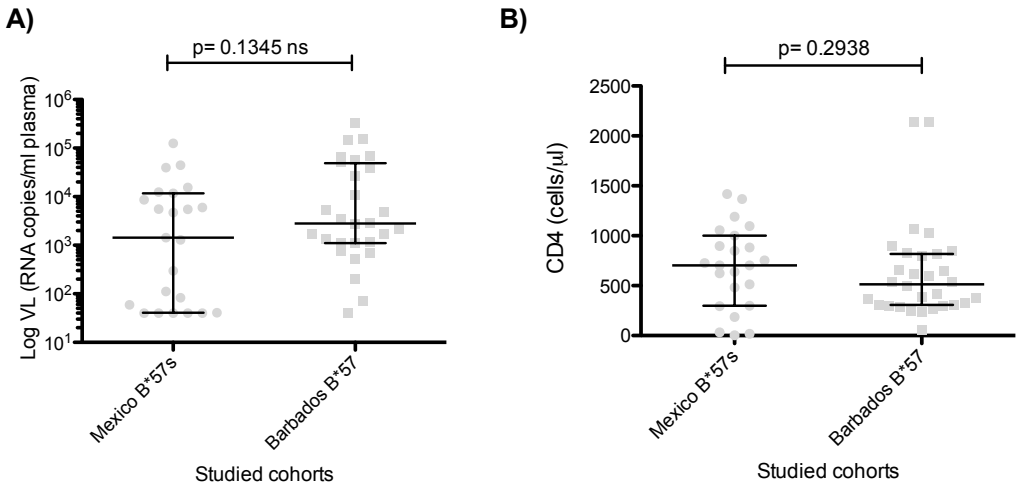


Figure 4.3.7: **Comparison of median VL and median CD4$^+$T cell in individuals with HLA-B*57 from Mexico and Barbados.** Markers of disease progression were compared between individuals with HLA-B*57 from Mexico and Barbados. Comparison of VL shown in panel A) and CD4$^+$ T cell counts in B). P values generated by a Mann-Whitney U test. Horizontal lines indicate median with interquartile range.

4.4 Discussion

The outcome of the transmission of escape mutations in not intuitive. It has been suggested that escape mutations that have a significant cost on VRC will most likely revert when selection pressure is lost, as it is transmitted to a mismatched host. However in a population where the frequency of protective alleles is high, it is likely that the virus transmitted most frequently will be that with the escape mutation and this may lead to this particular strain to persist in the population. Indeed, it has previously been suggested that escape variants can be transmitted and achieve fixation in certain populations in an allele frequency dependent manner [64]. These may be explained as successful transfer of escape mutations to new recipients has been shown to be possible, and is thought to be dependent on their frequency in the donor [44].

The present study looked to investigate the effect of frequent protective HLA class I B alleles, B*57/58:01/81:01, in HIV-1 VRC at a population level. We hypothesized that in vitro RC of HIV-1 may have decreased notably over the course of the epidemic in a populations with high frequency of protective alleles, when compared to a population with low frequency of protective alleles.

We compared two HIV-1 B clade infected cohorts in which protective HLA alleles (HLA-B*27/57/58:01/81:01) are expressed at 10% and 35% respectively. Consistent with our hypothesis we found that Mexico, the cohort with a lower frequency of protective alleles, had a worst outcome on markers of disease progression consistent with the infection by a fitter virus and exhibit a significantly higher median VL and significantly lower median $CD4^+$ T count (Figure 4.3.2).

We then analyzed the relationship between RC and markers of disease progression. We found that RC is negatively correlated with $CD4^+$ T cell counts in our study cohorts, but no correlation was found with VL (Figure 4.3.1).

VRC was therefore assessed in a subset of chimeric viruses from individuals with a $CD4^+$ T cell count between 300-500 cell/µl to avoid disease progression bias and due of the observed negative correlation of RC with CD4 $^+$ T cell counts. Our data show significant difference in the median VRC between the two study populations, Barbados having a less fit virus (p=0.003; Figure 4.3.3, A). This difference remained significant even when patients with protective alleles were removed from the analysis (p= 0.04; Figure 4.3.4, A). Thus the data presented here support our hypothesis that high prevalence of protective HLA alleles is associated with low population level viral fitness. Additionally, consistent with our hypothesis, a significantly higher median VL remained in the Mexican subset of subjects matched by $CD4^+$ T cell counts in comparison to the Barbados subset (Figure 4.3.4, B; and Figure 4.3.4, B). Again suggesting that the higher viraemia in Mexico is related to the lack of immune pressure for escape mutation with fitness cost, therefore a fitter virus in this cohort.

Moreover, we looked at the frequency of previously described Gag-escape variants with fitness cost driven by B*27/57/58:01/81:01. We found three variants that are significantly accumulated in the Barbados cohort in comparison to Mexico (p=0.004, Figure 4.3.6, A). Accumulation was at population level rather than just in individuals with the HLA driving the escape mutation (p=<0.0001, Figure 4.3.6, B). Supporting the notion that escape mutations with fitness cost can be transmitted and remain fixed in the population regardless of the fitness cost. This could possibly be as these are dominant variants in a donor and are therefore able to establish infection in recipients [44].

Additionally we hypothesized that if the virus in Barbados was better adapted to protective

alleles, such as HLA-B*57, due to the high frequency of the alleles in the population, then some loss of protection will be observed. We therefore compared individuals with HLA-B*57 in Barbados to individual with HLA-B*57 in Mexico, although no significant, a trend towards better protection in Mexico was observed with a lower median VL and higher median $CD4^+$ T cell count (Figure 4.3.7). Suggesting lost of protective capacity by HLA-B*57 in Barbados. This observation supports the notion that pathogens adapt to the most common alleles in a population.

The data presented in the study are the first suggesting that high frequency of protective alleles can shape viral evolution to accumulate a higher number of Gag escape mutants, resulting in alteration of VRC at a population level. Thus viral loads and disease progression rates may differ between distinct populations as a result of the frequency of protective alleles in the respective populations.

However there are important limitations to the study to consider. For example, studies have shown that transmission of viruses with escape mutants in critical Gag epitopes to individuals carrying the relevant restricting allele, leads to failure to control viremia [53, 52, 64]. Therefore, a selection of compensatory mutations need to be taken into account.

Resolving the impact of HLA alleles on HIV-1 evolution is difficult. It is obscured not only by the founders effect or by the proportion of mutations that revert (with different reversion rates) to WT after transmission; but also different alleles can drive the same mutation and only a small number of escape mutation with fitness cost have been analyzed and were included in this study study. Moreover non-reverting mutations that do not substantially contribute to the reduction of median VL (possibly as they do not cause a fitness cost or are corrected by compensatory mutations) might accumulate more readily. The difference in VRC may therefore not be as pronounced as expected. A longitudinal study comparing the VRC overtime in a population with high frequency of protective alleles may provide more conclusive data.

Nonetheless, the data presented here are consistent with the results from a recently carried out small study that analyzed Gag-Pro mediated HIV-1 RC during the course of the epidemic in Japan [105]. The study showed a significant reduction in RC of chimeric virus over 15 years of HIV-1 infection, thought to be caused by viral evolution in response to host immune selection pressure. However, these study did not observe a change in median VL overtime. We

hypothesize that this is due to the low frequency of protective alleles in Japan, a population in which HLA-B*27 and HLA-B*57 have previously been reported to be extremely rare [101].

Further larger studies from multiple geographical regions are required to confirm the effect that HLA have on viral evolution at a population level.

4.5 Summary and comments

The data presented in this chapter support the notion that HLA driven viral escape can influence HIV-1 evolution at the population level, to a less fit state in this case. Moreover this has an impact on clinical parameters and disease outcome, with lower median VL observed in Barbados, even for $CD4^+$ T cell matched individuals. Whether this will result in HIV-1 adaptation to HLA, leading to the loss of protection associated with B*57 over time in Barbados remains to be seen. The accumulation of escape variants in a population, even for rapidly reverting mutations, may be explained by the rate of mutant acquisition exceeding the reversion rate; or by the selection of compensatory mutations slowing or completely halting reversion.

VRC has been proven to be an important component of immune control of HIV-1, as such, more studies are needed to determined the rate at which HIV-1 adaptation to HLA is occurring in individuals as well as in populations. Whether adaptation is likely to drive viral evolution is dependent of the fate of the mutations after transmission. However the rate of reversion due to fitness cost is yet to be defined, and is one of the challenges to understand HLA driven HIV-1 evolution.

HIV-1 evolution in a population may be relevant to vaccine design as established associations with disease outcome may be altered over time [54, 64]. The long-term clinical benefit of immune-driven fitness costs also remains uncertain given the lack of correlation with longitudinal markers of disease progression. Little research has been carried out to analyze population level alterations in HIV-1 RC. To our knowledge none has examined the potential effect of immune escape fixation on population level viral fitness.

Chapter 5

Production of a C clade plasmid system to be used use as reference for fitness experiments

5.1 Introduction

The first full-length HIV-1 isolate, NL4-3, was constructed in 1986 [1]. Since then HIV-1 isolates have become important research tools used to investigate the biological characteristics of HIV-1. This includes the analysis of HIV-1 RC, drug resistance, testing of neutralizing antibodies and evaluation of drug and vaccine efficiency. Hence numerous molecular clones have since been produced, however most clones generated to date that are replication competent, are B clade virus.

Among the subtypes of the M group of HIV-1 the C clade is the most prevalent of the global HIV-1 infections. C clade HIV-1 accounts for approximately 70% of infections in Southern and Northern regions of Africa, a region where the epicenter of the HIV-1 epidemic is found [95, 103, 104, 58, 139]. Nevertheless MJ4 and Indie-C1 are the only C clade molecular clones produced to date that use CCR5 as co-receptor [104, 58]. Yet, both of these isolated have drawbacks that have limited their wide utilization. MJ4 growth is slow, with a replication peak significantly lower than NL43; while Indie-C1 sequence is consensus to the India strain [103, 104].

NL4-3, therefore continues to be widely used. However there is important constraints to consider while utilizing the backbone of an early laboratory B clade strain, and caution must be taken while conclusions are made regarding current circulating virus and more importantly

when studying the impact of C clade mutations. The challenges originated by the use of a B clade backbone have been partially tackled by the use of Gag-protease chimeric viruses. However this approach is limited to Gag-protease genes and the use of recombinant B and C clade clones may allow for the specific genetic context of the backbone to differentially interact with the tested mutation and result in a false measurement. For example, the effect of the Gag-A163G change on VRC analyzed by Boutwell et.al in the context of a B clade Backbone (NL4-3) is greater than the one previously reported in the context of a C clade backbone (MJ4) [16]. An other example is the Gag-T163S change, for which Wright et.al showed a difference in impact magnitude, when studied in the context of a B clade or C clade backbone [138]. Moreover the presence of compensatory mutations, which may just relate to differences in consensus sequence between clades, could also mask the impact of a mutation of interest.

In this chapter we therefore aimed to produce a C clade two plasmid system, similar to the p83-1 and p83-10 (NL4-3), to use as reference for replication experiments. The two plasmids system is preferred to facilitate mutagenesis, handling and biosafety.

5.2 Methods

The primary virus from an infant termed as 197B was selected from a well characterized cohort of C clade HIV-1 infected infants from South Africa, to serve as template for the two-plasmid system. The infant was a rapid progressor and the virus was shown to be fast replicating [112]. Virus stocks were provided by Julia Garcia Prado.

PBMC infection for template DNA extraction $10x10^6$ PBMCs from two healthy donors were stimulated together in 20 ml R10 media supplemented with 2 U/ml IL-2 and 5 µg/ml PHA for 3 days. PBMCs were then pelleted and incubated in 500 µl of 197B virus stock for 2 hours at 37 ℃, in a 50 ml falcon tube. PBMC's were then pelleted and washed with PBS. PBMCs were resuspended at 1 million/ml R10 media supplemented with 2 U/ml IL-2 and incubated at 37 ℃ and 5% CO_2. New stimulated donor PBMCs were added as feeders every 4 days and viral growth was analyzed by p24 ELISA. When the p24 concentration was found at >200,000 ng/ml, cells were pelleted for DNA extraction as described in chapter 2.

Amplification of the HIV-1 genome in two segments To obtain an infectious molecular clone of the infant 197B, we amplified overlapping 5' and 3' half genome from proviral DNA extracted from infected PBMCs by single round PCR using the High-Fidelity DNA Taq polymerase (Invitrogen, US). Both fragments contain a complete LTR element and overlap by 36 base pairs in Vpr. For the 5´ half genome, U3-R-U5, Gag, Pol, Vif, and Vpr (5.7kb) were amplified with the following primer set 5'LTRF and 5'VprR. For the 3´ half genome, Vpr, Tat, Rev, Vpu, Env, Nef, and U3-R-U5 (4kb) were amplified utilizing the following primer set 3'VprF and 3'LTR (Figure 5.2.1). As indicated in figure 5.2.1, MluI and NotI restriction sites were appended to the sense and antisense primers, respectively. The primers used for amplification were modified from Ndung'u et.al 2000 and Grisson et.al 2004, to complement the confirmed 197B sequence [103, 58].

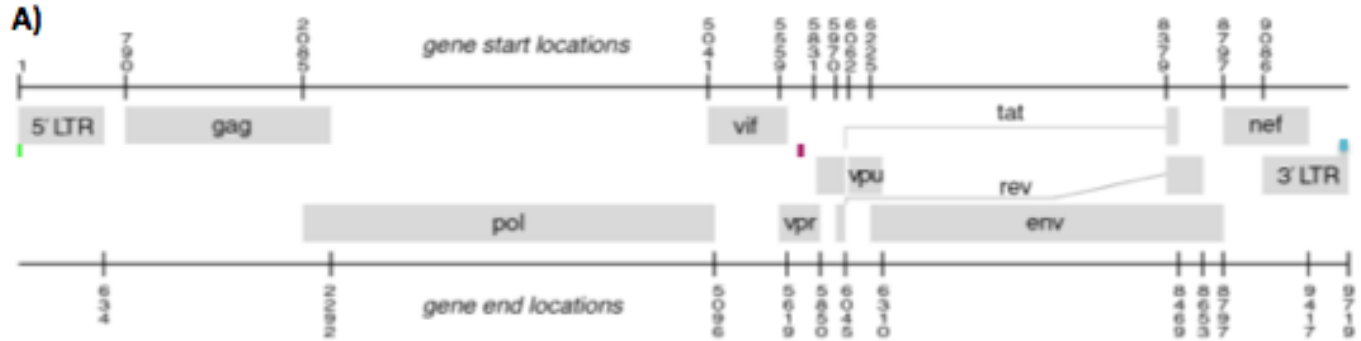
High fidelity Taq Polymerase was used for amplification. This enzyme was chosen because it is composed of a recombinant Taq DNA polymerase that possesses proofreading activity [132]. This in turn increases fidelity by six fold approximately from the regular taq polymerase.

50 µl reaction mix was used per amplification:

5 µl	10x High Fidelity PCR buffer
1 µl	5' primer (10 µM)
1 µl	3' primer (10 µM)
1.5 µl	$MgSO_4$ (50 mM)
1 µl	dNTP mix (10 mM)
0.2 µl	Platinum Taq High Fidelity
38.3 µl	RNase/DNase-free H_2O
2 µl	Template DNA

Amplification was carried out in the Peltier Thermal Cycler PTC-200.

1. Denaturation 94 °C 2 minutes
2. Denaturation 94 °C 30 seconds
3. Annealing 55 °C 30 seconds
4. Extension 68 °C 1 minute per kb of PCR product



B)

Primer	Sequence	Length	HXB2 Location	Alignment used
5'LTRF	CCGCGGGCACGCGTGTGGAAGGGTTAATTTACTCCAAGAAAAGGCAAG	48	1→33	HIV1 Complete DNA, 229 sequences*, Filtered web alignment
5'VprR	CGACTCCTGTCCAAGCGGCCGCATATGTTTCATAGAC	37	5694←5730	HIV1 Complete DNA, 229 sequences*, Filtered web alignment
3'VprF	GTCTATGAAACATATACGCGTTGGACAGGAGTCG	34	5694→5730	HIV1 Complete DNA, 229 sequences*, Filtered web alignment
3'LTRR	CGCGCGGCCGCACTACCACAAAGGGTCTGAGGGATCTCTAGTTAC	38	9669←9706	HIV1 Complete DNA, 229 sequences*, Filtered web alignment

MluI restriction site: ACGCGT

NotI restriction site: GCGGCCG

Figure 5.2.1: **Overview of primers utilized for the amplification of the 5' and 3' halves of the 197B viral genome.** Primer sequence was modified from Ndung'u et.al 2000 and Grisson et.al 2004 to accommodate MluI and Not I restriction sites and to match 197B sequence. Location of primer binding to HIV-1 was drawn using the Los Alamos database QuiclAlign tool www.hiv.lanl.gov.

5. Repeat steps 2-4 35 times
6. Final extension 68 °C 10 minutes
7. Cooling 4 °C

Full sequencing was carried out using 35 primers shown in table 5.1 as described in chapter 2.

Primer	Sequence	Length	Tm	GC%	HXB2 location
HIV1C-1F1	GAAGGAGAGAGATGGGTGC	19	54.8	57.9	779-797
HIV1C-1F2	CATCAAGCAGCCATGCAAATG	21	55.5	47.6	1369-1389
HIV1C-1F3	GGCACATAGCCAGAAATTGC	20	54.7	50.0	1985-2004
HIV1C-1F4	GTCCCATTGAAACTGTACCAG	21	53.4	47.6	2557-2577
HIV1C-1F5/2F2	GGGATTTACCACACCAGACAAG	22	55.6	50.0	3185-3206
HIV1C-1R1	TCCCACTCAGGAATCCAGG	19	56.2	57.9	3775-3793
HIV1C-1R2/2R5	GTTTCTTGTCTGGTGTGGTAAATCC	25	56.1	44.0	3186-3210
HIV1C-1R3	CTGGTACAGTTTCAATGGGAC	21	53.4	47.6	2557-2577
HIV1C-1R4	CCTGCAATTTCTGGCTATGTGC	22	56.9	50.0	1986-2007
HIV1C-1R5	CATTTGCATGGCTGCTTGATG	21	55.5	47.6	1369-1389
HIV1C-2F1	ACAATGGCCATTGACAGAAG	20	53.3	45.0	2615-2634
HIV1C-2F3	CCTGGATTCCTGAGTGGGAG	20	57.0	60.0	3775-3794
HIV1C-2F4	GACTGTAGTCCAGGGATATGGC	22	56.7	54.5	4392-4413
HIV1C-2F5/3F2	GTGACATAAAGGTAGTACCAAGGAGG	26	56.5	46.2	4993-5018
HIV1C-2R1	TGGTCTTCTGGGGCTTGTTC	20	57.4	55.0	5562-5581
HIV1C-2R2/3R6	CCTCCTTGGTACTACCTTTATGTC	24	54.4	45.8	4995-5018
HIV1C-2R3	GCCATATCCCTGGACTACAGTCTAC	25	58.0	52.0	4389-4413
HIV1C-2R4	CTCCCACTCAGGAATCCAGG	20	57.0	60.0	3775-3794
HIV1C-3F1	TGTAGTCCAGGAATATGGCA	20	53	45.0	4395-4414
HIV1C-3F3	CTCAAGCAGGAAGCTGTCAGAC	22	57.8	54.5	5634-5655
HIV1C-3F4/4F1	GCAGAAGACAGTGGCAATGAG	21	56.3	52.4	6209-6229
HIV1C-3F5/4F2	CAACCATAACACAAGCCTGTCC	22	56.3	50.0	6820-6841
HIV1C-3F6/4F3	GTGGAGGAGAATTTTTCTATTGC	23	52.0	39.1	7357-7379
HIV1C-3R1	GGTGCAGATGAGTTTTCCAG	20	53.8	50.0	8020-8039
HIV1C-3R2/4R5	TGCAATAGAAAAATTCTCCTCCAC	24	53.4	37.5	7357-7380
HIV1C-3R3/4R6	GGACAGGCTTGTGTTATGGTTG	22	56.3	50.0	6820-6841
HIV1C-3R4	CTCATTGCCACTGTCTTCTGC	21	56.3	52.4	6209-6229
HIV1C-3R5	GTCTGACAGCTTCCTGCTTGAG	22	57.8	54.5	5634-5655
HIV1C-4F4	GCTGAGGGCTATAGAGGCG	19	57.3	63.2	7889-7907
HIV1C-4F5	GTGGAGAGCAAGACAGAGACAG	22	57.0	54.5	8434-8455
HIV1C-4F6	GCTTTCCAGTCAGACCTCAG	20	55.2	55.0	8996-9015
HIV1C-4R1	CCTAGAGAGACCCAGTACAG	20	53.1	55.0	9532-9550
HIV1C-4R2	CTGAGGTCTGACTGGAAAGC	20	55.2	55.0	8996-9015
HIV1C-4R3	GTGCTAAGAATCCGCTCACTAATC	24	55.5	45.8	8464-8487
HIV1C-4R4	CGCCTCTATAGCCCTCAGC	19	57.3	63.2	7889-7907

Table 5.1: **List of primers used for full length sequencing of the 197B isolate.**

Cloning of fragments into pCR-XL-TOPO vector Both purified PCR fragments were digested with the MluI and NotI restriction enzymes. Band size was confirmed by gel electrophoresis followed by gel purification. Both halves were independently cloned into the MluI-NotI site of the pCR-XL-TOPO vector (Invitrogen, US), following manufacturer's instructions. The ligation mixture was transformed into XL2 Blue MRF competent cells and plated onto LB agar plates supplemented with 100 µg/ml of kanamycin and grown overnight at 30 °C. Single colonies were selected and grown overnight in LB medium with same concentration of kanamycin at 30 °C with constant shaking. Plasmid was isolated with a QIAprep spin Miniprep kit (QIAGEN,GE)

for sequencing. Selected clones were screened by enzyme digestion and gel electrophoresis as well as sequencing.

Transfection and Infection Once the sequence of both fragments was confirmed, a unique restriction site in the overlapping region of both segments was searched but found absent. Nevertheless the 3' half clones contained the EcoRI restriction site used on the p83-2/p83-10 plasmid system. We therefore decided to introduced this restriction site into the 5' half plasmid by site directed mutagenesis. As the insertion needed exceeded the QuikChange site directed mutagenesis kit recommendations, we employed a modified protocol in which megaprimers are utilized [48] (Figure 5.2.2). In short, the fragment to be inserted was amplified utilizing HiTaq polymerase using the 197B genomic DNA as template and the following primer set: MegaF (5'-CAGAGGGAACCATACAATGAATGGACACTAGAGC-3') and MegaR (5'-CGACTCACTA TAGGGCGAATTGGGAATTCTTATTATAGCTTCA-3'). The MegaF primer is consensus to the Vpr sequence upstream from the point of insertion in the destination vector. Accordingly, the 5´ end of the MegaR primer consists of 23 nucleotides downstream from the point of insertion matching the 5' clones (underlined), followed by 20 nucleotides of the target DNA fragment to be inserted. Fragment size was confirmed by gel electrophoresis, fragments were extracted using the gel extraction kit (QIAGEN, UK) according to manufacturer's instructions. The purified PCR fragment was then used for the mutagenesis reaction (QuikChangeII site directed mutagenesis, Agilent technologies, UK) according to manufacturer's instructions. In brief, we used 100 ng of the 5' half plasmid and 300 ng PCR fragment in a 50 µl reaction mix following the next program: 95 °C for 30 seconds; 18 cycles of 95 °C for 30 seconds, 55 °C for 30 seconds; and 68 °C for 2 min/kb. 10U *DpnI* were added and incubated for 3 hours at 37 °C. The plasmid was then transformed into *E. coli* XL1-Blue cells and the plasmid extracted by QIAmini prep (QIAGEN, GE). Mutagenesis was verified by EcoRI enzyme digestion followed by visualization by gel electrophoresis and by sequencing. Phylogenetic analysis was carried out as described in chapter 2 by the production of NJ trees in ClustalX.

Three clones for the 5' end of the HIV-1 genome were selected for further analysis and were called 5' clone 17, 18 and 20. Four clones were selected for further analysis for the 3' end of the HIV-1 genome and were named 3' clone 13, 23, 29 and 35.

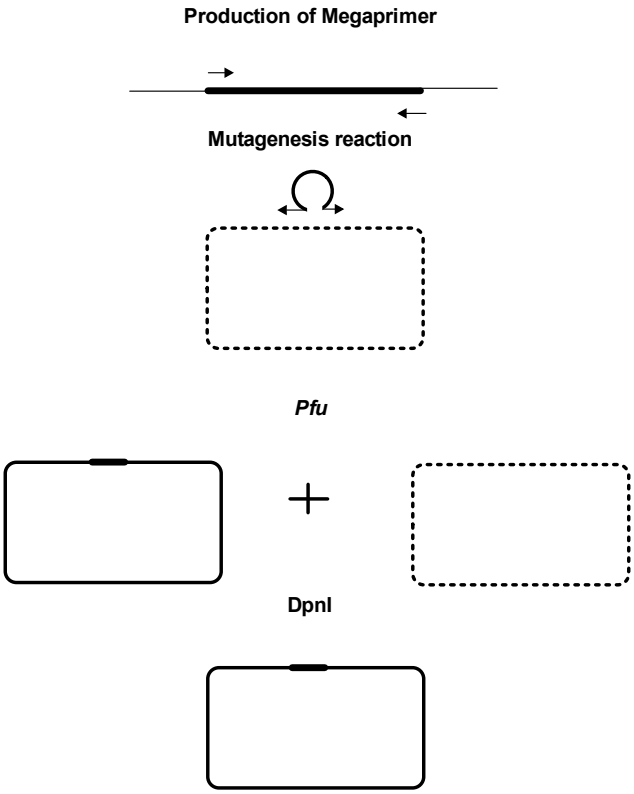


Figure 5.2.2: **Schematic description of the modified protocol for insertion of EcoRI restriction site to the 5' clones.** The DNA fragment to be inserted is amplified from template DNA. Primers used are complementary at its 5´ end, with additional nucleotides corresponding to the DNA sequence upstream and downstream, respectively, from the insertion site within the recipient plasmid (open boxes). The amplified fragment is annealed to the methylated template plasmid DNA (dotted line), and each strand is elongated with the Pfu DNA polymerase. After elongation, the newly created plasmid carries the insertion and is unmethylated (closed line). Upon DpnI digestion, the methylated template DNA is destroyed.

Viral production In order to produce viral stocks, both plasmids, 5' and 3', were digested with EcoRI and 1μg of each plasmids co-transected via electroporation at 300 V and 500 μF into $2.5x10^6$ GRX-CEM (BioRad GenePulser II). NL4-3, MJ4 and the p83-2/p83-10 plasmid system were used as positive control as CEM-GXR cells express both CCR5 and CXCR4 co-receptors. Viral growth was measured by measurement of GFP positive cells through Flowcytometry.

5.3 Results

A schematic representation of the cloning strategy used for the production of the two plasmid system is shown in figure 5.3.1. We successfully amplified both halves of the 197B genome and the visualization of the fragments can be seen in Figure 5.3.2. Both fragment were amplified using a sense primer with a MluI restriction site while the antisense primer had a NotI restriction site. Both fragments were then digested with these enzymes and cloned into the MluI and NotI sites of TOPO XL plasmid.

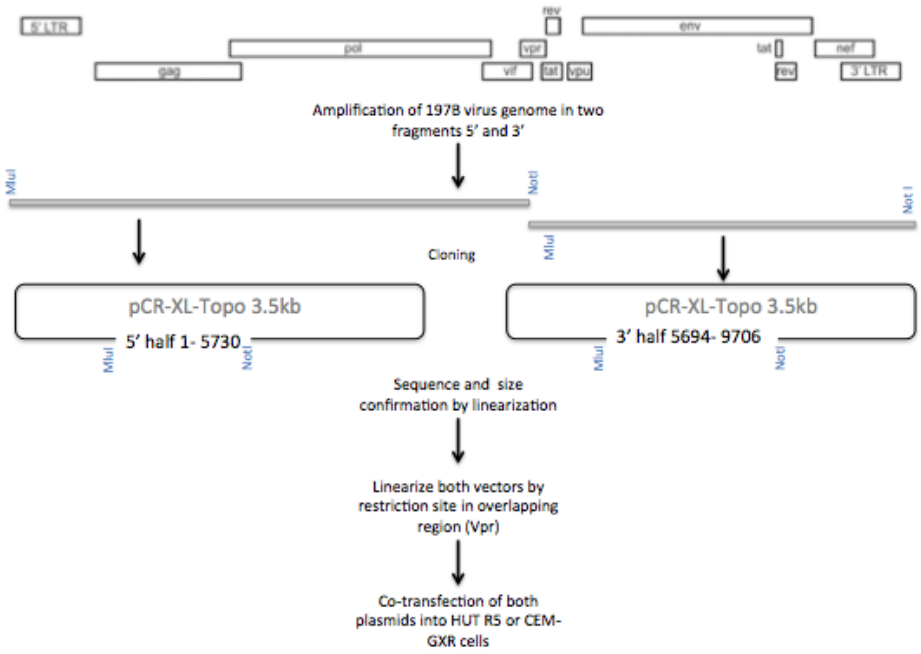


Figure 5.3.1: **Strategy for construction of the C-clade two plasmid system, similar to the p83-2 and p83-10.** 197B viral genome was amplified in two fragments and cloned into the MluI and NotI site of two pCR-XL-Topo plasmids. Each containing either the 5' or 3' half of the genome. Plasmid sequence will be verified and plasmids will be linearized by digestion with a restriction enzyme for which restriction site is located in the overlapping region of both plasmids, in the Vpr, followed by co-transfection into CEM-GXR cells to generate virus.

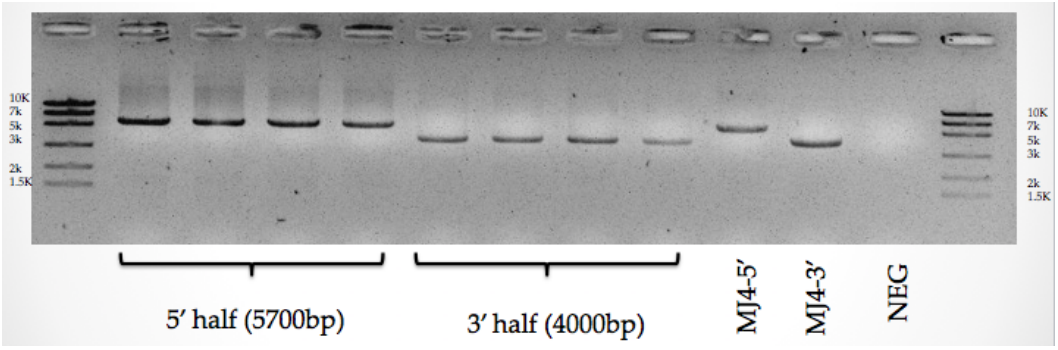


Figure 5.3.2: **Visualization of the amplified 197B half genome fragments through gel electrophoresis.**

After cloning, both plasmids were sequenced and we built NJ Phylogenectic trees of each protein to verify the presence of the full length HIV-1 genome. For the 5' clones we built phylogenetic trees for the aminoacid sequence of Gag, Pol and Vif (Figure 5.3.3). For the 3' clones we built phylogenetic trees for Tat, Rev, Vpu, Env and Nef (Figure 5.3.4). We verified that the full sequence for all of the genes was present and that in fact our sequences branched with the 197B parental sequence provided by Julia Garcia Prado, when available, and/or the PCR sequence of the amplified fragments. We also verified that the clones were C clade. We observed the conservation of Gag, as all the sequences are very closely related or the same. Pol and Vif are also very closely related. On the other hand greater variation is found between the 3' clones, this may be because the genes contained in the 3' clones are generally less conserved in HIV-1 replication.

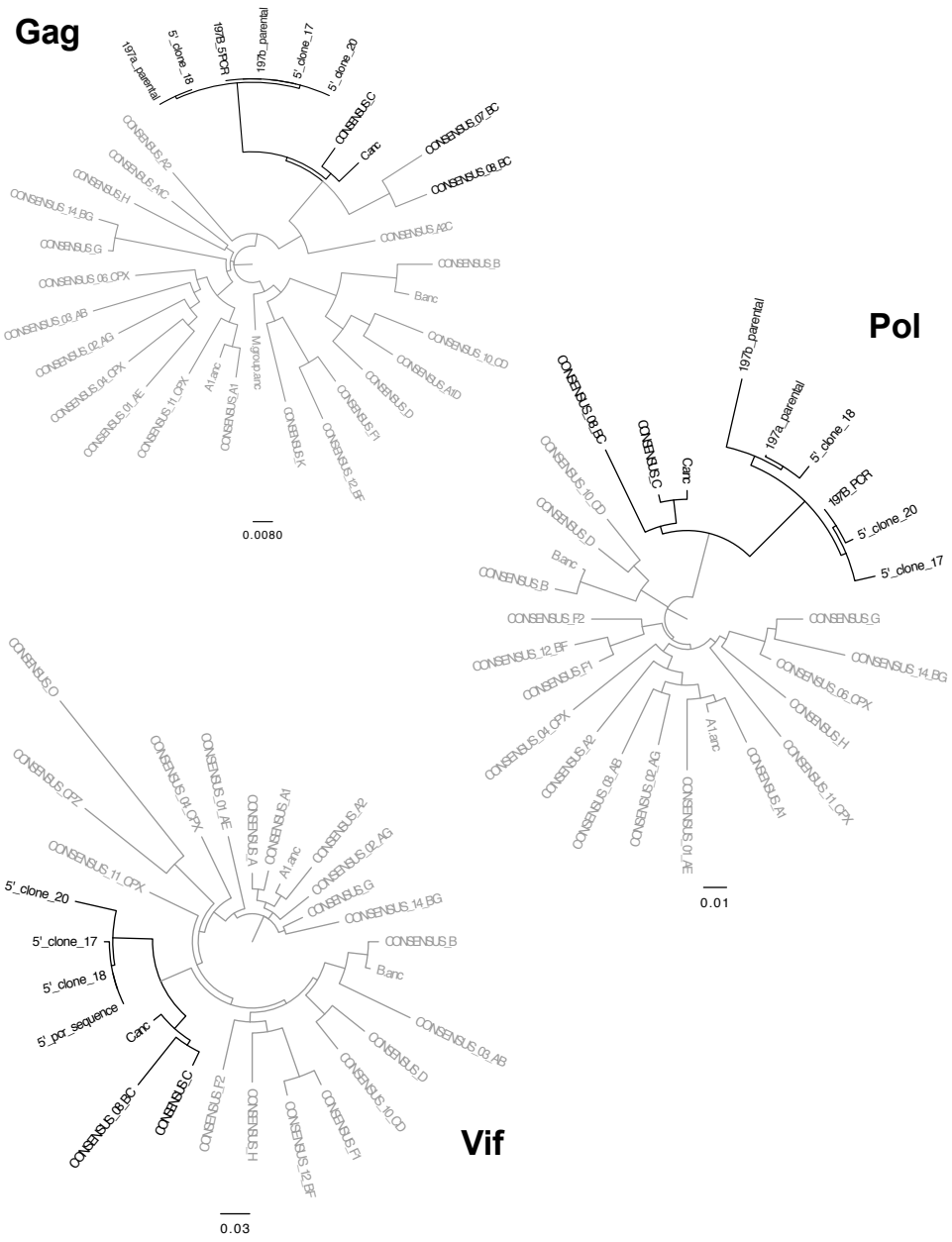


Figure 5.3.3: **Neighbor-joining phylogenetic trees of genes contained within the 5' clones.** Phylogenetic trees for Gag, Pol and Vif are shown. C clade consensus sequence and sequence for the 5' clones, parental and amplified fragment are shown in black. Sequence for other HIV-1 M group subtypes are shown in grey.

Sequence analysis of the overlapping region between 5' and 3' clones showed no shared restriction site that could be use for the construction of full length virus. However the EcoRI restriction site used for the p83-2/ p83-10 plasmid system was found on the 3' clones. We

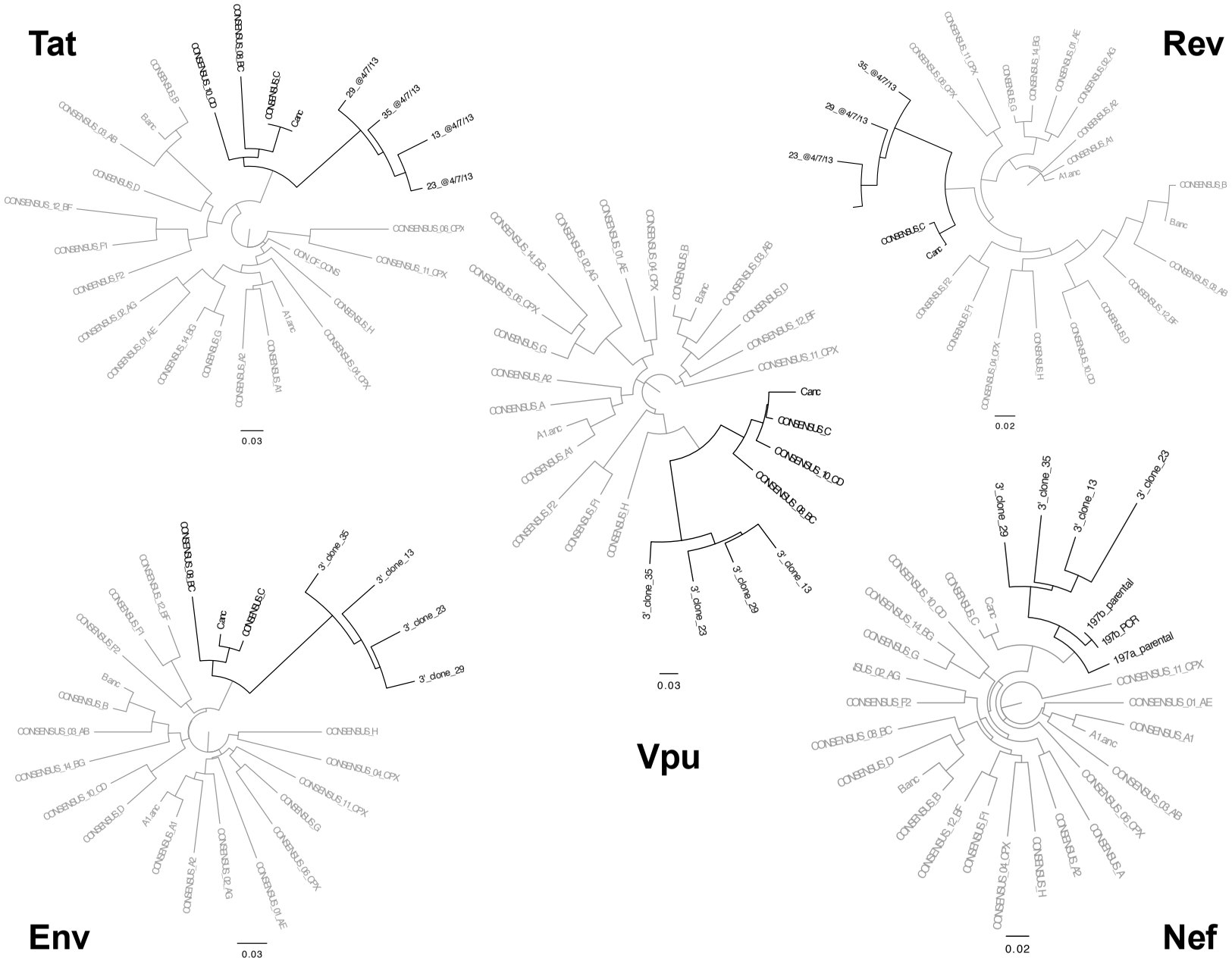


Figure 5.3.4: **Neighbor-joining phylogenetic trees of genes contained within the 3' clones.** Phylogenetic trees for Tat, Rev, Vpu, Env and Nef are shown. C clade consensus sequence and sequence for the 3' clones, parental and amplified fragment are shown in black. Sequence for other HIV-1 M group subtypes are in grey.

therefore decided to introduce the restriction site and missing Vpr sequence to the 5' clones. The successful mutation of the clones is shown in Figure 3.3.5.

Figure 5.3.5: **Sequence alignment of 5' clones after insertion of the EcoRI restriction site to the overlapping Vpr region.** The figure shows the 3' end sequence of the three 5' clones on the first rows, followed by the Vpr C clade consensus sequence and finally the 5' sequence of three 3' clones.

Once the restriction site was confirmed we followed by co-transfection of the 5' and 3' clones into CEM-GXR cell to produce the virus. Plasmids were linearized by EcoRI digestion and all combinations of the 5' and 3' clones were tested. NL4-3 full length plasmid, the p83-2 and p83-10 plasmids and MJ4 were used as positive controls. Unfortunately no combination of the clones produced led to a replication competent virus (Figure 5.3.6).

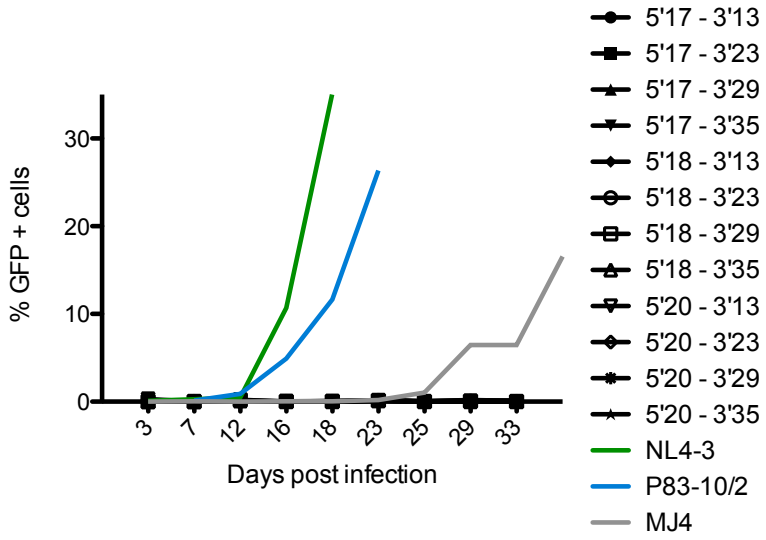


Figure 5.3.6: **Infectivity of CEM-GXR cells by the C clade two-plasmid system.** CEM-GXR cells were co-transfected with all possible combinations of the 5' and 3' clones for viral production and to test infectivity. Viral production was examined by measuring GFP-expressing cells through flowcytometry. The p83-2 and p83-10 plasmid system (shown in blue), the NL4-3 (shown in green) and the MJ4 (shown in grey) plasmids were used as positive controls. No combination of clones was found to be replication competent.

Several genetic features have been associated with lack of viral infectivity. We therefore followed by looking into the clones sequence in more detail. We carried out analysis of the full length sequence utilizing the HIV-1 Sequence Quality Analysis tool from Los Alamos Database. This tools allows for problems with the HIV-1 sequence to be found, including the presence of stop codons, frameshifts, hypermutations as well as providing you with subtype and Blast analysis. The sequence of two the 5' end clones (17 and 18) showed no particular quality problems with gene regions and did not have any obvious mutations. On the other hand the four 3' clones had a large quantity of stop codons, most located within Env (Table 5.2).

Sample	Stop Codons	Frameshifts	Hypermutation
MJ4	1	1	not detected
p83-2	0	1	not detected
5'17	0	1	not detected
5'18	1	0	not detected
5'20	0	3	not detected
p83-10	1	1	not detected
3'13	31	0	possible
3'23	32	0	possible
3'29	30	0	possible
3'35	23	0	possible

Table 5.2: **Quality control analysis of 5' and 3' clone sequences.** The analysis of the full length nucleotide sequence of the 5' and 3' clones produced was analyzed utilizing the HIV Sequence Quality tool from los Alamos Database. Stop codons, frameshifts and detection of hypermutations are shown for each clone and for the MJ4 and p83-2 and p83-10 plasmid system.

Nef polymorphisms or deletions have also been associated with lower VRC. We therefore looked at the Nef sequence of the 3' clones and found that all had a stop codon at position 109.

Other genetic factors, such as escape mutations have been implicated with lower VRC. We therefore looked at previously described escape mutations for which a fitness cost have previously been implicated. We found the Gag I147L, A163N and I247V changes previously shown to have a fitness cost. We also analyzed the Cyp A binding loop sequence to look for compensatory changes at H219, I223 and M228 (Table 5.3), no changes were found.

Sample	ISW9 (Gag 147-155)	KF11 (Gag 162-172)	CypA binding loop (Gag 212-228)	TW10 (Gag 240-249)	QW9 (Gag 308-316)
Consensus C	AISPRTLNAW	KAFSPEVIPMF	DRLHPVHAGPIAPGQM	TSTLQEQIAW	QATQDVKNW
197b parent	-L-------	-N----------	----------	-------V--	----------
5' clone 17	-L-------	-N----------	----------	-------V--	----------
5' clone 18	-L-------	-N----------	----------	-------V--	----------
5' clone 20	-L-------	-N----------	----------	-------V--	----------

Table 5.3: **Variation at HLA-B*57/58 CTL epitopes, and CypA binding loop.**

5.4 Discussion

This chapter aimed to produce an HIV-1 C clade DNA molecular isolate designed in a two-plasmid system to be used as a research tool for biological testing of C clade HIV-1. We chose the

primary virus from a progressor child, known as 197B, to use as template [112]. We successfully amplified and cloned the full length 197B viral genome, from pro-viral DNA extracted into two pCR-XL-TOPO vectors (Figure 5.3.2).

The 5' clones were modified to contain the EcoRI restriction site, which was already found in the 3' clones, and that is used in the p83-2/p83-10 plasmid system to produce full length virus (Figure 5.3.5). Three clones containing the 5' half of the 197B viral genome and four clones containing the 3' half of the 197B viral genome were tested for production of an infectious virus.

Although we found that all genes were enclosed in the clones produced (Figure 5.3.2 and 5.3.3), no combination of 5' and 3' clones were found to produce an infectious virus when tested in CEM-GXR cells (Figure5.3.6). Further analysis revelled a significant number of stop codons in the 3' clones, in particular in the *env* region (Table 5.2). Defects found were shared in the four 3' clones, we therefore concluded that the PCR amplification of the 3' end half needs to be repeated from the proviral DNA. However, it has previously been suggested that defective genomes are frequently found in HIV-1 infected PBMCs [120]. If this is the case, the amplification of the 3' half genome from the same pro-viral DNA may lead to more defective genomes and therefore the infection of healthy PBMCs with 197B virus needs to be repeated for the extraction of new DNA to be used as a template. In this case an alternative modification to our protocol would involve full-length amplification of the 197b genome from DNA instead of amplification in two fragments. This would allow faster testing of several clones for infectivity, followed by the construction of the two-plasmid system of infection-competent clones with the existing EcoRI restriction site.

Alternatively, the defective *env* gene region in our current clones can be replaced with a functional *env* region amplified from a clade C infectious virus using standard cloning techniques to construct an infectious chimeric virus [104].

Despite the genetic alterations present in our clones, our approach for the production of viral stocks can be improved by employing more suitable cell lines such as the COS-1 or 293T cell lines [51, 34], which were unavailable at the time that the project took place. COS-1 and 293T cell lines contain the SV40 Large T-antigen. SV40 allows for the generation of high-titers through stimulation of viral genome replication by direct interaction with DNA polymerase and by driving cells into S phase and thus promoting DNA replication [51, 34].

5.5 Summary comments

This chapter describes the methodology used in the construction of a DNA HIV-1 isolate. Although we were unsuccessful in producing a replication competent virus, the construction of a C clade backbone which replicates efficiently remain an aim due to its importance for research directed to the analysis of genetic determinants of clade C virus, as well as for the comparison of differences in phenotypic characteristics between clades. The data presented suggest that the *env* region of the HIV-1 genome needs to be replaced in order to obtain a replication competent virus, other improvements to the methodology applied are also mentioned for future work.

Chapter 6

Discussion

A central component of the immune response against HIV-1 infection is the HIV-specific $CD8^+$ T cell responses, which can recognize and kill HIV-1 infected $CD4^+$ target cells. Amongst the most valuable clues to understanding natural immune control of HIV-1 are the observed HLA class I associations with HIV-1 disease outcome, which restrict those $CD8^+$ T cell responses.

The studies presented in this thesis have examined the effect of HIV-1 specific $CD8^+$ T cell responses at the individual and population level. The projects were focused on responses restricted by HLA class I alleles and their link to disease outcome through functional and phenotypic analysis of protective immune responses and dynamics of viral escape. The ultimate goal of these studies was to help identify what $CD8^+$ T cell responses a vaccine needs to induce to achieve durable control of HIV-1 infection.

6.1 HLA-B*35 and Immune control

The initial focus of chapter 3 was on the Mexican HIV-1 epidemic. The frequency of HLA class I alleles expressed by the Mexican population are distinct from other populations, with particular high prevalence of HLA-B*35 alleles.

Four-digit analysis has revealed HLA subtype-specific effects that are invisible when only 2-digit typing is studied. For example, in Durban, HLA-B*58:01 is associated with low viral setpoint, whereas B*58:02 is associated with high viral setpoint [68]. The data shown in chapter 3 with regards to differences in markers of disease progression between individuals with different HLA-B*35 subtypes support this observation (Figure 3.3.2).

6.1.1 The Gag hypothesis

Evidence for the importance of targeting Gag in HIV-1 disease outcome has been mounting. Characterization of the $CD8^+$ T cell response in previous studies showed that increasing breadth of the Gag specific response is associated with decreasing viral loads [69]. Additionally a recent study by Boutwell et.al looked at the effect of 29 escape mutations in Gag and found that all carry a fitness cost, although some smaller than others [16]. These data offer an explanation as to why responses to Gag that drive escape are protective. The structural constraints from the conservation of this region make any change detrimental to the virus: the more changes are added, the less fit the virus.

With respect to HLA-B*35, detailed comparative studies of HLA-B*35 and the effect of Gag-specific $CD8^+$ T cell responses restricted by these alleles had not been done. The data presented in chapter 3 suggest a link between HLA-B*35 associated control of HIV-1 and number of Gag epitopes presented (Figure 3.3.6, A).

The discrepancy between disease progression in subjects with HLA-B*35:01 and individuals with HLA-B*35:08 (not associated with disease progression) suggests that HLA-B*35 restricted Gag epitopes may be successfully presented by HLA-B*35:08 but not by HLA-B*35:01. It has been suggested that immune control of C clade HIV-1 by HLA-B*35:01 may be mediated by a single effective response to an epitope in Gag (NY10), for which an escape variant is already in place in B clade virus (E260). We therefore tested the immunogenicity of this NY10-260E peptide when it is presented by HLA-B*35:08. Preliminary data shown here support the protective role of the NY10-specific $CD8^+$ T cell response. We show data from an individual with HLA-B*35:08 who responds to the NY10-260E peptide but does not respond to the NY10-260D peptide (Figure 3.3.9). Additionally a trend for selection of a B clade escape mutation of the NY10 epitope by individuals with B*35:08 was observed: ~10% of individuals without B*35:08 have the E260D (NPPIPVG<u>E</u>IY) variant, compared to 30% of individuals with B*35:08 (Figure 3.3.10).

These data highlight the importance of micropolymorphisms in epitope presentation and its effect on targeting by $CD8^+$ T cells when presented by closely related HLA-B alleles. These data illustrate that small differences, even of a single amino acid, between HLA subtypes can

profoundly influence rate of HIV-1 disease progression within a population.

It is important to note, however, that the functional and phenotypic markers of this NY10 specific $CD8^+$ T cell response were not assessed. The limitation of utilizing IFNγ ELISpot and tetramer staining assays to measure immune responses is that effective responses are not exclusively identified.

Analysis of $CD8^+$ T cell responses by flowcytometry methods have shown that $CD8^+$ T cell responses from LTNP have more polyfunctionality, defined as co expression of CD107a, MIP1β, IFNγ and IL2, upon antigen stimulation compared to responses from progressors [13]. Studies have also shown that rectal mucosal derived from elite controllers and viraemic controllers (VL$<$75 and 75-2000 copies/ml respectively) are more polyfunctional than those from non controllers (VL $>$ 10,000 copies/ml) [42]. Hence particular $CD8^+$ T cell phenotypes, not specific to Gag, may be correlated with immune protection.

Additionally, studies on $CD8^+$ T cell avidity of multiple $CD8^+$ T cell clones specific for the same epitope show that decreasing polyfunctionality correlates with decreasing TCR avidity. Therefore suggesting that avidity also has a role in effective $CD8^+$ T cell responses. However assays used to analyze the $CD8^+$ T cell responses here used supraphysiological concentrations of peptides that do not reflect the physiological levels of peptides processed and presented in vivo [3, 4].

Although proliferation, polyfunctionality and lytic granule content are found to be enhanced in LTNPs it is not clear if these are the cause or effect of low viral antigen. Whilst these characteristics may have an important role in viral control, peptide processing and presentation may be the rate limiting factor in allowing an epitope specific $CD8^+$ T cell response to fulfill its anti-viral potential.

The data presented in this thesis offers a mechanistic explanation for the difference in disease outcome observed between individuals with HLA-B*35:01 or HLA-B*35:08. Previous studies support these data and have also shown that peptide conformation, and not only peptide processing, can affect immunogenicity of an epitope [134]. On the other hand, a main factor controlling the magnitude of $CD8^+$ T cell responses is the immunogenicity of the pMHC complexes presented on the cell surface. However it may also be limited by the TCR repertoire imposed by positive and negative selection. The TCR clonotypes of LTNP have been found to display

a better capacity and higher avidity to cross-recognize naturally occurring viral variants and upregulate perforin and granzyme B than the TCR clonotypes of progressors [119]. But once again, the direction of causality is difficult to asses.

6.2 The role of viral replication capacity as a correlate of disease protection

The extensive diversity of the HIV-1 genome and rapid viral adaptation are the main challenges to vaccine design. Results from the large-scale study carried out by Kawashima, et.al., suggested that WT epitopes may gradually be lost in the course of the epidemic, in particular epitopes that are restricted by common alleles.

Chapter 4 therefore looked at the impact of HLA driven evolution of HIV-1 in VRC at a population level. We hypothesized that the frequency of viral escape mutations with a fitness cost restricted by HLA-B*57, B*58:01 and B*81:01 would be higher in Barbados than Mexico, due to the higher frequency of the restricting alleles in the former population. Results in chapter 4 show that this may be the case as we observe a significant difference in the population viral fitness (Figure 4.3.6). Therefore suggesting that as the epidemic progresses commonly selected escape mutations will accumulate at a population level.

Nevertheless the frequency of reverting mutation has been shown to also correlate with the frequency of the restricting allele [64]. If escape is selected more rapidly than reversion this may explain accumulation, or accumulation may be mediated through the selection of compensatory mutations. In any case, the transmission of escape mutatio ns with a fitness cost to HLA mismatched individuals, may be beneficial for the recipient even if they revert rapidly, as a lower VL set point may be achieved during acute infection.

We observed a decreased viraemia in Barbados (significant lower median VL and significantly higher median $CD4^+$ T cell counts) associated with decreased VRC, which may have occurred due to selection pressure for mutations with a fitness cost by protective alleles occurring in the population over time (Figure 4.3.2). Although RC is significantly lower and markers of disease progression are significantly better in Barbados than in Mexico, when we compare individuals

with HLA-B*57 alleles between the two study cohorts, we observe a trend towards higher median VL and lower median $CD4^+$ T cell counts in Barbados. This suggests that the virus has indeed adapted better to the common HLA-B*57 alleles in Barbados and therefore a loss of protection is observed when compared to the Mexican cohort, in which HLA-B*57 is rare.

The RC of HIV-1 is correlated with disease pathogenesis. Infection with low fitness viruses or the emergence of immune pressure-mediated low fitness viruses is associated with improved control of the VL [94]. It has been suggested, therefore, that vaccine-induced immune responses should aim to focus on vulnerable regions of the virus. These are regions that can not escape without a high fitness cost and with a complex and difficult selection of compensatory mutations. Targeting of such regions will allow for immune protection to remain even through adaptation. Critical $CD8^+$ T cell responses may be lost through accumulation of escape mutations, but not without inducing a high fitness cost that will cripple the virus [54, 136].

The analysis of such mutations is therefore imperative. However there are several deficiencies in the current methods utilized to measure the impact of escape mutations in VRC. It is important to note that the RC method utilized in this thesis does not necessarily equate with viral virulence, as the former assesses the ability of a recombinant virus to replicate in a controlled *in vitro* environment devoid of host or other selection pressures, while the latter reflects the far more complex ability of the virus to cause disease in its host [105].

This study analyzed fitness of B clade virus by the construction of Gag-protease chimeric virus in a B clade backbone. However the analysis of the impact of C clade escape mutations by construction of chimeric virus in a B clade backbone may give false results due to unknown genetic characteristics which may be present in the B clade backbone but are absent from C clade sequence. The production of a replication competent C clade isolate, is therefore imperative and remains an aim for the project presented in chapter 5. Although I was unsuccessful at producing a replication competent virus, defects have been identified and will be tackled.

6.3 Remaining questions

Why are certain HLA alleles associated with better control of HIV-1? Targeting of epitopes in conserved region with structural constraints, leaving HIV-1 unable to escape or

to escape only with a significant fitness cost is thought to contribute to control of the virus. However, the reason why some alleles may target these conserved regions more readily that other alleles is yet to be understood. It is plausible that certain binding motifs may generate more immunogenic pMHC complexes and therefore elicit more effective $CD8^+$ T cell responses. Therefore the different associations between HLA alleles and disease progression may be due to the immunogenic epitopes they are able to present. GWAS support this notion as it has been shown that the SNPs most strongly associated with protection in the Caucasian population lie within the coding region for the HLA-B binding groove. The specific amino acid composition of the protective HLA-B peptide-binding groove may favor an optimal presentation of viral peptides.

What are the advantages of a $CD8^+$ T cell based vaccine? To date HLA class I expression and their restricting HIV-1 specific $CD8^+$ T cell responses have the strongest single impact on the immune control of HIV-1 infection. Therefore a T cell based vaccine holds potential for induced control of HIV-1, especially if combined with appropriate innate and humoral responses.

The RV144 trial gave the first positive results in prevention of HIV-1 acquisition associated with Env Abs. However, no association with $CD8^+$ T cell responses was found in the analysis of correlates of protection. This vaccine may therefore prove ineffective if infected cells start infection and follows cell to cell spread. Additionally, vaccinated individuals that later got infected by HIV-1 did not show reduction in viraemia. It can therefore be argued that the lack of $CD8^+$ T cell responses may have contributed to the lack of significant protection, again suggesting the need for appropriate $CD8^+$ T cell responses to be elicited by an effective vaccine.

Data from a study looking into exposed sero negative individuals who have no neutralizing Abs but exhibit variable levels of HIV-1 specific $CD8^+$ T cells [116] also support the role of cellular immune responses in protection from HIV-1. Nevertheless, $CD8^+$ T cells are only activated once target cells are infected and therefore responses need to lyse the infected cells before progeny virus production. In the SIV model Gag-pecific $CD8^+$ T cell responses have been detected two hours post infection, before integration of viral genome and viral protein synthesis [117]. This very early evolution of the adaptive immune response to HIV-1 capsid protein may

help to prevent the spread and dissemination of HIV-1 and may explain the positive effect of broad Gag specific $CD8^+$ T cell responses in HIV-1 infection.

A therapeutic vaccine that elicits effective $CD8^+$ T cell responses, which constrain the virus, may also prove helpful for infected patients. Such a vaccine could lower viraemia and allow for recovery of $CD4^+$ T cell numbers, as well as providing a public health benefit by reducing chances of onward transmission.

Can a $CD8^+$ T cell vaccine be beneficial for individuals without protective alleles? The data presented here, in support of previous studies, have suggested that protection is associated with targeting of Gag [69]. My data demonstrate that this is also the case for alleles that have been described as having a detrimental effect in HIV-1 disease progression. The data presented in chapter 3 suggest that the breadth of Gag responses is associated with significantly lower median VL and higher median $CD4^+$ T cell counts in individual with HLA-B*35 alleles. These data also suggest that targeting of a single effective Gag-specific (NY10) CD8+ T cell response is associated with control of HIV-1 by HLA-B*35:01 and HLA-B*35:08 alleles. Therefore even 'bad' alleles are capable of inducing effective $CD8^+$ T cell responses. The number of polymorphisms where reversion is inferred is also correlated with lower VL across all HLA type, confirming that disease control is not specific to only a select few alleles. Furthermore, natural immune control of HIV-1 has been shown to occur in the absence of specific protective HLA class I alleles in LTNP; therefore a controller phenotype may be therapeutically inducible in a larger number of patients [136].

Immune selection pressure that promotes outgrowth of viral quasispecies with a significant loss of fitness has emerged as a correlate of immune protection [84, 30]. Thus all HIV-1 infected individuals could potentially benefit from a HIV-1 vaccine designed to alter $CD8^+$ T cell responses to target vulnerable regions of the HIV-1 genome. Identification of such detrimental mutations could be of value for the design of $CD8^+$ T cell vaccines. Nevertheless, further studies on the complex mechanisms involved in escape and reversion rates as well as rate of selection of compensatory mutations are required.

The data presented here and reviewed from previous studies, suggest a complex model of interaction between host and viral genetics in the course of infection that are yet to be fully

understood. Several mechanisms may act together or specific mechanism may prevail in different individuals. The outcome will depend on the fitness of the transmitted virus and genetics of the host involved in regulation of immune activation and effectivity of anti-viral responses.

6.4 Final remarks

Although ART has reduced the rate of mortality and morbidity caused by HIV-1 infection, problems such as drug toxicity, patients adherence, drug resistance and accessibility due to the high cost of ART treatment still remain.

A prophylactic or therapeutic vaccine offers the best hope to restrain the HIV-1 epidemic. However a major challenge is yet to be tackled, the extreme viral sequence variability among strains. Still, LTNPs provide tangible evidence that some humans are able to effectively and durably control HIV-1 replication for extended periods of time [119]. Further understanding of the contribution of genetic factors and of immune responses to viral control is still required.

Although much work remains to be done to achieve an effective $CD8^+$ T based vaccine, hope that induction of HIV-1 control may be possible remains. There is mounting evidence to support that $CD8^+$ T cell responses are important contributors to HIV-1 disease control [55].

The data presented here form part of a larger body of studies, most ongoing at the time of this thesis submission. The work presented adds to the field by providing further understanding of immunological and virological correlates of disease progression. However much remains to be learned about this sophisticated pathogen.

Whilst important research gains have been made through small scientific groups working independently, there is growing recognition for the need of large-scale collaborative studies. The combined effords from the HIV-1 research community worldwide are needed to accelerate the development of an effective HIV-1 vaccine.

Chapter 7

Publications

- Impact of HLA-selection pressure on HIV-1 fitness at a population level. **Claudia I Juarez Molina**, Rebecca Payne, Maribel Soto-Nava, Santiago Avila-Rios, Humberto Valenzuela-Ponce, Emily Adland, Ellen Leitman, Jacqui Brener, Maximilian Muenchhoff, Songee Branch, Gustavo Reyes-Teran, Clive Landis and Philip Goulder. *J. Virol.* Published ahead of print 9 July 2014.

- Impact of HLA-B*35 subtype differences on HIV disease outcome in Mexico. **Claudia Juarez Molina**, Humberto Valenzuela-Ponce, Santiago Avila-Rios, Daniela Garrido-Rodriguez, Thalia Garcia-Tellez, Maribel Soto-Nava, Claudia Garcia-Morales, Philip Goulder and Gustavo Reyes-Teran. *AIDS*. 2014, 28 (11).

- Differential Clade-Specific HLA-B*35:01 Association with HIV-1 Disease Outcome Is Linked to Immunogenicity of a Single Gag Epitope. Philippa C. Matthews, Madoka Koyanagi, Henrik N. Kløverpris, Mikkel Harndahl, Anette Stryhn, Tomohiro Akahoshi, Hiroyuki Gatanaga, Shinichi Oka, **Claudia Juarez Molina**, Humberto Valenzuela Ponce, Santiago Avila Rios, David Cole, Jonathan Carlson, Rebecca P. Payne, Anthony Ogwu, Alfred Bere, Thumbi Ndung'u, Kamini Gounder, Fabian Chen, Lynn Riddell, Graz Luzzi, Roger Shapiro, Christian Brander, Bruce Walker, Andrew K. Sewell, Gustavo Reyes Teran, David Heckerman, Eric Hunter, Søren Buus, Masafumi Takiguchi and Philip J. R. Goulder *J. Virol.* 2012, 86(23).

- Part of the results presented in chapter 3 are in the process of writing to be submitted as a manuscript for peer-reviewed journal publication (*Nature*). The proposed list of co-authors includes: **Claudia I Juarez Molina,** David Cole (equal contribution as first author), Humberto Valenzuela-Ponce, Santiago Avila-Rios, Anna Fuller, Maribel Soto-Nava, Claudia Garcia-Morales, Daniela Garrido-Rodriguez, Gustavo Reyes Teran, Philip Goulder (equal contribution as last author) and Andrew Sewell.

7.0.1 Other publications

- Replicative Capacity of Mother-to-Child transmitted virus is associated with pediatric HIV-1 disease progression rate. Julia G. Prado, Andrew Prendergast, Christina Thobakgale, **Claudia Molina**, Gareth Tudor-Williams, Thumbi Ndung'u, Bruce D. Walker and Philip Goulder. *J. Virol*, 2010, (84)

- Impact of HLA in Mother and Child on Disease Progression of Pediatric Human Immunodeficiency Virus Type 1 Infection. Christina F. Thobakgale, Andrew Prendergast, Hayley Crawford, Nompumelelo Mkhwanazi, Danni Ramduth, Sharon Reddy, **Claudia Molina**, Zenele Mncube, Alasdair Leslie, Julia Prado, Fundi Chonco, Wendy Mphatshwe, Gareth Tudor-Williams, Prakash Jeena, Natasha Blanckenberg, Krista Dong, Photini Kiepiela, Hoosen Coovadia, Thumbi Ndung'u, Bruce D. Walker, and Philip J. R. Goulder. *J. Virol.* 2009 (83).

Bibliography

[1] A Adachi, H E Gendelman, S Koenig, T Folks, R Willey, A Rabson, and M a Martin. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *Journal of virology*, 59(2):284–91, August 1986.

[2] Shizuo Akira, Satoshi Uematsu, and Osamu Takeuchi. Pathogen recognition and innate immunity. *Cell*, 124(4):783–801, February 2006.

[3] Jorge R Almeida, David a Price, Laura Papagno, Zaïna Aït Arkoub, Delphine Sauce, Ethan Bornstein, Tedi E Asher, Assia Samri, Aurélie Schnuriger, Ioannis Theodorou, Dominique Costagliola, Christine Rouzioux, Henri Agut, Anne-Geneviève Marcelin, Daniel Douek, Brigitte Autran, and Victor Appay. Superior control of HIV-1 replication by CD8+ T cells is reflected by their avidity, polyfunctionality, and clonal turnover. *The Journal of experimental medicine*, 204(10):2473–85, October 2007.

[4] Jorge R Almeida, Delphine Sauce, David a Price, Laura Papagno, So Youn Shin, Arnaud Moris, Martin Larsen, Gianfranco Pancino, Daniel C Douek, Brigitte Autran, Asier Sáez-Cirión, and Victor Appay. Antigen sensitivity is a major determinant of CD8+ T-cell polyfunctionality and HIV-suppressive activity. *Blood*, 113(25):6351–60, June 2009.

[5] Julia K Archbold, Whitney a Macdonald, Stephanie Gras, Lauren K Ely, John J Miles, Melissa J Bell, Rebekah M Brennan, Travis Beddoe, Matthew C J Wilce, Craig S Clements, Anthony W Purcell, James McCluskey, Scott R Burrows, and Jamie Rossjohn. Natural

micropolymorphism in human leukocyte antigens provides a basis for genetic control of antigen recognition. *The Journal of experimental medicine*, 206(1):209–19, January 2009.

[6] Haim Ashkenazy, Osnat Penn, Adi Doron-Faigenboim, Ofir Cohen, Gina Cannarozzi, Oren Zomer, and Tal Pupko. FastML: a web server for probabilistic reconstruction of ancestral sequences. *Nucleic acids research*, 40(Web Server issue):W580–4, July 2012.

[7] T. Baba, Y. Jeong, D Pennick, R Bronson, M. Greene, and R. Ruprecht. Pathogenicity of live, attenuated SIV after mucosal infection of neonatal macaques. *Science*, 267(5205):1820–1825, March 1995.

[8] Daniel L Barber, E John Wherry, David Masopust, Baogong Zhu, James P Allison, Arlene H Sharpe, Gordon J Freeman, and Rafi Ahmed. Restoring function in exhausted CD8 T cells during chronic viral infection. *Nature*, 439(7077):682–7, February 2006.

[9] W M Barnes. The fidelity of Taq polymerase catalyzing PCR is improved by an N-terminal deletion. *Gene*, 112(1):29–35, 1992.

[10] Dan H Barouch, Kara L O'Brien, Nathaniel L Simmons, Sharon L King, Peter Abbink, Lori F Maxfield, Ying-Hua Sun, Annalena La Porte, Ambryice M Riggs, Diana M Lynch, Sarah L Clark, Katherine Backus, James R Perry, Michael S Seaman, Angela Carville, Keith G Mansfield, James J Szinger, Will Fischer, Mark Muldoon, and Bette Korber. Mosaic HIV-1 vaccines expand the breadth and depth of cellular immune responses in rhesus monkeys. *Nature medicine*, 16(3):319–23, March 2010.

[11] Rodrigo Barquera, Joaquín Zúñiga, Raquel Hernández-Díaz, Victor Acuña Alonzo, Karla Montoya-Gama, Juan Moscoso, Diana Torres-García, Claudia García-Salas, Beatriz Silva, David Cruz-Robles, Antonio Arnaiz-Villena, Gilberto Vargas-Alarcón, and Julio Granados. HLA class I and class II haplotypes in admixed families from several regions of Mexico. *Molecular immunology*, 45(4):1171–8, February 2008.

[12] Phillip W. Berman, Timothy J. Gregory, Lavon Riddle, Gerald R. Nakamura, Mark A. Champe, James P. Porter, Florain M. Wurm, Robert D. Hershberg, E. Kathy Cobb, and

Jorg W. Eichberg. Protection of chimpanzees from infection by HIV-1 after vaccination with recombinant glycoprotein gp120 but not gp160. 345(6276):622–625, June 1990.

[13] Michael R Betts, Martha C Nason, Sadie M West, Stephen C De Rosa, Stephen a Migueles, Jonathan Abraham, Michael M Lederman, Jose M Benito, Paul a Goepfert, Mark Connors, Mario Roederer, and Richard a Koup. HIV nonprogressors preferentially maintain highly functional HIV-specific CD8+ T cells. *Blood*, 107(12):4781–9, June 2006.

[14] Tanmoy Bhattacharya, Marcus Daniels, David Heckerman, Brian Foley, Nicole Frahm, Carl Kadie, Jonathan Carlson, Karina Yusim, Ben McMahon, Brian Gaschen, Simon Mallal, James I Mullins, David C Nickle, Joshua Herbeck, Christine Rousseau, Gerald H Learn, Toshiyuki Miura, Christian Brander, Bruce Walker, and Bette Korber. Founder effects in the assessment of HIV polymorphisms and HLA allele associations. *Science (New York, N.Y.)*, 315(5818):1583–6, March 2007.

[15] P Borrow, H Lewicki, B H Hahn, G M Shaw, and M B Oldstone. Virus-specific CD8+ cytotoxic T-lymphocyte activity associated with control of viremia in primary human immunodeficiency virus type 1 infection. *Journal of virology*, 68(9):6103–10, September 1994.

[16] Christian L Boutwell, Jonathan M Carlson, Tien-ho Lin, Aaron Seese, Karen A Power, Jian Peng, Yanhua Tang, Zabrina L Brumme, David Heckerman, Arne Schneidewind, and Todd M Allen. Fitness Costs in the Human Immunodeficiency Virus Type 1 Subtype. *Journal of virology*, 87(7):3952–3965, 2013.

[17] Jason M Brenchley, Timothy W Schacker, Laura E Ruff, David a Price, Jodie H Taylor, Gregory J Beilman, Phuong L Nguyen, Alexander Khoruts, Matthew Larson, Ashley T Haase, and Daniel C Douek. CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. *The Journal of experimental medicine*, 200(6):749–59, September 2004.

[18] Mark A Brockman, Arne Schneidewind, Matthew Lahaie, Aaron Schmidt, Toshiyuki Miura, Ivna Desouza, Faina Ryvkin, Cynthia a Derdeyn, Susan Allen, Eric Hunter, Joseph Mulenga, Paul a Goepfert, Bruce D Walker, and Todd M Allen. 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. *Journal of virology*, 81(22):12608–18, November 2007.

[19] Jacqueline M Burrows, Katherine K Wynn, Fleur E Tynan, Julia Archbold, John J Miles, Melissa J Bell, Rebekah M Brennan, Susan Walker, James McCluskey, Jamie Rossjohn, Rajiv Khanna, and Scott R Burrows. The impact of HLA-B micropolymorphism outside primary peptide anchor pockets on the CTL response to CMV. *European journal of immunology*, 37(4):946–53, April 2007.

[20] Pedro Cano, William Klitz, Steven J Mack, Martin Maiers, Steven G E Marsh, Harriet Noreen, Elaine F Reed, David Senitzer, Michelle Setterholm, Anajane Smith, and Marcelo Fernández-Viña. Common and well-documented HLA alleles: report of the Ad-Hoc committee of the american society for histocompatiblity and immunogenetics. *Human immunology*, 68(5):392–417, May 2007.

[21] Jonathan M Carlson, Zabrina L Brumme, Christine M Rousseau, Chanson J Brumme, Philippa Matthews, Carl Kadie, James I Mullins, Bruce D Walker, P Richard Harrigan, Philip J R Goulder, and David Heckerman. Phylogenetic dependency networks: inferring patterns of CTL escape and codon covariation in HIV-1 Gag. *PLoS computational biology*, 4(11):e1000225, November 2008.

[22] M. Carrington. HLA and HIV-1: Heterozygote Advantage and B*35-Cw*04 Disadvantage. *Science*, 283(5408):1748–1752, March 1999.

[23] Mary Carrington and Stephen J O'Brien. The influence of HLA genotype on AIDS. *Annual review of medicine*, 54:535–51, January 2003.

[24] Huabiao Chen, Chun Li, Jinghe Huang, Thai Cung, Katherine Seiss, Jill Beamon, Mary F Carrington, Lindsay C Porter, Patrick S Burke, Yue Yang, Bethany J Ryan, Ruiwu Liu, Robert H Weiss, Florencia Pereyra, William D Cress, Abraham L Brass, Eric S Rosenberg, Bruce D Walker, Xu G Yu, and Mathias Lichterfeld. CD4+ T cells from elite controllers resist HIV-1 infection by selective upregulation of p21. *The Journal of clinical investigation*, 121(4):1549–60, April 2011.

[25] Huabiao Chen, Zaza M Ndhlovu, Dongfang Liu, Lindsay C Porter, Justin W Fang, Sam Darko, Mark A Brockman, Toshiyuki Miura, Zabrina L Brumme, Arne Schneidewind, Alicja Piechocka-Trocha, Kevin T Cesa, Jennifer Sela, Thai D Cung, Ildiko Toth, Florencia Pereyra, Xu G Yu, Daniel C Douek, Daniel E Kaufmann, Todd M Allen, and Bruce D Walker. TCR clonotypes modulate the protective effect of HLA class I molecules in HIV-1 infection. *Nature immunology*, 13(7):691–700, July 2012.

[26] Nicolas Chomont, Mohamed El-Far, Petronela Ancuta, Lydie Trautmann, Francesco a Procopio, Bader Yassine-Diab, Geneviève Boucher, Mohamed-Rachid Boulassel, Georges Ghattas, Jason M Brenchley, Timothy W Schacker, Brenna J Hill, Daniel C Douek, Jean-Pierre Routy, Elias K Haddad, and Rafick-Pierre Sékaly. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nature medicine*, 15(8):893–900, August 2009.

[27] David K Cole, Emily S J Edwards, Katherine K Wynn, Mathew Clement, John J Miles, Kristin Ladell, Julia Ekeruche, Emma Gostick, Katherine J Adams, Ania Skowera, Mark Peakman, Linda Wooldridge, David a Price, and Andrew K Sewell. Modification of MHC anchor residues generates heteroclitic peptides that alter TCR binding and T cell recognition. *Journal of immunology (Baltimore, Md. : 1950)*, 185(4):2600–10, August 2010.

[28] Christopher M Coleman and D Ph. HIV Interactions with Dendritic Cells. 762:1–20, 2013.

[29] Hayley Crawford, Wendy Lumm, Alasdair Leslie, Malinda Schaefer, Debrah Boeras, Julia G Prado, Jianming Tang, Paul Farmer, Thumbi Ndung'u, Shabir Lakhi, Jill Gilmour, Paul Goepfert, Bruce D Walker, Richard Kaslow, Joseph Mulenga, Susan Allen, Philip J R Goulder, and Eric Hunter. Evolution of HLA-B*5703 HIV-1 escape mutations in HLA-B*5703-positive individuals and their transmission recipients. *The Journal of experimental medicine*, 206(4):909–21, April 2009.

[30] Hayley Crawford, Julia G Prado, Alasdair Leslie, Stéphane Hué, Isobella Honeyborne, Sharon Reddy, Mary van der Stok, Zenele Mncube, Christian Brander, Christine Rousseau, James I Mullins, Richard Kaslow, Paul Goepfert, Susan Allen, Eric Hunter, Joseph Mulenga, Photini Kiepiela, Bruce D Walker, and Philip J R Goulder. 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. *Journal of virology*, 81(15):8346–51, August 2007.

[31] D. Persaud, H. Gay, C. Ziemniak, Y.H. Chen, M. Piatak, T-W. Chun, M. Strain, D. Richman and K. Luzuriaga. *Conference on Retroviruses Opportunistic infections*. 2013.

[32] Peter C. Doherty and Jan P. Christensen. Accessing Complexity: The Dynamics of Virus-Specific T Cell Responses. November 2003.

[33] Daniel C Douek, Mario Roederer, and Richard a Koup. Emerging concepts in the immunopathogenesis of AIDS. *Annual review of medicine*, 60(December 2008):471–84, January 2009.

[34] R B DuBridge, P Tang, H C Hsia, P M Leong, J H Miller, and M P Calos. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Molecular and cellular biology*, 7(1):379–87, January 1987.

[35] B Dupont. "Phototyping" for HLA: the beginning of the end of HLA typing as we know it. *Tissue antigens*, 46(5):353–4, November 1995.

[36] BH Edwards, Anju Bansal, and Steffanie Sabbaj. Magnitude of Functional CD8+ T-Cell Responses to the Gag Protein of Human Immunodeficiency Virus Type 1 Correlates Inversely with Viral Load in Plasma. *Journal of . . .*, 2002.

[37] M. A. Egan, W. A. Charini, M. J. Kuroda, J. E. Schmitz, P. Racz, K. Tenner-Racz, K. Manson, M. Wyand, M. A. Lifton, C. E. Nickerson, T. Fu, J. W. Shiver, and N. L. Letvin. Simian Immunodeficiency Virus (SIV) gag DNA-Vaccinated Rhesus Monkeys Develop Secondary Cytotoxic T-Lymphocyte Responses and Control Viral Replication after Pathogenic SIV Infection. *Journal of Virology*, 74(16):7485–7495, August 2000.

[38] T G Evans, M C Keefer, K J Weinhold, M Wolff, D Montefiori, G J Gorse, B S Graham, M J McElrath, M L Clements-Mann, M J Mulligan, P Fast, M C Walker, J L Excler, A M Duliege, and J Tartaglia. A canarypox vaccine expressing multiple human immunodeficiency virus type 1 genes given alone or with rgp120 elicits broad and durable CD8+

cytotoxic T lymphocyte responses in seronegative volunteers. *The Journal of infectious diseases*, 180(2):290–8, August 1999.

[39] ME Feeney, Y Tang, KA Roosevelt, AJ Leslie, K. McIntosh, N Karthas, BD Walker, and PJR Goulder. 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. *Journal of virology*, 78(16):8927, 2004.

[40] Jacques Fellay, Dongliang Ge, Kevin V Shianna, Sara Colombo, Bruno Ledergerber, Elizabeth T Cirulli, Thomas J Urban, Kunlin Zhang, Curtis E Gumbs, Jason P Smith, Antonella Castagna, Alessandro Cozzi-Lepri, Andrea De Luca, Philippa Easterbrook, Huldrych F Günthard, Simon Mallal, Cristina Mussini, Judith Dalmau, Javier Martinez-Picado, José M Miro, Niels Obel, Steven M Wolinsky, Jeremy J Martinson, Roger Detels, Joseph B Margolick, Lisa P Jacobson, Patrick Descombes, Stylianos E Antonarakis, Jacques S Beckmann, Stephen J O'Brien, Norman L Letvin, Andrew J McMichael, Barton F Haynes, Mary Carrington, Sheng Feng, Amalio Telenti, and David B Goldstein. Common genetic variation and the control of HIV-1 in humans. *PLoS genetics*, 5(12):e1000791, December 2009.

[41] Jacques Fellay, Kevin V Shianna, Dongliang Ge, Sara Colombo, Bruno Ledergerber, Mike Weale, Kunlin Zhang, Curtis Gumbs, Antonella Castagna, Andrea Cossarizza, Alessandro Cozzi-Lepri, Andrea De Luca, Philippa Easterbrook, Patrick Francioli, Simon Mallal, Javier Martinez-Picado, José M Miro, Niels Obel, Jason P Smith, Josiane Wyniger, Patrick Descombes, Stylianos E Antonarakis, Norman L Letvin, Andrew J McMichael, Barton F Haynes, Amalio Telenti, and David B Goldstein. A whole-genome association study of major determinants for host control of HIV-1. *Science (New York, N.Y.)*, 317(5840):944–7, August 2007.

[42] April L Ferre, Peter W Hunt, J William Critchfield, Delandy H Young, Megan M Morris, Juan C Garcia, Richard B Pollard, Hal F Yee, Jeffrey N Martin, Steven G Deeks, and Barbara L Shacklett. Mucosal immune responses to HIV-1 in elite controllers: a potential correlate of immune control. *Blood*, 113(17):3978–89, April 2009.

[43] Will Fischer, Simon Perkins, James Theiler, Tanmoy Bhattacharya, Karina Yusim, Robert Funkhouser, Carla Kuiken, Barton Haynes, Norman L Letvin, Bruce D Walker, Beatrice H Hahn, and Bette T Korber. Polyvalent vaccines for optimal coverage of potential T-cell epitopes in global HIV-1 variants. *Nature Medicine*, 13(1):100–106, December 2006.

[44] A J Frater, C T 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. Passive sexual transmission of human immunodeficiency virus type 1 variants and adaptation in new hosts. *Journal of virology*, 80(14):7226–34, July 2006.

[45] J Frater, H Brown, A Oxenius, H F Günthard, B Hirschel, N Robinson, a J Leslie, R Payne, H Crawford, A Prendergast, C Brander, P Kiepiela, B D Walker, P J R Goulder, A McLean, and R E Phillips. Effective T-cell responses select human immunodeficiency virus mutants and slow disease progression. *Journal of virology*, 81(12):6742–51, June 2007.

[46] M. J. Fuller, A. Khanolkar, A. E. Tebo, and A. J. Zajac. Maintenance, Loss, and Resurgence of T Cell Responses During Acute, Protracted, and Chronic Viral Infections. *The Journal of Immunology*, 172(7):4204–4214, March 2004.

[47] X Gao, 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. Effect of a single amino acid change in MHC class I molecules on the rate of progression to AIDS. *The New England journal of medicine*, 344(22):1668–75, May 2001.

[48] M Geiser, R Cèbe, D Drewello, and R Schmitz. Integration of PCR Fragments at Any Specific Site within Cloning Vectors without the Use of Restriction Enzymes and DNA Ligase. *Biotechniques*, 2001.

[49] Christof Geldmacher, Ian S Metzler, Sodsai Tovanabutra, Tedi E Asher, Emma Gostick, David R Ambrozak, Constantinos Petrovas, Alexandra Schuetz, Njabulo Ngwenyama, Gustavo Kijak, Leonard Maboko, Michael Hoelscher, Francine McCutchan, David a Price, Daniel C Douek, and Richard a Koup. Minor viral and host genetic polymorphisms

can dramatically impact the biologic outcome of an epitope-specific CD8 T-cell response. *Blood*, 114(8):1553–62, August 2009.

[50] Ronald B Geskus, Maria Prins, Jean-Baptiste Hubert, Frank Miedema, Ben Berkhout, Christine Rouzioux, Jean-Francois Delfraissy, and Laurence Meyer. The HIV RNA set-point theory revisited. *Retrovirology*, 4:65, January 2007.

[51] Yakov Gluzman. SV40-transformed simian cells support the replication of early SV40 mutants. *Cell*, 23(1):175–182, January 1981.

[52] Paul a Goepfert, Wendy Lumm, Paul Farmer, Philippa Matthews, Andrew Prendergast, Jonathan M Carlson, Cynthia a Derdeyn, Jianming Tang, Richard a Kaslow, Anju Bansal, Karina Yusim, David Heckerman, Joseph Mulenga, Susan Allen, Philip J R Goulder, and Eric Hunter. Transmission of HIV-1 Gag immune escape mutations is associated with reduced viral load in linked recipients. *The Journal of experimental medicine*, 205(5):1009–17, May 2008.

[53] P J Goulder, C Brander, Y Tang, C Tremblay, R a Colbert, M M Addo, E S Rosenberg, T Nguyen, R Allen, a Trocha, M Altfeld, S He, M Bunce, R Funkhouser, S I Pelton, S K Burchett, K McIntosh, B T Korber, and B D Walker. Evolution and transmission of stable CTL escape mutations in HIV infection. *Nature*, 412(6844):334–8, July 2001.

[54] Philip J R Goulder and David I Watkins. HIV and SIV CTL escape: implications for vaccine design. *Nature reviews. Immunology*, 4(8):630–40, August 2004.

[55] Philip J R Goulder and David I Watkins. Impact of MHC class I diversity on immune control of immunodeficiency virus replication. *Nature reviews. Immunology*, 8(8):619–30, August 2008.

[56] P.J.R. Goulder, R.E. Phillips, R.A. Colbert, S. McAdam, G. Ogg, M.A. Nowak, P. Giangrande, G. Luzzi, B. Morgan, A. Edwards, and Others. late escape from an immunodominant cytotoxic T lymphocyte response associated with progression to AIDS. *Nature medicine*, 3(2):212–217, 1997.

[57] Kate J Green, John J Miles, Judy Tellam, Wendy J M van Zuylen, Geoff Connolly, and Scott R Burrows. Potent T cell response to a class I-binding 13-mer viral epitope and the influence of HLA micropolymorphism in controlling epitope length. *European journal of immunology*, 34(9):2510–9, September 2004.

[58] R.D. Grisson, A.L. Chenine, L.Y. Yeh, Jun He, Charles Wood, G.J. Bhat, Weidong Xu, Chipepo Kankasa, and R.M. Ruprecht. Infectious molecular clone of a recently transmitted pediatric human immunodeficiency virus clade C isolate from Africa: evidence of intraclade recombination. *Journal of virology*, 78(24):14066–14069, 2004.

[59] Scott G Hansen, Julia C Ford, Matthew S Lewis, Abigail B Ventura, Colette M Hughes, Lia Coyne-Johnson, Nathan Whizin, Kelli Oswald, Rebecca Shoemaker, Tonya Swanson, Alfred W Legasse, Maria J Chiuchiolo, Christopher L Parks, Michael K Axthelm, Jay a Nelson, Michael a Jarvis, Michael Piatak, Jeffrey D Lifson, and Louis J Picker. Profound early control of highly pathogenic SIV by an effector memory T-cell vaccine. *Nature*, 473(7348):523–7, May 2011.

[60] CM Hogan and SM Hammer. Host Determinants in HIV Infection and Disease: Part 1: Cellular and Humoral Immune Responses. *Annals of internal medicine*, pages 761–776, 2001.

[61] Jinghe Huang, Patrick S Burke, Thai Duong Hong Cung, Florencia Pereyra, Ildiko Toth, Bruce D Walker, Luis Borges, Mathias Lichterfeld, and Xu G Yu. Leukocyte immunoglobulin-like receptors maintain unique antigen-presenting properties of circulating myeloid dendritic cells in HIV-1-infected elite controllers. *Journal of virology*, 84(18):9463–71, September 2010.

[62] Jinghe Huang, James J Goedert, Eric J Sundberg, Thai Duong Hong Cung, Patrick S Burke, Maureen P Martin, Liliana Preiss, Jeffrey Lifson, Mathias Lichterfeld, Mary Carrington, and Xu G Yu. HLA-B*35-Px-mediated acceleration of HIV-1 infection by increased inhibitory immunoregulatory impulses. *The Journal of experimental medicine*, 206(13):2959–66, December 2009.

[63] Xia Jin, Xiaojiang Gao, M. Ramanathan Jr, G.R. Deschenes, G.W. Nelson, S.J. O'Brien, J.J. Goedert, D.D. Ho, T.R. O'Brien, and Mary Carrington. Human Immunodeficiency Virus Type 1 specific CD8+ T cell responses for groups of HIV 1 infected individuals with different HLA B 35 genotypes. *Journal of virology*, 76(24):12603, 2002.

[64] Yuka Kawashima, Katja Pfafferott, John Frater, Philippa Matthews, Rebecca Payne, Marylyn Addo, Hiroyuki Gatanaga, Mamoru Fujiwara, Atsuko Hachiya, Hirokazu Koizumi, Nozomi Kuse, Shinichi Oka, Anna Duda, Andrew Prendergast, Hayley Crawford, Alasdair Leslie, Zabrina Brumme, Chanson Brumme, Todd Allen, Christian Brander, Richard Kaslow, James Tang, Eric Hunter, Susan Allen, Joseph Mulenga, Songee Branch, Tim Roach, Mina John, Simon Mallal, Anthony Ogwu, Roger Shapiro, Julia G Prado, Sarah Fidler, Jonathan Weber, Oliver G Pybus, Paul Klenerman, Thumbi Ndung, Rodney Phillips, David Heckerman, P Richard Harrigan, Bruce D Walker, Masafumi Takiguchi, Philip Goulder, and Thumbi Ndung'u. Adaptation of HIV-1 to human leukocyte antigen class I. *Nature*, 458(7238):641–645, April 2009.

[65] Brandon F Keele, Elena E Giorgi, Jesus F Salazar-Gonzalez, Julie M Decker, Kimmy T Pham, Maria G Salazar, Chuanxi Sun, Truman Grayson, Shuyi Wang, Hui Li, Xiping Wei, Chunlai Jiang, Jennifer L Kirchherr, Feng Gao, Jeffery a Anderson, Li-Hua Ping, Ronald Swanstrom, Georgia D Tomaras, William a Blattner, Paul a Goepfert, J Michael Kilby, Michael S Saag, Eric L Delwart, Michael P Busch, Myron S Cohen, David C Montefiori, Barton F Haynes, Brian Gaschen, Gayathri S Athreya, Ha Y Lee, Natasha Wood, Cathal Seoighe, Alan S Perelson, Tanmoy Bhattacharya, Bette T Korber, Beatrice H Hahn, and George M Shaw. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proceedings of the National Academy of Sciences of the United States of America*, 105(21):7552–7, May 2008.

[66] A D Kelleher, 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. Clustered mutations in HIV-1 gag are consistently required for escape from HLA-B27-restricted cytotoxic T lymphocyte responses. *The Journal of experimental medicine*, 193(3):375–86, February 2001.

[67] Peter Kelleher. HIV immunology: what lessons can we learn from recent vaccine trials? *Journal of HIV therapy*, 15(1):2–5, March 2010.

[68] Photini Kiepiela, Alasdair J Leslie, Isobella Honeyborne, Danni Ramduth, Christina Thobakgale, Senica Chetty, Prinisha Rathnavalu, Corey Moore, Katja J Pfafferott, Louise Hilton, Peter Zimbwa, Sarah Moore, Todd Allen, Christian Brander, Marylyn M Addo, Marcus Altfeld, Ian James, Simon Mallal, Michael Bunce, Linda D Barber, James Szinger, Cheryl Day, Paul Klenerman, James Mullins, Bette Korber, Hoosen M Coovadia, Bruce D Walker, and Philip J R Goulder. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. *Nature*, 432(7018):769–75, December 2004.

[69] Photini Kiepiela, Kholiswa Ngumbela, Christina Thobakgale, Dhanwanthie Ramduth, Isobella Honeyborne, Eshia Moodley, Shabashini Reddy, Chantal de Pierres, Zenele Mncube, Nompumelelo Mkhwanazi, Karen Bishop, Mary van der Stok, Kriebashnie Nair, Nasreen Khan, Hayley Crawford, Rebecca Payne, Alasdair Leslie, Julia Prado, Andrew Prendergast, John Frater, Noel McCarthy, Christian Brander, Gerald H Learn, David Nickle, Christine Rousseau, Hoosen Coovadia, James I Mullins, David Heckerman, Bruce D Walker, and Philip Goulder. CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. *Nature medicine*, 13(1):46–53, January 2007.

[70] Henrik N Kloverpris, Anette Stryhn, Mikkel Harndahl, Mary van der Stok, Rebecca P Payne, Philippa C Matthews, Fabian Chen, Lynn Riddell, Bruce D Walker, Thumbi Ndung'u, Sø ren Buus, and Philip Goulder. HLA-B*57 micropolymorphism shapes HLA allele-specific epitope immunogenicity, selection pressure and HIV immune control. *Journal of virology*, 824(November), November 2011.

[71] R A Koup, J T Safrit, Y Cao, C a Andrews, G McLeod, W Borkowsky, C Farthing, and D D Ho. Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. *Journal of virology*, 68(7):4650–5, July 1994.

[72] Smita Kulkarni, Maureen P Martin, and Mary Carrington. The Yin and Yang of HLA and KIR in human disease. *Seminars in immunology*, 20(6):343–52, December 2008.

[73] Peter D Kwong, Michael L Doyle, David J Casper, Claudia Cicala, Stephanie A Leavitt, Shahzad Majeed, Tavis D Steenbeke, Miro Venturi, Irwin Chaiken, Michael Fung, Hermann Katinger, Paul W I H Parren, James Robinson, Donald Van Ryk, Liping Wang, Dennis R Burton, Ernesto Freire, Richard Wyatt, Joseph Sodroski, Wayne A Hendrickson, and James Arthos. HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites. *Nature*, 420(6916):678–82, December 2002.

[74] A J Leslie, 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 R Goulder. HIV evolution: CTL escape mutation and reversion after transmission. *Nature medicine*, 10(3):282–9, March 2004.

[75] Alasdair Leslie, Daniel Kavanagh, Isobella Honeyborne, Katja Pfafferott, Charles Edwards, Tilly Pillay, Louise Hilton, Christina Thobakgale, Danni Ramduth, Rika Draenert, Sylvie Le Gall, Graz Luzzi, Anne Edwards, Christian Brander, Andrew K Sewell, Sarah Moore, James Mullins, Corey Moore, Simon Mallal, Nina Bhardwaj, Karina Yusim, Rodney Phillips, Paul Klenerman, Bette Korber, Photini Kiepiela, Bruce Walker, and Philip Goulder. Transmission and accumulation of CTL escape variants drive negative associations between HIV polymorphisms and HLA. *The Journal of experimental medicine*, 201(6):891–902, March 2005.

[76] Alasdair Leslie, Philippa C Matthews, Jennifer Listgarten, Jonathan M Carlson, Carl Kadie, Thumbi Ndung'u, Christian Brander, Hoosen Coovadia, Bruce D Walker, David Heckerman, and Philip J R Goulder. Additive contribution of HLA class I alleles in the immune control of HIV-1 infection. *Journal of virology*, 84(19):9879–88, October 2010.

[77] Norman L Letvin. Strategies for an HIV vaccine. *The Journal of clinical investigation*, 110(1):15–20, July 2002.

[78] JA Levy. *HIV and the Pathogenesis of AIDS*. 2007.

[79] Hui Li, Katharine J Bar, Shuyi Wang, Julie M Decker, Yalu Chen, Chuanxi Sun, Jesus F Salazar-Gonzalez, Maria G Salazar, Gerald H Learn, Charity J Morgan, Joseph E Schumacher, Peter Hraber, Elena E Giorgi, Tanmoy Bhattacharya, Bette T Korber, Alan S Perelson, Joseph J Eron, Myron S Cohen, Charles B Hicks, Barton F Haynes, Martin Markowitz, Brandon F Keele, Beatrice H Hahn, and George M Shaw. High Multiplicity Infection by HIV-1 in Men Who Have Sex with Men. *PLoS pathogens*, 6(5):e1000890, May 2010.

[80] R Liu, W a Paxton, S Choe, D Ceradini, S R Martin, R Horuk, M E MacDonald, H Stuhlmann, R a Koup, and N R Landau. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell*, 86(3):367–77, August 1996.

[81] K S Lundberg, D D Shoemaker, M W W Adams, J M Short, J A Sorge, and E J Mathur. High-Fidelity Amplification Using a Thermostable DNA-Polymerase Isolated from Pyrococcus-Furiosus. *Gene*, 108(1):1–6, 1991.

[82] Steven G.E. Parham Peter Marsh and Barber Linda D. *The HLA factsbook*. London, 2000.

[83] Maureen P Martin and Mary Carrington. Immunogenetics of HIV disease. *Immunological reviews*, 254(1):245–64, July 2013.

[84] J. Martinez-Picado, J.G. Prado, E.E. Fry, Katja Pfafferott, Alasdair Leslie, Senica Chetty, Christina Thobakgale, Isobel Honeyborne, Hayley Crawford, Philippa Matthews, and Others. Fitness cost of escape mutations in p24 Gag in association with control of human immunodeficiency virus type 1. *Journal of virology*, 80(7):3617, 2006.

[85] John R Mascola and Barton F Haynes. HIV-1 neutralizing antibodies: understanding nature's pathways. *Immunological reviews*, 254(1):225–44, July 2013.

[86] Philippa C Matthews, Madoka Koyanagi, Henrik N Klø verpris, Mikkel Harndahl, Anette Stryhn, Tomohiro Akahoshi, Hiroyuki Gatanaga, Shinichi Oka, Claudia Juarez Molina,

Humberto Valenzuela Ponce, Santiago Avila Rios, David Cole, Jonathan Carlson, Rebecca P Payne, Anthony Ogwu, Alfred Bere, Thumbi Ndung'u, Kamini Gounder, Fabian Chen, Lynn Riddell, Graz Luzzi, Roger Shapiro, Christian Brander, Bruce Walker, Andrew K Sewell, Gustavo Reyes Teran, David Heckerman, Eric Hunter, Sø ren Buus, Masafumi Takiguchi, and Philip J R Goulder. Differential clade-specific HLA-B*3501 association with HIV-1 disease outcome is linked to immunogenicity of a single Gag epitope. *Journal of virology*, 86(23):12643–54, December 2012.

[87] Philippa C Matthews, Alasdair J Leslie, Aris Katzourakis, Hayley Crawford, Rebecca Payne, Andrew Prendergast, Karen Power, Anthony D Kelleher, Paul Klenerman, Jonathan Carlson, David Heckerman, Thumbi Ndung'u, Bruce D Walker, Todd M Allen, Oliver G Pybus, and Philip J R Goulder. HLA footprints on human immunodeficiency virus type 1 are associated with interclade polymorphisms and intraclade phylogenetic clustering. *Journal of virology*, 83(9):4605–15, May 2009.

[88] Philippa C Matthews, Andrew Prendergast, Alasdair Leslie, Hayley Crawford, Rebecca Payne, Christine Rousseau, Morgane Rolland, Isobella Honeyborne, Jonathan Carlson, Carl Kadie, Christian Brander, Karen Bishop, Nonkululeko Mlotshwa, James I Mullins, Hoosen Coovadia, Thumbi Ndung'u, Bruce D Walker, David Heckerman, and Philip J R Goulder. Central role of reverting mutations in HLA associations with human immunodeficiency virus set point. *Journal of virology*, 82(17):8548–59, September 2008.

[89] Allan J. Zajac Maureen A. Cox. Shaping Successful and Unsuccessful CD8 T Cell Responses Following Infection, 2010.

[90] Andrew J McMichael, Persephone Borrow, Georgia D Tomaras, Nilu Goonetilleke, and Barton F Haynes. The immune response during acute HIV-1 infection: clues for vaccine development. *Nature reviews. Immunology*, 10(1):11–23, January 2010.

[91] Saurabh Mehandru, Michael a Poles, Klara Tenner-Racz, Amir Horowitz, Arlene Hurley, Christine Hogan, Daniel Boden, Paul Racz, and Martin Markowitz. Primary HIV-1 infection is associated with preferential depletion of CD4+ T lymphocytes from effector sites in

the gastrointestinal tract. *The Journal of experimental medicine*, 200(6):761–70, October 2004.

[92] Stephen A Migueles, Christine M Osborne, Cassandra Royce, Alex A Compton, Rohan P Joshi, Kristin A Weeks, Julia E Rood, Amy M Berkley, Jonah B Sacha, Nancy A Cogliano-Shutta, Margaret Lloyd, Gregg Roby, Richard Kwan, Mary McLaughlin, Sara Stallings, Catherine Rehm, Marie A O'Shea, JoAnn Mican, Beverly Z Packard, Akira Komoriya, Sarah Palmer, Ann P Wiegand, Frank Maldarelli, John M Coffin, John W Mellors, Claire W Hallahan, Dean A Follman, and Mark Connors. Lytic granule loading of CD8+ T cells is required for HIV-infected cell elimination associated with immune control. *Immunity*, 29(6):1009–21, December 2008.

[93] John J Miles, Natalie a Borg, Rebekah M Brennan, Fleur E Tynan, Lars Kjer-Nielsen, Sharon L Silins, Melissa J Bell, Jacqueline M Burrows, James McCluskey, Jamie Rossjohn, and Scott R Burrows. TCR alpha genes direct MHC restriction in the potent human T cell response to a class I-bound viral epitope. *Journal of immunology (Baltimore, Md. : 1950)*, 177(10):6804–14, November 2006.

[94] T. Miura, M.A. Brockman, Z.L. Brumme, C.J. Brumme, F. Pereyra, A. Trocha, B.L. Block, A. Schneidewind, T.M. Allen, D. Heckerman, and Others. HLA-Associated Alterations in Replication Capacity of Chimeric NL4-3 Viruses Carrying gag-protease from Elite Controllers of Human Immunodeficiency Virus Type 1. *Journal of virology*, 83(1):140, 2009.

[95] N Mochizuki, N Otsuka, K Matsuo, T Shiino, a Kojima, T Kurata, K Sakai, N Yamamoto, S Isomura, T N Dhole, Y Takebe, M Matsuda, and M Tatsumi. An infectious DNA clone of HIV type 1 subtype C. *AIDS research and human retroviruses*, 15(14):1321–4, September 1999.

[96] Trine H Mogensen. Pathogen recognition and inflammatory signaling in innate immune defenses. *Clinical microbiology reviews*, 22(2):240–73, Table of Contents, April 2009.

[97] Corey B Moore, Mina John, Ian R James, Frank T Christiansen, Campbell S Witt, and Simon a Mallal. Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. *Science (New York, N.Y.)*, 296(5572):1439–43, May 2002.

[98] John P Moore, P J Klasse, Matthew J Dolan, and Sunil K Ahuja. AIDS/HIV. A STEP into darkness or light? *Science (New York, N.Y.)*, 320(5877):753–5, May 2008.

[99] M Murphey-Corb, L. Martin, B Davison-Fairburn, R. Montelaro, M Miller, M West, S Ohkawa, G. Baskin, J. Zhang, S. Putney, and al. Et. A formalin-inactivated whole SIV vaccine confers protection in macaques. *Science*, 246(4935):1293–1297, December 1989.

[100] L Musey, J Hughes, T Schacker, T Shea, L Corey, and M J McElrath. Cytotoxic-T-cell responses, viral load, and disease progression in early human immunodeficiency virus type 1 infection. *The New England journal of medicine*, 337(18):1267–74, October 1997.

[101] Takuya Naruto, Hiroyuki Gatanaga, George Nelson, Keiko Sakai, Mary Carrington, Shinichi Oka, and Masafumi Takiguchi. HLA class I-mediated control of HIV-1 in the Japanese population, in which the protective HLA-B*57 and HLA-B*27 alleles are absent. *Journal of virology*, 86(19):10870–2, October 2012.

[102] Zaza M Ndhlovu, Jacqueline Proudfoot, Kevin Cesa, Donna Marie Alvino, Ashley McMullen, Seanna Vine, Eleni Stampouloglou, Alicja Piechocka-Trocha, Bruce D Walker, and Florencia Pereyra. Elite controllers with low to absent effector CD8+ T cell responses maintain highly functional, broadly directed central memory responses. *Journal of virology*, 86(12):6959–69, June 2012.

[103] T Ndung'u, B Renjifo, V a Novitsky, M F McLane, S Gaolekwe, and M Essex. Molecular cloning and biological characterization of full-length HIV-1 subtype C from Botswana. *Virology*, 278(2):390–9, December 2000.

[104] T. Ndung'u, Boris Renjifo, and M. Essex. Construction and analysis of an infectious human immunodeficiency virus type 1 subtype C molecular clone. *Journal of virology*, 75(11):4964, 2001.

[105] Shigeru Nomura, Noriaki Hosoya, Zabrina L Brumme, Mark a Brockman, Tadashi Kikuchi, Michiko Koga, Hitomi Nakamura, Tomohiko Koibuchi, Takeshi Fujii, Jonathan M Carlson, David Heckerman, Ai Kawana-Tachikawa, Aikichi Iwamoto, and Toshiyuki Miura. Significant reductions in Gag-protease-mediated HIV-1 replication capacity during the course of the epidemic in Japan. *Journal of virology*, 87(3):1465–76, February 2013.

[106] S J O'Brien, X Gao, and M Carrington. HLA and AIDS: a cautionary tale. *Trends in molecular medicine*, 7(9):379–81, September 2001.

[107] Karen A O'Connell, Yefei Han, Thomas M Williams, Robert F Siliciano, and Joel N Blankson. Role of natural killer cells in a cohort of elite suppressors: low frequency of the protective KIR3DS1 allele and limited inhibition of human immunodeficiency virus type 1 replication in vitro. *Journal of virology*, 83(10):5028–34, May 2009.

[108] Afam a Okoye and Louis J Picker. CD4(+) T-cell depletion in HIV infection: mechanisms of immunological failure. *Immunological reviews*, 254(1):54–64, July 2013.

[109] Florencia Pereyra, Marylyn M Addo, Daniel E Kaufmann, Yang Liu, Toshiyuki Miura, Almas Rathod, Brett Baker, Alicja Trocha, Rachel Rosenberg, Elizabeth Mackey, Peggy Ueda, Zhigang Lu, Daniel Cohen, Terri Wrin, Christos J Petropoulos, Eric S Rosenberg, and Bruce D Walker. Genetic and immunologic heterogeneity among persons who control HIV infection in the absence of therapy. *The Journal of infectious diseases*, 197(4):563–71, February 2008.

[110] J Phair, L Jacobson, R Detels, and C Rinaldo. Acquired Immune Deficiency Syndrome Occurring Within 5 Years of Infection with Human Immunodeficiency Virus Type-1: The Multicenter AIDS Cohort Study. *... Deficiency Syndromes*, 1992.

[111] RE Phillips, S Rowland-Jones, and DF Nixon. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. *Nature*, 1991.

[112] Julia G Prado, Andrew Prendergast, Christina Thobakgale, Claudia Molina, Gareth Tudor-Williams, Thumbi Ndung'u, Bruce D Walker, and Philip Goulder. Replicative capacity of human immunodeficiency virus type 1 transmitted from mother to child is

associated with pediatric disease progression rate. *Journal of virology*, 84(1):492–502, January 2010.

[113] Michael Quigley, Florencia Pereyra, Björn Nilsson, Filippos Porichis, Catia Fonseca, Quentin Eichbaum, Boris Julg, Jonathan L Jesneck, Kathleen Brosnahan, Sabrina Imam, Kate Russell, Ildiko Toth, Alicja Piechocka-Trocha, Douglas Dolfi, Jill Angelosanto, Alison Crawford, Haina Shin, Douglas S Kwon, Jennifer Zupkosky, Loise Francisco, Gordon J Freeman, E John Wherry, Daniel E Kaufmann, Bruce D Walker, Benjamin Ebert, and W Nicholas Haining. Transcriptional analysis of HIV-specific CD8+ T cells shows that PD-1 inhibits T cell function by upregulating BATF. *Nature medicine*, 16(10):1147–51, October 2010.

[114] S Rerks-Ngarm. Vaccination with ALVAC and AIDSVAX to Prevent HIV-1 Infection in Thailand Supachai. *. . . England Journal of . . .*, pages 2209–2220, 2009.

[115] C Rinaldo, X L Huang, Z F Fan, M Ding, L Beltz, a Logar, D Panicali, G Mazzara, J Liebmann, and M Cottrill. High levels of anti-human immunodeficiency virus type 1 (HIV-1) memory cytotoxic T-lymphocyte activity and low viral load are associated with lack of disease in HIV-1-infected long-term nonprogressors. *Journal of virology*, 69(9):5838–42, September 1995.

[116] S L Rowland-Jones and a McMichael. Immune responses in HIV-exposed seronegatives: have they repelled the virus? *Current opinion in immunology*, 7(4):448–55, August 1995.

[117] Jonah B Sacha, Chungwon Chung, Eva G Rakasz, Sean P Spencer, Anna K Jonas, Alexander T Bean, Wonhee Lee, Benjamin J Burwitz, Jason J Stephany, John T Loffredo, David B Allison, Sama Adnan, Akihiko Hoji, Nancy a Wilson, Thomas C Friedrich, Jeffrey D Lifson, Otto O Yang, and David I Watkins. Gag-specific CD8+ T lymphocytes recognize infected cells before AIDS-virus integration and viral protein expression. *Journal of immunology (Baltimore, Md. : 1950)*, 178(5):2746–54, March 2007.

[118] Asier Sáez-Cirión, Chiraz Hamimi, Anna Bergamaschi, Annie David, Pierre Versmisse, Adeline Mélard, Faroudy Boufassa, Françoise Barré-Sinoussi, Olivier Lambotte, Christine

Rouzioux, and Gianfranco Pancino. Restriction of HIV-1 replication in macrophages and CD4+ T cells from HIV controllers. *Blood*, 118(4):955–64, July 2011.

[119] Asier Sáez-Cirión and Gianfranco Pancino. HIV controllers: a genetically determined or inducible phenotype? *Immunological reviews*, 254(1):281–94, July 2013.

[120] Giselle Sanchez, Xiaoyuan Xu, JC Chermann, and I Hirsch. Accumulation of defective viral genomes in peripheral blood mononuclear cells of human immunodeficiency virus type 1-infected individuals. *Journal of virology*, 1997.

[121] G Scarlatti, E Tresoldi, and ÅBjörndal. In vivo evolution of HIV co-receptor usage and sensitivity to chemokine-mediated suppression. *Nature medicine*, 1997.

[122] C. S. Schmidt and M. F. Mescher. Peptide Antigen Priming of Naive, But Not Memory, CD8 T Cells Requires a Third Signal That Can Be Provided by IL-12. *The Journal of Immunology*, 168(11):5521–5529, June 2002.

[123] JE Schmitz, MJ Kuroda, Sampa Santra, and VG Sasseville. ControL of Viremia in Simian Immunodeficiency Virus Infection by CD8 + Lymphocytes. *Science*, 1999.

[124] Arne Schneidewind, Mark a Brockman, Ruifeng Yang, Rahma I Adam, Bin Li, Sylvie Le Gall, Charles R Rinaldo, Sharon L Craggs, Rachel L Allgaier, Karen a Power, Thomas Kuntzen, Chang-Shung Tung, Montiago X LaBute, Sandra M Mueller, Thomas Harrer, Andrew J McMichael, Philip J R Goulder, Christopher Aiken, Christian Brander, Anthony D Kelleher, and Todd M Allen. Escape from the dominant HLA-B27-restricted cytotoxic T-lymphocyte response in Gag is associated with a dramatic reduction in human immunodeficiency virus type 1 replication. *Journal of virology*, 81(22):12382–93, November 2007.

[125] Arne Schneidewind, Zabrina L Brumme, Chanson J Brumme, Karen a Power, Laura L Reyor, Kristin O'Sullivan, Adrianne Gladden, Ursula Hempel, Thomas Kuntzen, Yaoyu E Wang, Cesar Oniangue-Ndza, Heiko Jessen, Martin Markowitz, Eric S Rosenberg, Rafick-Pierre Sékaly, Anthony D Kelleher, Bruce D Walker, and Todd M Allen. Transmission and

long-term stability of compensated CD8 escape mutations. *Journal of virology*, 83(8):3993–7, April 2009.

[126] PM Sharp, E Bailes, and F Gao. Origins and evolution of aids viruses: estimating the time scale. *Biochemical society ...*, (I 998), 2000.

[127] L Shen, Z. Chen, M. Miller, V Stallard, G. Mazzara, D. Panicali, and N. Letvin. Recombinant virus vaccine-induced SIV-specific CD8+ cytotoxic T lymphocytes. *Science*, 252(5004):440–443, April 1991.

[128] R Shibata, T Igarashi, N Haigwood, a Buckler-White, R Ogert, W Ross, R Willey, M W Cho, and M a Martin. Neutralizing antibody directed against the HIV-1 envelope glycoprotein can completely block HIV-1/SIV chimeric virus infections of macaque monkeys. *Nature medicine*, 5(2):204–10, February 1999.

[129] John W. Shiver, Tong-Ming Fu, Ling Chen, Danilo R. Casimiro, Mary-Ellen Davies, Robert K. Evans, Zhi-Qiang Zhang, Adam J. Simon, Wendy L. Trigona, Sheri A. Dubey, Lingyi Huang, Virginia A. Harris, Romnie S. Long, Xiaoping Liang, Larry Handt, William A. Schleif, Lan Zhu, Daniel C. Freed, Natasha V. Persaud, Liming Guan, Kara S. Punt, Aimin Tang, Minchun Chen, Keith A. Wilson, Kelly B. Collins, Gwendolyn J. Heidecker, V. Rose Fernandez, Helen C. Perry, Joseph G. Joyce, Karen M. Grimm, James C. Cook, Paul M. Keller, Denise S. Kresock, Henryk Mach, Robert D. Troutman, Lynne A. Isopi, Donna M. Williams, Zheng Xu, Kathryn E. Bohannon, David B. Volkin, David C. Montefiori, Ayako Miura, Georgia R. Krivulka, Michelle A. Lifton, Marcelo J. Kuroda, J|[ouml]|rn E. Schmitz, Norman L. Letvin, Michael J. Caulfield, Andrew J. Bett, Rima Youil, David C. Kaslow, and Emilio A. Emini. Replication-incompetent adenoviral vaccine vector elicits effective anti-immunodeficiency-virus immunity. *Nature*, 415(6869):331–335, January 2002.

[130] Guido Silvestri. Embracing the complexity of HIV immunology. *Immunological reviews*, 254(1):5–9, July 2013.

[131] Christina F Thobakgale, Andrew Prendergast, Hayley Crawford, Nompumelelo Mkhwanazi, Danni Ramduth, Sharon Reddy, Claudia Molina, Zenele Mncube, Alasdair

Leslie, Julia Prado, Fundi Chonco, Wendy Mphatshwe, Gareth Tudor-Williams, Prakash Jeena, Natasha Blanckenberg, Krista Dong, Photini Kiepiela, Hoosen Coovadia, Thumbi Ndung'u, Bruce D Walker, and Philip J R Goulder. Impact of HLA in mother and child on disease progression of pediatric human immunodeficiency virus type 1 infection. *Journal of virology*, 83(19):10234–44, October 2009.

[132] K R Tindall and T A Kunkel. Fidelity of DNA synthesis by the Thermus aquaticus DNA polymerase. *Biochemistry*, 27(16):6008–6013, 1988.

[133] Elizabeth Trachtenberg, Bette Korber, Cristina Sollars, Thomas B Kepler, Peter T Hraber, Elizabeth Hayes, Robert Funkhouser, Michael Fugate, James Theiler, Yen S Hsu, Kevin Kunstman, Samuel Wu, John Phair, Henry Erlich, and Steven Wolinsky. Advantage of rare HLA supertype in HIV disease progression. *Nature medicine*, 9(7):928–35, July 2003.

[134] Fleur E Tynan, Diah Elhassen, Anthony W Purcell, Jacqueline M Burrows, Natalie a Borg, John J Miles, Nicholas a Williamson, Kate J Green, Judy Tellam, Lars Kjer-Nielsen, James McCluskey, Jamie Rossjohn, and Scott R Burrows. The immunogenicity of a viral cytotoxic T cell epitope is controlled by its MHC-bound conformation. *The Journal of experimental medicine*, 202(9):1249–60, November 2005.

[135] Fleur E Tynan, Hugh H Reid, Lars Kjer-Nielsen, John J Miles, Matthew C J Wilce, Lyudmila Kostenko, Natalie a Borg, Nicholas a Williamson, Travis Beddoe, Anthony W Purcell, Scott R Burrows, James McCluskey, and Jamie Rossjohn. A T cell receptor flattens a bulged antigenic peptide presented by a major histocompatibility complex class I molecule. *Nature immunology*, 8(3):268–76, March 2007.

[136] Bruce D Walker and Xu G Yu. Unravelling the mechanisms of durable control of HIV-1. *Nature reviews. Immunology*, 13(7):487–98, July 2013.

[137] Jaclyn K Wright, Zabrina L Brumme, Jonathan M Carlson, David Heckerman, Carl M Kadie, Chanson J Brumme, Bingxia Wang, Elena Losina, Toshiyuki Miura, Fundisiwe Chonco, Mary van der Stok, Zenele Mncube, Karen Bishop, Philip J R Goulder, Bruce D Walker, Mark a Brockman, and Thumbi Ndung'u. Gag-protease-mediated replication

capacity in HIV-1 subtype C chronic infection: associations with HLA type and clinical parameters. *Journal of virology*, 84(20):10820–31, October 2010.

[138] Jaclyn K Wright, Vanessa L Naidoo, Zabrina L Brumme, Jessica L Prince, Daniel T Claiborne, Philip J R Goulder, Mark a Brockman, Eric Hunter, and Thumbi Ndung'u. Impact of HLA-B*81-associated mutations in HIV-1 Gag on viral replication capacity. *Journal of virology*, 86(6):3193–9, March 2012.

[139] Shou-li Wu, Yan-sheng Yan, Ping-ping Yan, Hai-long Huang, and Hui-rong Wang. Construction and characterization of a full-length infectious clone from a fast-replicating, X4-tropic HIV-1 subtype B' isolate. *Archives of virology*, 155(12):1923–31, December 2010.

[140] Huanbin Xu, Xiaolei Wang, and Ronald S Veazey. Mucosal immunology of HIV infection. *Immunological reviews*, 254(1):10–33, July 2013.

[141] Jiyu Yan, Steffanie Sabbaj, Anju Bansal, Nilesh Amatya, John J Shacka, Paul A Goepfert, and Sonya L Heath. HIV-specific CD8+ T cells from elite controllers are primed for survival. *Journal of virology*, 87(9):5170–81, May 2013.

[142] Simone C Zimmerli, Alexandre Harari, Cristina Cellerai, Florence Vallelian, Pierre-Alexandre Bart, and Giuseppe Pantaleo. HIV-1-specific IFN-gamma/IL-2-secreting CD8 T cells support CD4-independent proliferation of HIV-1-specific CD8 T cells. *Proceedings of the National Academy of Sciences of the United States of America*, 102(20):7239–44, May 2005.