THE IMPACT OF HLA-DRIVEN ESCAPE MUTATION ON VIRAL REPLICATIVE CAPACITY AND IMMUNE CONTROL IN HIV INFECTION

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

University of Oxford, Hilary Term 2017

Ming-Han Chloe Tsai

Wolfson College, Department of Paediatrics

Supervised by:

Professor Philip J. R. Goulder

Department of Paediatrics

THE IMPACT OF HLA-DRIVEN ESCAPE MUTATION ON VIRAL REPLICATIVE CAPACITY AND IMMUNE CONTROL IN HIV INFECTION

Ming-Han Chloe Tsai,
D.Phil. in Paediatrics, Wolfson College,
University of Oxford, Hilary 2017

ABSTRACT

Despite the introduction of antiretroviral therapy, the HIV/HIV epidemic remains an unsolved global health problem. Amongst all the host defence mechanisms, HLA class I molecules have shown the strongest genetic association with delayed disease progression, in particular HLA-B alleles. Numerous studies have shown that the HLA-mediated CD8+ T cell responses play a central role in the immune control of HIV. Yet our understanding of HLA-mediated immune control of HIV remains incomplete, even when considering the best-defined epitopes restricted by the protective HLA alleles at a population level.

The studies I have conducted and describe herein focus on two well-charaterised protective HLA-B molecules, HLA-B*81:01 and HLA-B*27:05; a third protective molecule, HLA-B*52:01, that has not been well-studied hitherto; and finally the most prevalent HLA-B allele in many Asian populations such as Taiwan, HLA-B*40:01, which has an apparently neutral effect on viral replication. This thesis is centred on the Gag-specific immune response, since previous studies have shown the benefits of CD8+ T-cell responses targeting this conserved and immunogenic region of the HIV proteome, in particular the p24 capsid protein. I have investigated here HLA footprints driven by CD8+ T-cell pressure on HIV that are evident in the viral sequences of individuals expressing these HLA molecules. These footprints include novel escape and putative compensatory mutations. The impact of these variants on viral replicative capacity (VRC) and on HIV disease outcome clinical outcomes was examined via fitness assays. These studies identified several escape mutations that effectively cripple HIV. The distinct compensatory pathways available to the virus to mitigate the fitness cost of particular escape mutations were evaluated. In the course of these analyses I have demonstrated the critical influence of the viral backbone, including HIV clade, in combination with particular viral variants, on VRC. Computational modelling analysis has been applied to facilitate understanding of the mechanism by which certain mutants affect the stability of interactions between HLA and viral capsid protein. This thesis offers novel insights into immune control of the key HIV subtypes – B- and C-clade – and of the most severely affected populations – in Africa (South Africa) and Asia (India and Taiwan) – within the global epidemic. This work helps to better define the viral mutation landscape that is essential both for future vaccines designed to corner the virus, and for successful HIV cure strategies.

CONTRIBUTIONS TO THIS THESIS

Much appreciated for all the participants giving the consents and efforts in the studies. My appreciation also goes to the doctors, nurses, and research assistants for their excellent work at the local hospitals in Taiwan, India, South Africa, and the UK.

The work described in this thesis is my own work, except where specified in each chapter where the contribution of others has been detailed.

This thesis would have been impossible without the expertise of the following individuals:

Chapter 3:

- **Maximilian Muenchhoff** contributed to the conception of the study, performed the ELISPOT;
- **Jaclyn Wright** provided the plasmids and guidance for the assays;
- **Emily Adland** and **Anna Carlqvist** participated in mutagenesis;
- **David Cole** and **Andrew Sewell** performed the structural modeling;
- **Julia Roider**, **Jonathan Carlson**, and **Thumbi Ndung'u** contributed to the data analyses.

Chapter 4:

- **Supriya Singh** contributed to the conception of the study, performed the p24 Gag sequencing and phylogenetic analyses;
- **Masahiko Mori** performed the statistical analyses.

Chapter 5:

- **Jacqueline Brener** contributed to the conception of the study, provided some mutant viruses and guidance for the assays, and performed data analyses;
- **Jonathan Carlson** performed the statistical (co-variation) analyses.

Chapter 6:

- **Jason Huang, Wing-wai Wong, Shu-Sheng Chen, Josephine Wu, Ken Li** and **Philip Chiu** contributed to recruitment of the samples and cohort management.

Separately, I would like to acknowledge **Krystyna Nahlik** for her expertise in primer design for site-directed mutagenesis. She provided guidance, feedback, and discussions thereafter. In addition, I would like to thank **Lynne Richardson** and her amazing group at Zoology for helping me with all the public sequencing. Without them, none of the work is possible.

I would also like to thank **Matthew Pace**, **Masahiko Mori**, and **Rebecca Allen** for proofreading the following chapters and contributing many critical insights in this thesis.

This D.Phil. was funded by the scholarship provided by the Taiwan Bio-Development Foundation.

ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisor, Philip, for all the supports and the insightful advices throughout these four years in Oxford. Thank you, Philip, for being willing to let me conduct my studies in this amazing research group. I enjoyed every meaningful discussion and thank you for the trust, guidance, challenges, encouragements, and patience you have kindly given me, in particular during one of my lowest point in life as the start of the D.Phil. Thank you for offering me many opportunities of applying for scholarships, funding, conferences, and talks. All these unforgettable experiences have forged me to a stronger person, whose passion toward science and life will not be easily diminished upon frustrations.

I would like to thank my master supervisor Adriano Boasso for introducing HIV immunology to me and being very supportive in both my master and D. Phil. I would also like to thank Barbara and Adel from the same lab, thank you both so much for the amazing supports and meaningful discussions.

My passion toward medical science was inspired by my undergraduate mentor Jason Huang who also contributed to this thesis by building the cooperation between UK and Taiwan. I would like to thank the excellence of the medical team at the site in Taiwan. Together we started this cohort study from scratch.

Chan, I cannot thank you enough for literally everything. I could not have completed this journey without you. You aspired me to a better person and scientist. You have opened the world for me since day one you started your work in Peter Medawar. Thank you for your passion, kindness, inspiration, and wisdom. Words cannot express my gratitude to you, Chan.

Gahn, Masa and Matt P, thank you all so much for the help and support you kindly offered. I really appreciate your expertise in data analysis and your critical insights. I would also like to thank all the people in the Medawar, including Becky, Supriya, Emily H, Jelke, Lisa, Tianqi, JT, Debbie, Madeleine, Matt J, Anna, Emily A, Cathy, Ellen and Jacqui, not only for your help in the lab but also your friendship. Especially I would like to thank Amna, who has always been very supportive and offered her heart-wormed accompany during my lowest moments. Thank you for everything, Amna. I could not have come this far without you.

Rachel, thank you for going through everything with me. You are the gem. Thank you for the encouragement and the support throughout these four years. Thank you for always being there for me for better for worse. It is my greatest pleasure to spend my days in Oxford with you and Ed. Kris, Candece, Min-Wen, Fu-Sheng, Shao-Yang, Hsiao-hang, Yvette, Doris, Bo, Po-Hsiang, Yegor, Andrea, Greg, Fiona, Sara, Jared, Kyle, Alban, Rudy, Ryotaro, Ting, Takenori, Qingji, Kazuko, Seth, Patience, Hiroki, Steven, Elaine, and Aaron, thank you all, you are like my family in Oxford. Andy, Fran, Marina, Rita, Janice, Ming-Chia, Chia-yi, Ellie, Ayami and Lei-mei, thank you all for giving me supports and believing in me despite the distance.

Thank you, my dearest mum, dad, and Lu-Ming. You are the best family. Thank you for everything you have sacrificed and done for me in order to support my D. Phil. I am so blessed to have you in my life. Thank you for always being there and offering a shelter where I can always go back. I love you. Yi-Ran and Hsu-Wei, thank you for being my amazing sisters and always taking good care of me.

ABBREVIATIONS

°C – degree Celsius

A
aa – amino acid
ADCC – antibody-dependent cell-mediated cytotoxicity
APCs – antigen presenting cells
AIDS – acquired immune deficiency syndrome
ART – antiretroviral therapy
ATP – adenosine triphosphate

B
bp – base pair
bNAbs – broadly neutralising antibodies
BSA – bovine serum albumin

C
CA – capsid
CCR5 – chemokine co-receptor C-C chemokine receptor 5
CDC – centers for disease control
CRF – circulating recombinant form
CTD – C-terminal domain
CTL – cytotoxic T-lymphocytes
CXCR4 – C-X-C chemokine receptor 4
CypA loop – cyclophilin A-binding loop

D
D1 – Daughter 1
D2 – Daughter 2
ddH_2O – double-distilled aater
DC – dendritic cell
DNA – deoxyribonucleotide triphosphate
dNTP – deoxynucleoride triphosphate
ds – double-strand

E
EC – elite controller
ELISA - enzyme-linked immunosorbent assay
ELISPOT – enzyme-linked immunospot assay
EM – electroporation media
ER – endoplasmic reticulum

F
FCS – fetal calf serum

G
GALT – gut-associated lymphoid tissue
GD – Grand-daughter
GM – Grand-mother
GWAS – genome-wide association study

H
HIV – human immunodeficiency virus

HLA – human leukocyte antigen
HPD – height posterior density
HTLV-III – human T lymphotropic virus type III

I
IC – Indian Consensus
IDUs – intravenous drug users
ILCs – innate lymphoid cells
IN – integrase
IQR – interquartile range

K
kb – kilobase
KIR – killer cell immunoglobulin-like receptor

L
LAV – lymphadenopathy-associated virus
LILR – leukocyte immunoglobulin-like receptor
LOD – limit of detection
LTR – long terminal repeats
LTNPs – long-term non-progressors

M
MA – matrix
MHC – major histocompatibility complex
MHR – major homology region
MSM – men who have sex with men
MRCA – most recent common ancestor

N
NAbs – neutralising antibodies
NC – nucleocapsid
NFAT – nuclear factor activated T cells
NFκB – nuclear factor kappa B
NGS – normal goat serum
NMS – normal mouse serum
NTD – N-terminal domain / amino-terminal domain

O
OD – optical density
OLPs – overlapping peptides
OR – odds ratio

P
P-TEFb – positive transcription elongation factor B
PBMCs – peripheral blood mononuclear cells
PBS – phosphate buffered saline
PCR – polymerase chain reaction
PHA – phytohaemagglutinin
PR – protease

R
RT – real-time / reverse transcriptase
RNA – ribonucleic acid
RNAPII – RNA polymerase II

RPMI – roswell park memorial Institute medium

S
SD – standard deviation
SFC – spot-forming cells
SIV – simian immunodeficiency viruses
ss – single standed

T
TAP – transporter for antigen processing
TB – tuberculosis
TBE – tris-borate-EDTA
TCR – T cell receptors
Treg – regulatory T cell

U
URF – unique recombinant form

V
VRC – viral replicative capacity

W
WHO – world health organization

D. Phil. PUBLICATIONS

- **Tsai M**, Muenchhoff M, Adland E, Carlqvist A, Roider J, Cole D, Sewell A, Carlson J, Ndung'u T, Goulder P. 2016. Paediatric non-progression following grandmother-to-child HIV transmission. Retrovirology 13(1):65.

Contributing Author Publications (Under Review)

- Leitman E, Willberg C, **Tsai M**, Chen H, Buus S, Chen F, Riddell L, Haas D, Fellay J, Goedert J, Trocha A, Walker B, Martin J, Deeks S, Wolinsky S, Martinson J, Martin M, Qi Y, Saez-Cirion A, Yang O, Matthews P, Carrington M, Goulder P. 2017. HLA-B*14:02-Restricted Env-Specific CD8+ T-Cell Activity Has Highly Potent Antiviral Efficacy Associated With Control of HIV. Journal of Virology.

- Leitman E, Thobakgale C, Ansari M, Raghwani J, Prendergast A, Tudor-Williams G, Kiepiela P, Hemelaar J, Brener J, **Tsai M**, Adland E, Mori M, Riddell L, Luzzi G, Jooste P, Ndung'u T, Walker B, Pybus O, Kellam P, Matthews P, Gall A, Goulder P. Role for Virus-Specific Cytotoxic T Lymphocytes in Paediatric HIV Cure Strategies Following Widespread Early Escape. The Journal of Experimental Medicine.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1 HIV/AIDS Epidemic

1.1.1 Origin

Human Immunodeficiency Virus (HIV) is the etiologic agent responsible for AIDS (Acquired Immune Deficiency Syndrome), the most advanced stage of infection with severe loss of immunocompetence. Phylogenetic and statistical studies revealed that the origin of main group of HIV-1 (further referred to as HIV) was between 1910 and 1930 (Korber *et al*., 2000) through cross-species transmission in central Africa (Faria *et al*., 2014). The first documented case of HIV was a male patient with sickle-cell-trait and glucose-6-phosphate-dehydrogenase deficiency in now Kinshasa, Democratic Republic of Congo (Motulsky *et al*., 1966; Nahmias *et al*., 1986). HIV infection was retrospectively confirmed by obtaining the partial sequence from the preserved plasma sample (Zhu *et al*., 1997). Since then, the growing epidemic was contingent upon an active transportation network, high population of migrants, colonial social factors and sex trade (Faria *et al*., 2014).

The HIV pandemic officially began in 1981, when five homosexual men with opportunistic infections were reported in the United States (CDC, 1981; Gottlieb *et al*., 1981). The term 'AIDS' was coined by CDC one year later, as the severe immunosuppression syndromes such as Kaposi's sarcoma, pneumocystis jirovecii pneumonia and persistent lymphadenopathy were shown not only in homosexual population but also in haemophiliacs and intravenous drug users (IDUs) (CDC, 1982). In 1983, the virus was isolated by Dr Luc Montagnier at the Pasteur Institute in France (Barre-Sinoussi *et al*., 1983; Gallo *et al*., 1984). Three years later, the International Committee on Taxonomy of Viruses adjudicated that the etiologic agent

of AIDS should be renamed HIV, replacing both of the earlier designations LAV and HTLV-III (Coffin *et al*., 1986).

1.1.2 HIV classification

There are two lineages of HIV, HIV-1 and HIV-2 (Claver *et al*., 1986; Guyader *et al*., 1987). Both fall within the phylogenetic tree of simian immunodeficiency viruses (SIVs) that are isolated from other primates (**Figure 1.1**), which represent independent zoonotic events (Rambaut *et al*., 2004). In particular, HIV-1 is most closely related to SIVcpz, which is found in some sub-species of chimpanzee that inhabit parts of equatorial Western and Central Africa (Santiago *et al*., 2002; Rambaut *et al*., 2004). HIV-1 is classified into four groups: M (major), O (outlier), N (non-M, non-O) and recently discovered P (Plantier *et al*., 2009; Vallari *et al*, 2011). The M group can be further classified into 10 subtypes, namely A-D and F-H, as well as recombinant forms according to the phylogenetic relatedness (**Figure 1.2**; Sharp *et al*., 1994). SIVcpz from *P. t. troglodytes* clusters most closely to subtype B and C of the most virulent HIV-1 M group. Subtype C is responsible for almost 50% of the infections worldwide and subtype B is the most geographically dispersed (Gao *et al*., 1999; Hemelaar *et al*., 2011).

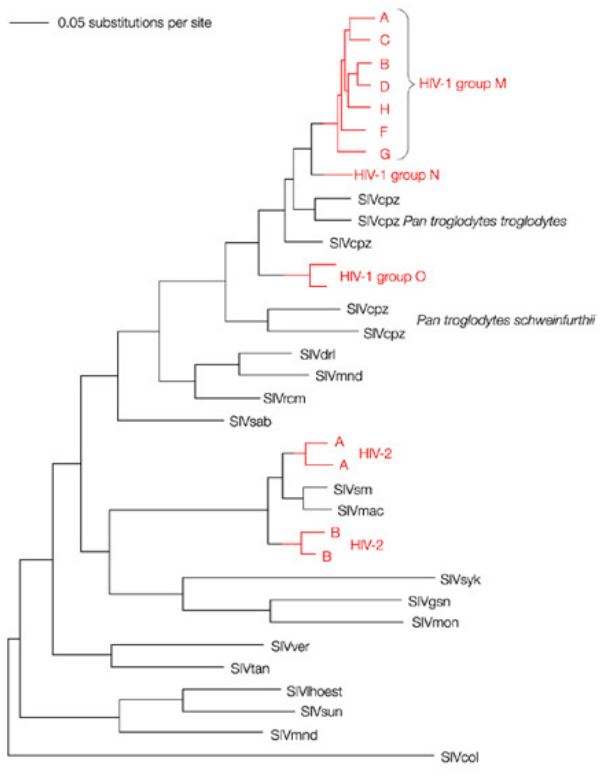


Figure 1.1. Neighbour-joining tree of the primate lentiviruses. Red highlights are the HIV lineages. Figure is adapted from Rambaut *et al*., 2004.

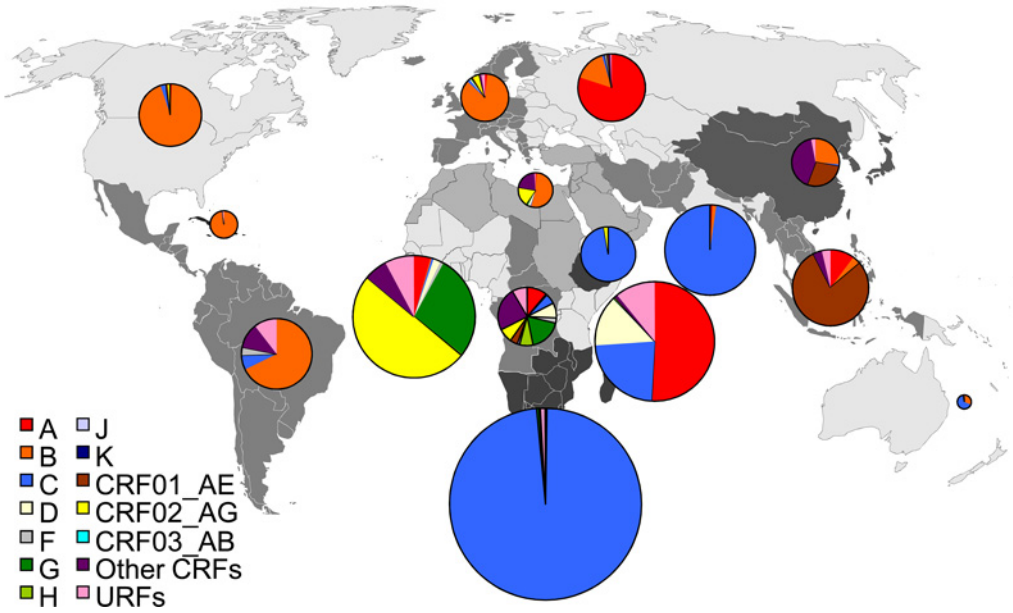


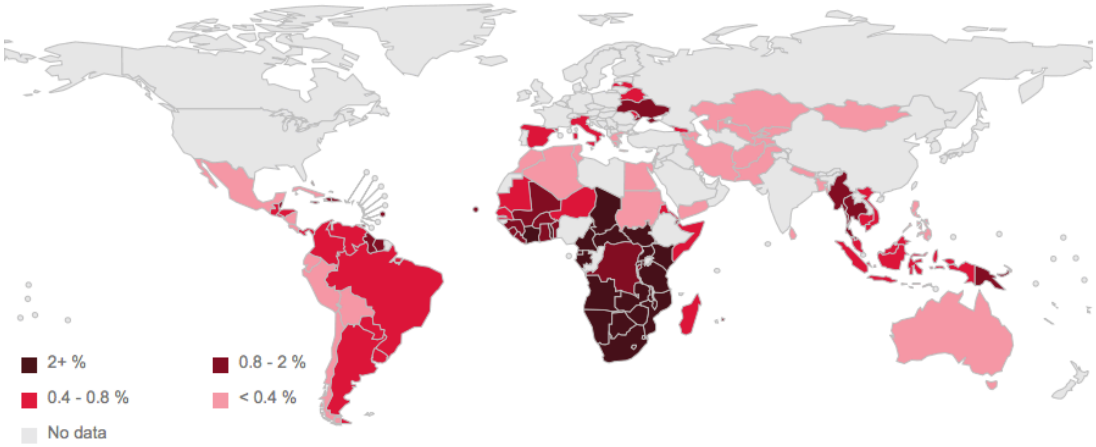
Figure 1.2. Geographical distribution of HIV-1 subtypes from 2004 to 2007. The world was divided into 15 regions, shaded in the same colour according to UNAIDS classification. Pie charts representing the distribution of HIV-1 subtypes and recombinants in each region are shown. The relative surface areas of the pie charts correspond to the relative numbers of people living with HIV in the regions. Figure is adapted from Hemelaar *et al*., 2011.

1.1.3 Current HIV epidemic

There were 36.7 million people living with HIV by the end of 2015, among which 1.8 million were children (UNAIDS AIDS by the number, 2016). Sub-Saharan Africa accounted for nearly 70% of the epidemic and 66% of the new infections arising in 2015 (**Figure 1.3**; UNAIDS AIDS by the number, 2016). South Africa, where most of our cohort studies have taken place (**Section 2.1**; Chapter 3 and 5), had 7 million people living with HIV and only 48% adults had access to antiviral therapy (UNAIDS Prevention Gap Report, 2016). Asia and the Pacific region had the second highest number of current infections (**Figure 1.3**; Chapter 4 and 6). The epidemic was concentrated among large populations such as Chinese, Indian and Indonesian where the treatment coverage was low (36%) (UNAIDS Prevention Gap Report, 2016). Clearly there are still challenges to be met, such as access to antiviral therapy (90-90-90 goal by 2020), adherence, stigma and the expense of antiretroviral therapy

(ART) (UNAIDS 90-90-90, 2014). In addition, although ART has saved millions of lives and normalised survival, life expectancy is still shorter than the normal population, as a result of the fact that, even when viral replication is successfully suppressed by ART, immune activation levels remain higher than in HIV-uninfected individuals (Jones and Núñez, 2012; Samji *et al.*, 2013; Cenderello and Maria, 2015; Muenchhoff *et al.*, 2016; Sluis-Cremer *et al.*, 2015; Wang *et al.*, 2015). In the developed countries with high treatment coverage, the focus has been shifted from AIDS-relating illness and morbidity to liver diseases, cardiovascular diseases and cancer (Farahani *et al.*, 2016).

A



B

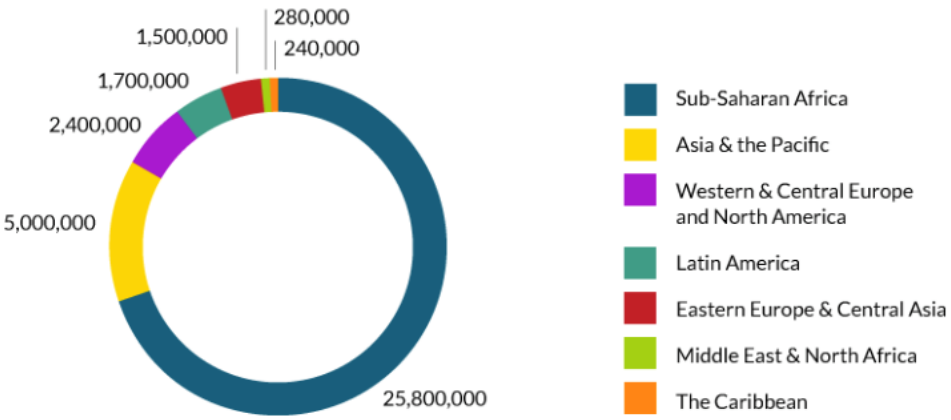


Figure 1.3. HIV prevalence and number of HIV-infected people worldwide. (A) HIV prevalence according to the most updated data from 1990 to 2015. Prevalence in different regions is marked in different colours accordingly. Figure is adapted from UNAIDS (http://aidsinfo.unaids.org). **(B)** Number of HIV-infected people based on the 2014 updated data. Figure is adapted from AVERT (http://www.avert.org/global-hiv-and-AIDS-statistics).

1.2 Structure and Function of HIV

1.2.1 Structure

HIV belongs to *Retroviridae* family, a member of the genus *Lentiviridae*. It is composed of two copies of positive sense and single-stranded (ss) RNA, closely bound to p7 nucleocapsid protein and late-assembly p6 protein (Granoff and Webster, 1999; Lu *et al.*, 2011). The viral genome is enclosed by conical capsid protein p24, along with accessory proteins (Vif, Vpu, Vpr and Nef), regulatory proteins (Tat and Rev) and enzymes (protease, integrase and reverse transcriptase) (**Figure 1.4**). The capsid protein is further surrounded by the matrix protein p17 and with the outer lipid membrane-anchoring envolope glycoprotein complex and some components from the host cells (**Figure 1.4**; Granoff and Webster, 1999; Murphy *et al.*, 2007)

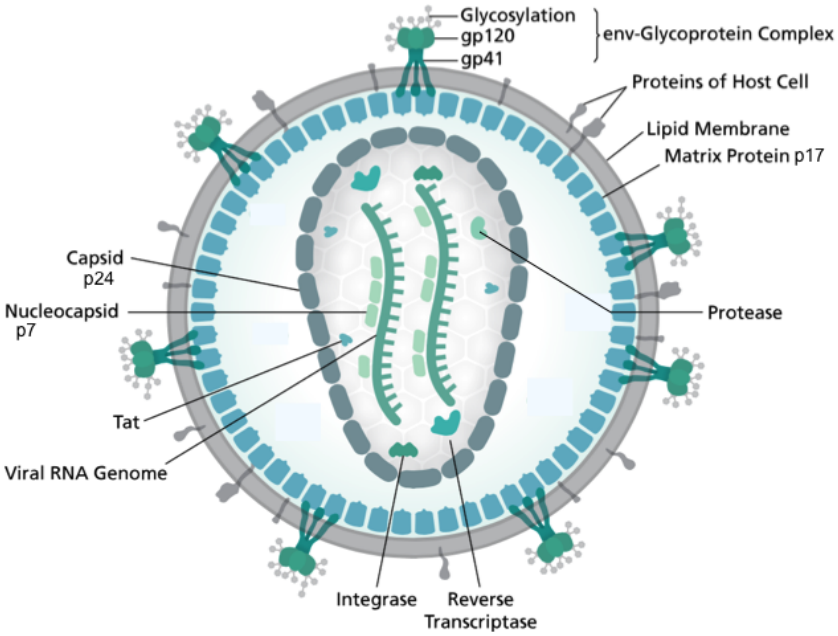


Figure 1.4. Schematic structure of HIV. The mature virion shows the outer surface Env glycoproteins (gp120 and gp41) and the inner core consisting of: a) Four Gag proteins (p24 Capsid, p17 matrix and p7 nucleocapsid. Noted that p6 protein is not shown on the figure but exists within the capsid). b) Two copies of ssRNA bound to nucleocapsid. c) Three enzymes resulting from the cleavage of the polyprotein Pol (Protease, Integrase and Reverse Transcriptase). d) Two regulatory proteins (Tat and Rev not shown). e) Accessory proteins (Vif, Vpu, Vpr and Nef all not shown). Adapted and modified based on Robinson, 2002 and Splettstoesser (http://www.scistyle.com/).

1.2.2 Genome

The 9.7-kilobase HIV genome comprises nine genes and encodes fifteen proteins in total, as illustrated in **Figure 1.5** (Epstein *et al.*, 1991). As retroviruses, HIV possesses three principal genes: *gag*, *pol* and *env*. The *gag* gene encodes p17 matrix, p24 capsid, p7 nucleocapsid which is adjacent to p2 'spacer' peptide (Accola *et al.*, 1998; Greene and Peterlin, 2002). The gene also contains p6 that is involved in viral assembly and budding, separated by p1 peptide that directs translational frameshifting into the overlapping *pol* gene (Jacks *et al.*, 1988; Wilson *et al.*, 1988). The *pol* gene encodes viral protease, reverse transcriptase, integrase and RNase (**Figure 1.5**). The products of *gag* and *pol* are from the precursor Gag-Pol polyprotein, processed by the protease (Könnyű *et al.*, 2013). As for the *env* gene, the precursor Env glycoprotein gp160 is later cleaved into the surface glycoprotein gp120 and the transmembrane protein gp41 (**Figure 1.4** and **1.5**). Details of the functions of each main protein as well as the other accessory and regulatory proteins are listed in **Table 1.1**. The HIV genome also encodes some genomic structural elements apart from the functional proteins. At either end of the genome, the identical long terminal repeats (5'LTR and 3'LTR) are involved in integration of viral DNA (Krebs *et al.*, 2001). Once integrated, 5'LTR acts as a promoter for the entire downstream viral genome, while 3'LTR is responsible for nascent RNA polyadenylation and encodes the accessory protein Nef (Krebs *et al.*, 2001). Other genomic structural elements are detailed in **Table 1.2**.

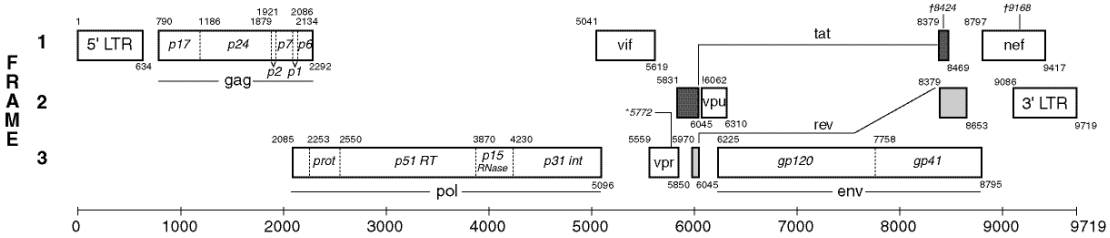


Figure 1.5. HIV genome, reference sequence HXB2 (Genbank K03455). The Open reading frames are shown as rectangles. X-axis is the base pair of nucleotides. Figure is adapted from LANL HIV Sequence Database, https://www.hiv.lanl.gov/content/sequence/HIV/MAP/landmark.html.

Table 1.1. HIV gene products and function.

Gene			Gene product/function
gag	Group-specific antigen	p17, Matrix (MA)	Forms viral matrix layer, targeting Gag polyprotein to lipid rafts. Nuclear import of provirus.
		p24, Capsid (CA)	NTD binds CypA to protect virus from host cell derived restriction factors. CTD is for particle assembly and core formation.
		p7, Nucleocapsid (NC)	Essential for specific packaging two copies of viral RNA into virion.
		p6	Unique to HIV-1, variable in length and sequence. Promotes detachment of assembled virions form cell. Binds Vpu.
pol	Polymerase	p10, Protease (PR)	Gag/Pol cleavage, maturation of translated viral proteins.
		p66, Reverse Transcriptase and RNase H (RT)	Reverse transcription of ssRNA into dsDNA.
		p31, Integrase (IN)	Integration of the provirus into the host genome.
env	Envelope	gp120	Binds to CD4 and chemokine receptors CCR5 and CXCR4.
		gp41	Mediates cell fusion.
tat	Transactivator	p16/14	Activates viral gene expression.
rev	Regular of viral expression	p19	Inhibits viral RNA splicing and promotes nuclear export of incompletely spliced, stablised viral mRNAs. Binds to RRE.
vif	Viral infectivity factor	p23	Accelerates virion maturation and infectivity; inhibits host derived restriction factor APOBEC3G.
vpr	Viral protein R	p15	Nuclear import of provirus; increase virus production; promotes cell-cycle arrest.
vpu	Viral protein U	p23	Promotes CD4 degradation in the ER and viral budding.
nef	Negative-regulation factor	p27	Downregulates CD4 and HLA-A/B. Increases T-cell activation. Increases infectivity.

Based on LANL HIV Sequence Database,

https://www.hiv.lanl.gov/content/sequence/HIV/MAP/landmark.html.

Table 1.2. HIV genomic structural elements.

Element		Function
LTR	Long terminal repeat	634 nucleotides. Contains important regulatory regions, especially those for transcription initiation and polyadenylation.
TAR	Target sequence for viral transactivation	45 nucleotides of viral mRNAs. TAR RNA forms a hairpin structure with a side bulge, which is Tat and cellular proteins binding site.
RRE	Rev responsive element	Approximately 200 nucleotides, located in *env* region. Contains 7 binding sites for Rev.
PE	Psi element	Presence only in unspliced genomic transcrips. A set of 4 stem-loop structures, preceding and overlapping the Gag start codon. Recognised by a conserved motif called the cysteine histidine box, present in the Gag p7 protein.
SLIP	TTTTTT slippery site	Followed by a stem-loop structure. Responsible for regulating ribosomal frameshift out of the Gag reading frame into the Pol reading frame.
CRS	Cis-acting repressive sequences	Located in *pol*. Inhibit structural protein expression in the absence of Rev. The exact function has not been defined.
INS	Inhibitory/Instability RNA sequence	Multiple INSs exist within the genome. Inhibit expression posttranscriptionally. Mutation of the RNA elements was shown to lead to INS inactivation and up-regulation of gene expression.

Based on LANL HIV Sequence Database,
https://www.hiv.lanl.gov/content/sequence/HIV/MAP/landmark.html.

1.2.3 Molecular structure of p24 Gag

Among all protein-coding regions, p24 Gag (CA protein) is one of the most conserved regions. It contains many of the immunodominant T-cell epitopes and therefore it is in particular of interest (Foley, 2000). This conical capsid core consists of approximately 1,500 p24 Gag monomers (Briggs *et al*., 2004), which are assembled predominantly into hexamers, with a few pentamers facilitating the curvature on the top and bottom of the core necessary to form a closed structure (**Figure 1.6 A** and **B**; Ganser *et al*., 1999; Li *et al*., 2000; Campbell and Hope, 2015). Each monomer contains two domains: an approximately 150aa-amino-terminal domain (NTD; CA 1-145, HXB2 133-277) and an 80aa-amino-acid carboxy-terminal domain (CTD; CA 151-231, HXB2 Gag 283-363) (**Figure 1.6 B**). When the monomer

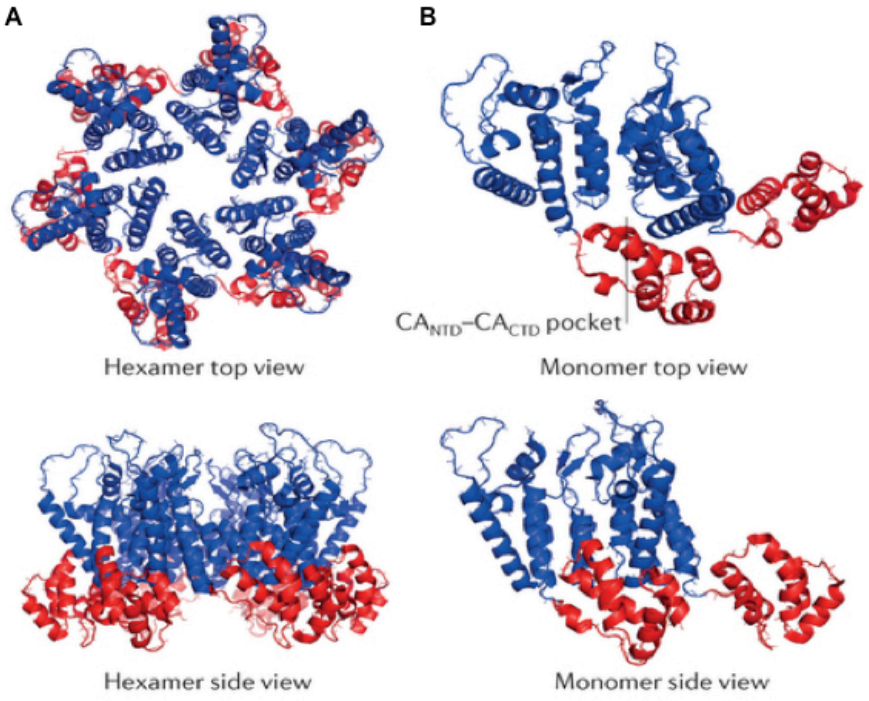
assembles with each other, NTD is facing toward the external surface, whereas CTD structures face inward (**Figure 1.6 A**).

The NTD domain consists of a β-hairpin, 7 α-helices packed in the shape of an arrowhead, with an extended loop connecting helices 4 and 5 that binds the prolyl isomerase, cyclophilin A which is associated with viral assembly (**Figure 1.6 C**; von Schwedler *et al*., 2003; Ganser-Pornillos *et al*., 2008). The cyclophilin A (CypA) loop is known to interact with the host cell factors; for example, it can promote the infection in some cell types with the proline-rich loop present in the viral capsid (CA) (Gamble *et al*., 1997; Tower *et al*., 2003; Sokolskaja *et al*., 2004).

The CTD domain is consisted of a short 3_{10} helix, extended strand and four α-helices (CA helices 8-11) corresponding to its symmetry-related unit (Gamble *et al*., 1997; Ganser-Pornillos *et al*., 2008). The CTD domain also contains a highly conserved 20aa-long major homology region (MHR; CA 153-172, Gag 285-304), which is found in most onco- and lentiviruses (Patarca and Haseltine, 1985; Wills and Craven, 1991) and is crucial for viral assembly, maturation and early steps of infectivity (Strambio-de-Castillia and Hunter, 1992; Mammano *et al*., 1994; Craven et al., 1995). The CTD dimersies the CA in solution (Gamble *et al*., 1997). The conserved MHR not only provides an extended dimer interface (Ivanov *et al*., 2007) but also serves as a high-affinity binding site for cellular elements (Gamble *et al*., 1997).

The flexible link between the two domains – NTD-CTD contacts (CA 145-151, HXB2 278-282) further stablises the hexameric or pentameric subunits (Byeon *et al*., 2009; Zhao *et al*., 2013). This interdomain interface also forms a binding pocket where many cellular factors interact during infection and where two recent antiviral compounds are targeted (Gamble *et al*., 1997; Bhattacharya *et al*., 2014; Price *et al*., 2014;). Together with NTD and CTD, the interdomain linker is crucial for the

structural stability and core assembly of viral genome, enzymes and accessory proteins.



C

```
1
PIVQNLQGQM VHQAISPRTL NAWVKVVEEK AFSPEVIPMF SALSEGATPQ DLNTMLNTVG
GHQAAMQMLK ETINEEAAEW DRLHPVHAGP IAPGQMREPR GSDIAGTTST LQEQIGWMTH
NPPIPVGEIY KRWIILGLNK IVRMYSPTSI LDIRQGPKEP FRDYVDRFYK TLRAEQASQE
VKNWMTETLL VQNANPDCKT ILKALGPGAT LEEMMTACQG VGGPGHKARV L
                                                        231
```

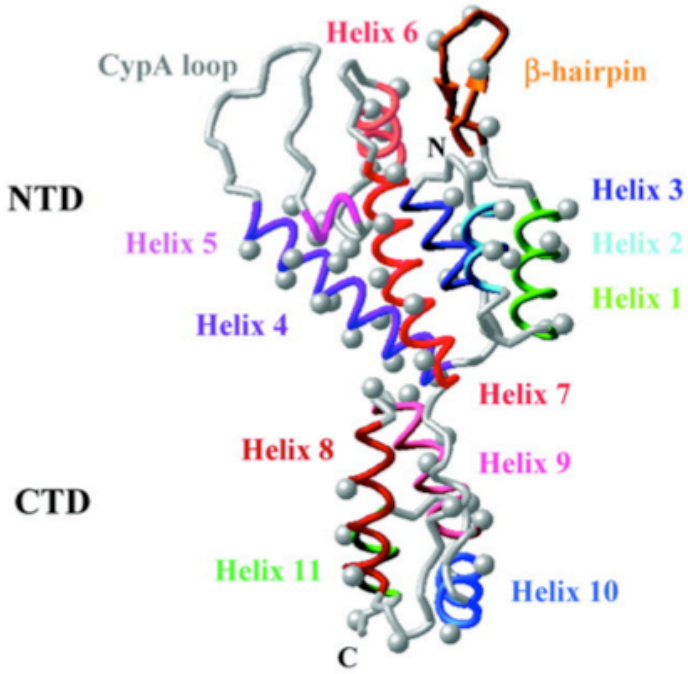


Figure 1.6. Structure and assembly of the HIV-1 p24 Gag protein. (A) Top view and side view of the hexameric subunits that form the primary capsid core of HIV-1 (Protein Data Bank (PDB) identifier: 3GV2). Interactions between the p24 Gag amino-terminal domain (NTD; blue) and the p24 Gag carboxy-terminal domain (CTD; red) of adjacent monomers stabilise the assembled hexamer. **(B)** Structure of two p24 Gag monomers of a hexamer, illustrating the CANTD–CACTD pocket, which mediates interactions between p24 Gag and host cell proteins. **(C)** HIV-1 NL4-3 p24 Gag amino acid sequence. Secondary structures are colour-coded. Structural model for the HIV-1 p24 Gag protein based upon crystal structures of the protein's NTD and CTD. Secondary structures are coded as follows. Orange, β-hairpin; light green, helix 1; turquoise, helix 2; dark blue, helix 3; purple, helix 4; grey, cyclophilin A-binding loop (CypA loop); pink, helix 5; rose, helix 6; red, helix 7; dark red, helix 8; salmon, helix 9; blue, helix 10; and green, helix 11. A short 3_{10} helix in the CTD (CA residues 149 to 152) is not highlighted and the disordered C-terminal 11 amino acids (including P224) are not shown. Figures are adapted and modified from Campbell and Hope, 2015 and von Schwedler *et al*., 2003).

1.2.4 Life cycle

The overview of HIV life cycle is shown in **Figure 1.7**. HIV cell infection begins with the binding of Env glycoprotein complex gp140 and gp41 to the CD4 receptor along with the chemokine co-receptor CCR5 or CXCR4, depending on the viral tropism (Choe *et al*., 1996; Feng *et al*., 1996; Oberlin *et al*., 1996). Most of the circulating HIV strains enter the cell via CCR5 (R5 viruses), while the CXCR4-tropic strains (X4 viruses) are associated with worse disease outcomes, and the use of CXCR4 typically arises in the later stages of the disease (Connor *et al*., 1986). Apart from CD4+ T cells, macrophages and dendritic cells are also vulnerable to HIV since they all express CD4 and CCR5 on the membrane (Gartner *et al*., 1986; Koenig *et al*., 1986; Patterson and Knight, 1987; Tschachler *et al*., 1987).

The binding of the two receptors leads to fusion of the viral and cellular membranes by gp41. The viral capsid core is thus released into the host cytoplasm. During the Vpr-assisted microtubule-based transportation to the nucleus, the viral RNA is reverse-transcribed into double stranded cDNA by reverse transcriptase (RT) and

integrated into the host genome by integrase (IN) (Chan and Kim, 1998). The integrated HIV is referred to as a provirus, which either carries on its life cycle upon cell activation or persists in the host genome in a state of latency (Marcello, 2006).

Once the cell is activated, the up-regulated host transcription factors NFκB and NFAT will bind to the LTR on the viral cDNA end and commence transcription (Nabel and Baltimore *et al*., 1987; Siekevitz *et al*., 1987; Duh *et al*., 1989). Along with several transcription factors, LTR offers the correct binding position to RNA Polymerase II (RNAP2) (**Figure 1.7**; Ott *et al*., 2011). Tat and its cellular co-factor P-TEFb then bind to TAR with high affinity to ensure the elongation of RNAP2-transcription (Kao *et al*., 1987; Wei *et al*., 1998; Price *et al*., 2000). HIV has three mRNA transcripts: approximately 9kb-unspliced mRNA encoding for Gag-Pol precursors; 4.5kb-single-sliced mRNA encoding Env, Vif, Vpu and Vpr; and 2kb-multiply-sliced mRNA encoding Tat, Rev and Nef (Feinberg and Greene, 1992). The export of mRNAs, in particular the incomplete- and un-spliced transcripts encoding for structural proteins, are dependent on the level of Rev (Cullen, 1998). Once the concentration of Rev reaches the threshold, the transcripts are released into cytoplasm and then translated into different proteins (Pollard and Malim. 1998). Followed translation, viral proteins travel to the plasma membrane to assemble as immature virions and become enveloped (Nguyen and Hildreth, 2000). Vpu and p6 Gag assist the budding process by engaging with components of multivesicular bodies (Strack *et al*., 2000; Peterlin and Trono, 2003). Finally, the processing of Gag and Pol from their precursor establishes mature viral particles, which have the capability of infecting other cells (Peterlin and Trono, 2003).

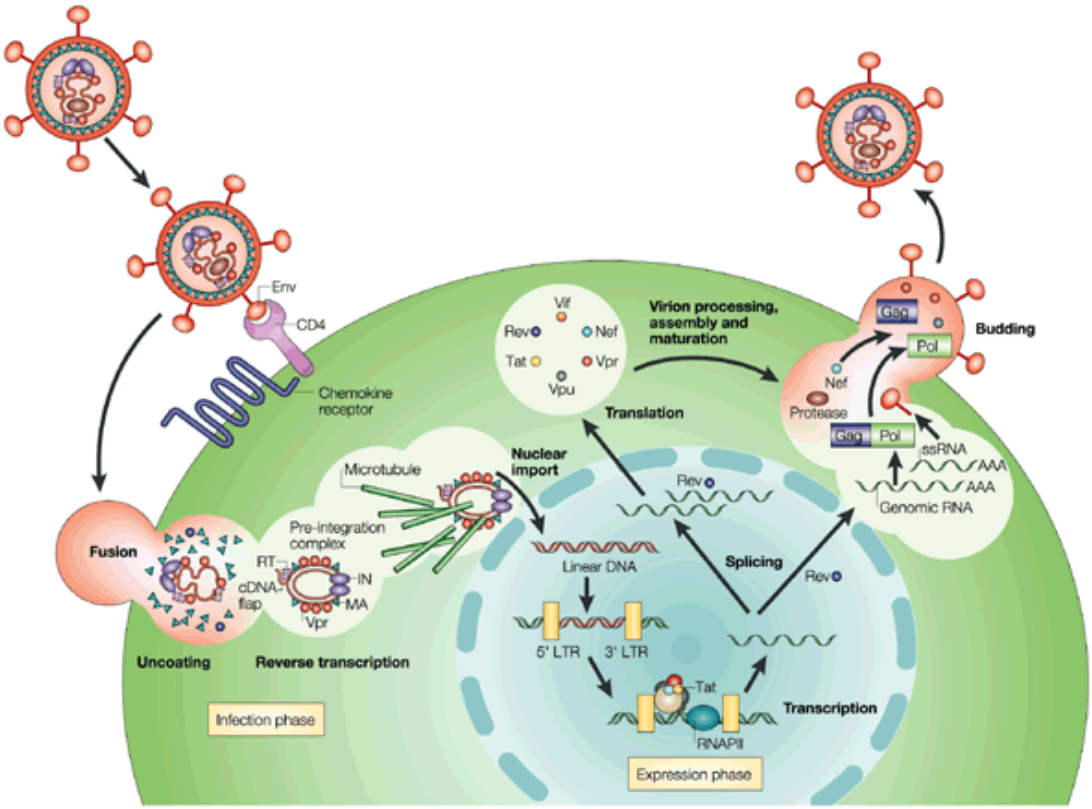


Figure 1.7. HIV replicative cycle. The Envelope protein binds to CD4 and either CCR5 or CXCR4 receptor to let the virus enter the cell by fusion of viral and cellular membrane. The viral capsid is uncoated and viral genome is linearised and reverse transcripted into double-strand cDNA, which is further transported by Rev into the nucleus and integrated into the host genome (provirus). Once the cell activates, proviral DNA produced spliced transcripts and encoded accessary and structural proteins. These are exported from the nucleus and resemble into virion, finally budding out from the cell. Figure is adapted from Peterlin and Trono, 2003.

1.2.5 Evolution

The rate of HIV evolution is remarkably fast. HIV can develop genetic heterogeneity at the speed of 2.4×10^{-3} substitutions per bp per year for Env gp160 and 1.9×10^{-3} substitutions per bp per year for Gag (Korber *et al*., 2000), which is about 10 million times faster than its human host ($\sim 1.1 \times 10^{-8}$ per bp per generation; Roach *et al*., 2010). There are three main variables contributing to this high mutation rate. First is the short time of each viral replicative generation. It takes 1.5-2.6 days for the virus to finish one life cycle, which is equal to 10^{10} to 10^{12} new virions per day (Perelson *et al*., 1996; Rodrigo *et al*., 1999; Salami, 2013). Second, RT is extremely error-prone

without a proof-reading function (Salami, 2013). The error rate is approximately 0.3 nucleotide substitution per second per newly synthesized genome (Preston *et al*., 1988; Mansky and Temin, 1995). The third variable is the high recombination rate. The number of crossovers between two viral RNA templates is estimated at 3-9 per genome per replication (Jetzt *et al*., 2000; Salami, 2013). The recombination rate even outstrips the mutation rate (Jost *et al*., 2002). As a given cell is not limited to infection by only one virus, the 'template switching' of RT between two or multiple viral genomes is likely to create a new recombinant form of virus known as unique recombinant form (URF), especially when the parental templates are subtypes within one group (Jung *et al*., 2002; LANL HIV Sequence Database, http://www.hiv.lanl.gov/). Once an inter-subtype URF is transmitted to three or more people without epidemiological relevance, it is called a circulating recombinant form (CRF) (Robertson *et al*., 2000). For instance, CRF01_AE was the first CRF discovered; and is a common circulating strain in Asia (Murphy *et al*., 1993; Carr *et al*., 1996; Gao *et al*., 1996). To date, there are at least 88 reported CRFs in HIV-1 and numerous URFs (LANL HIV Sequence Database, http://www.hiv.lanl.gov/). Taken together, the fast evolution rate of the virus creates the incredible diversity of the viral genome, which is not only ongoing and associating with different host factors, but also forms a key challenge for HIV vaccine design (Carr *et al*., 2006; Fisher *et al*., 2008).

1.3 HIV Infection

1.3.1 Course of disease

HIV is transmitted via the exposure to body fluids or tissue containing infectious viral particles. Sexual contact is the most common transmission route (Boily *et al*., 2009). Transmission is also common via blood products including needle-sharing among IDUs, transfusion and tissue transplantation (Baggaley *et al*., 2006). Vertical transmission, mother-to-child-transmission (MTCT), accounts for >90% paediatric

infections (WHO. 2017). MTCT is considered to be the transmission at any time point during pregnancy (*in utero*), labour and delivery (*intro-partum*), or breast-feeing (*post-partum*). It remains uncertain if the HIV transmission between two individuals is through a cell-bound or cell-free virus, but *ex vivo* and animal studies showed both HIV/SIV can be transmitted via either form (Sodora *et al*., 1998; Barreto-de-Souza *et al*., 2014).

Following transmission, there is a clinically silent period (~10 days) known as the eclipse phase between infection and the first detection of viral RNA in plasma (Cohen *et al*., 2010). Once the dissemination to lymphoid tissue and systemic circulation occurs, plasma viral load peaks on the third week and positive reaction begins to show in the clinical diagnostic assays (**Figure 1.8**; McMichael *et al*., 2009; Cohen *et al*., 2010). Patients can be classified into Fiebig stages I-VI, based on the detection of viral RNA (PCR), p24 (ELISA), HIV-specific antibody (ELISA/Western blot) (McMichael *et al*., 2009). Typical symptoms such as fever, rash and muscle pain after this seroconversion are similar to those of general viral infections (Soogoor and Daar, 2005). 80% of the circulating viruses at this point are initiated by a single virus referred as the primary founder virus (Keele *et al*., 2008; Salazar-Gonzalez *et al*., 2009). Part of the virus quasispecies become latent and might arise later in the infection (Salazar-Gonzalez *et al*., 2009). Viral latency is the key reason why HIV cannot be eradicated by either immune system or antiretroviral therapy (ART) (Dinoso *et al*., 2009). Acute infection is defined as the duration – approximately 4 weeks – from the acquisition of HIV to the establishment of viral set-point, following viral load decrease from peak viraemia (Schacker *et al*., 1998). Viral set-point is an important indicator as it negatively correlates with time to AIDS (**Figure 1.8**; Wei *et al*., 1995; Mellors *et al*., 1996; Lyles *et al*., 2000). Without ART intervention, the length of this period can vary from <5 years to >10 years (median time: 8-10 years; Pantaleo and Fauci, 1996; Payne *et al.*, 2014). Treatment-naïve individuals who are

asymptomatic for a long time with normal CD4 count are termed as long-term non-progressors (LTNPs) (Poropatich and Sullivan, 2009; Gaardbo *et al*., 2012). The exact definition of LTNPs still remains debatable, but it is generally acknowledged as ART-naïve individuals whose CD4 count is >500 cells/mm^3 for more than ten years without developing AIDS, accounts for 2% population (Okulicz *et al*., 2009). Elite controllers (EC) are defined as LTNPs whose viral load remains undetectable. EC are rare, accounting for <1% in the population (Deeks and Walker, 2007; Hubert *et al*., 2000). Finally, AIDS is defined as the last 1 to 2 years before death when either the CD4 count drops below 200 cell/mm^3 or an opportunistic illness develops (**Figure 1.8**; Osmond *et al*., 1994; Gail *et al*., 1997).

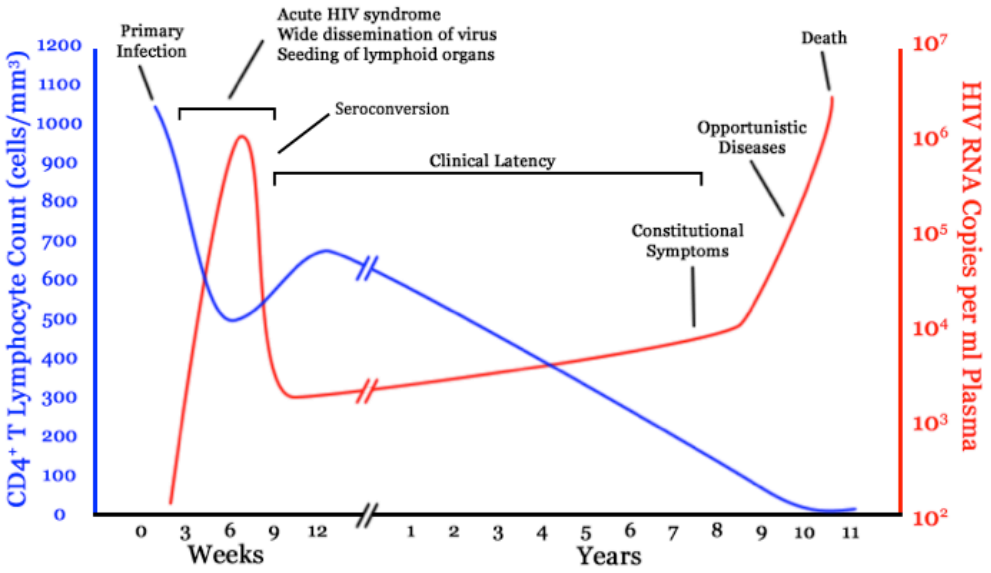


Figure 1.8. Natural course of HIV infection. The primary infection starts from the acquisition of HIV with a window period until seroconversion, followed by the peak of viraemia and the decline of CD4 count. Symptoms of the acute infection show during this phase and the virus spreads systemically. The chronic infection is established after the viral set-point when the viral load decreases and stablises, whereas CD4 count is at a gradual decline rate for the 8-11 years. The final stage of the infection is the developments of AIDS with opportunistic infection and/or CD4 count <200 cell/mm^3, eventually leading to death. Figure developed and modified from Pantaleo *et al*., 1995.

1.3.2 HIV infection in children

The MTCT transmission and natural course of HIV infections in children differ from adults greatly. First, most of HIV acquisition happens during the development of the child's immune system. The fact that infected children are born to HIV-infected mothers who attend hospitals or clinics to deliver, provides the chance of early initiation of ART (Jani *et al*., 2014; Goulder *et al*., 2016). There are prominent case-study reports such as Mississippi child and the VISCONTI child (Persaud *et al*., 2013; Luzuriaga *et al*., 2015.) Frange *et al*., 2015 suggests that the possibility that early treatment in children may be more likely to result in post-treatment control. The timing and route of infection (*in utero, intro-partum*, or breast-feeding), host, viral and maternal factors as well as conditions all play a role in transmission rate, disease progression and treatment strategy (John and Kreiss, 1996; Tobin and Aldrovandi, 2013; Adland *et al*., 2015; Muenchhoff *et al*., 2016). Typical paediatric infection in the absence of ART is distinct from that in adults, marked by their rapid progression which leads to death in more than half of children by the age of two years (**Figure 1.9**; Dunn, 2003; Goulder *et al*., 2016). The reason for the fast deterioration remains unclear. Factors contributing include the immaturity of the immune system and the expansion of susceptible CD4+ T cells (Sperduto *et al*., 1993; Pillay *et al*., 2005; Gosselin *et al*., 2009; Monteiro *et al*., 2011).

The underlying mechanisms of the immune control of HIV in paediatric and adult individuals are also profoundly different. A typical paediatric non-progressor is defined as ART-naïve children remaining clinically healthy with a normal-for-age CD4 count. This normal CD4 count varies with age, but is >750 cells/mm^3 in children aged 5-10 years (Shearer *et al*., 2003). Unlike adult elite controllers (ECs) with undetectable viral load, the median plasma viral load of paediatric non-progressors is ~30,000 copies/ml (**Figure 1.9**; Muenchhoff *et al*., 2016, Goulder *et al*., 2016). In ECs, low immune activation is achieved by HLA-restricted, CTL-mediated control of

viraemia. In paediatric non-progressors, viraemia remains high and the mechanism may be related to low expression of CCR5 on central memory CD4 T cells, (Muenchhoff *et al*., 2016). This scenario observed in the paediatric non-progessors resembles that described in the natural hosts of SIV, e.g. sooty mangabey and African green monkey (Dunham *et al*., 2006; Paiardini *et al*., 2011; Silvestri *et al*., 2003). Of note, that the HLA-, especially HLA-B allele-, related viral control and better disease outcomes observed in adult infection have less impact in paediatric HIV infection, perhaps similar to adult post-treatment controllers in this respect (Sáez-Cirión *et al*., 2013; Adland *et al*., 2015).

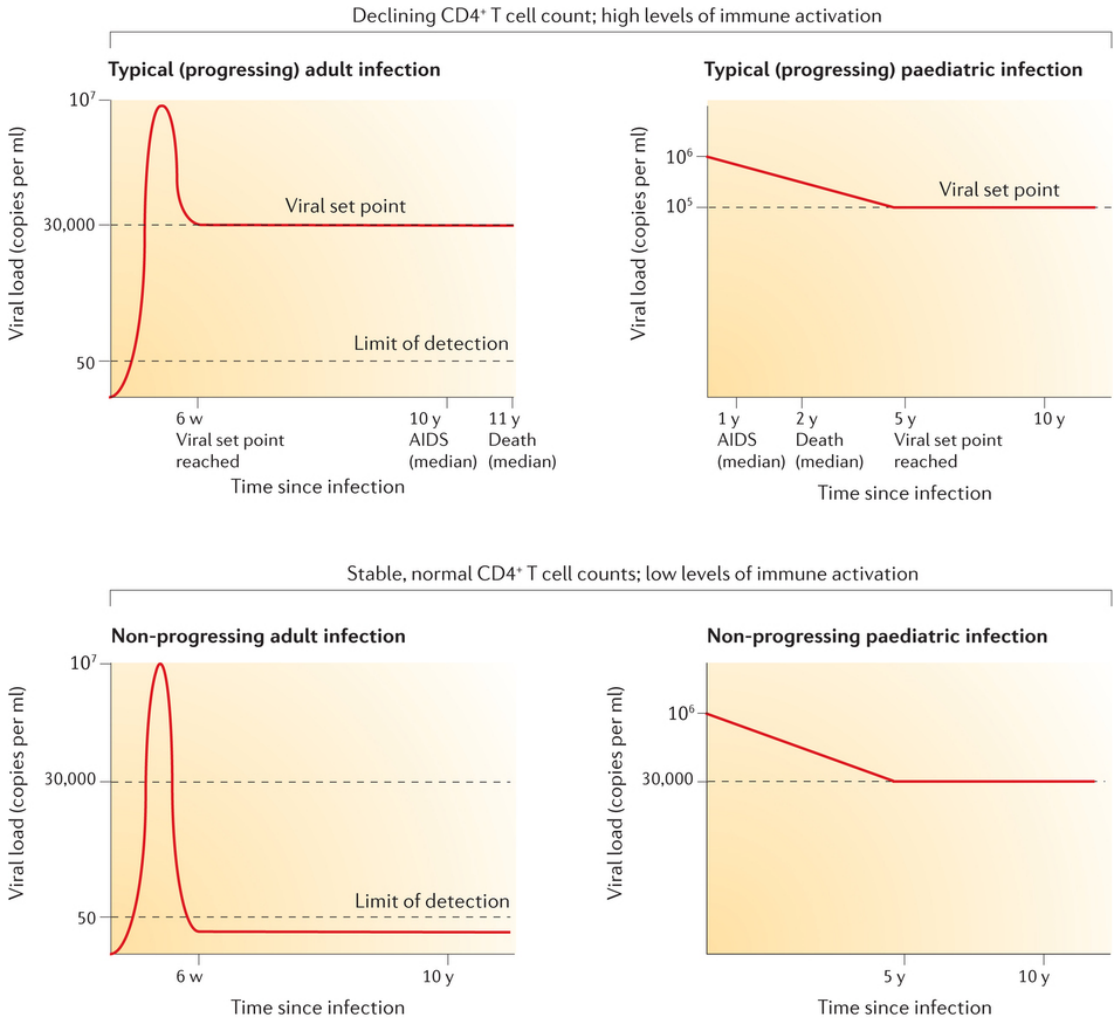


Figure 1.9. Viral loads and progression rate in adult and paediatric infection. Typical progressing adult (left) and paediatric (right) infection are shown in the upper panels. Non-progressing adult (left) and paediatric (right) infection are shown in the lower panels. In the

absence of ART, viral set-point is determined 6 weeks after the acquisition and the chronic infection phase is 8-11 years. In children, the viral set-point is reached after 5 years and 2 years as the median time to progress to death. In non-progressing adults, viral load maintains undetectable whereas in children, the median viral load is ~30,000 copies/ml. Figure is adapted from Goulder *et al*., 2016.

1.4 Innate Immune Response to HIV

As the front line of the defence system, innate immune cells such as mucosal dendritic cells (DCs) and tissue macrophages, make the initial contacts with virus following HIV infection (Espíndola *et al*., 2016). Following the primary local tissue infection, DCs further facilitate the spread of virus by binding to HIV through DC-SIGN (CD229) and migration to lymph nodes or gut-associated lymphoid tissue (GALT) where CCR5+CD4+ naïve and memory T cells reside (Geijtenbeek *et al*., 2000; Turville *et al*., 2002). Innate immune cells enhance both innate and adaptive immune responses to counter the infection via the secreting cytokines and chemokines such as IL-15, IL-18 and type I IFNs from activated DCs, monocytes, NK and NKT cells (Lapenta *et al*., 1999; Stacey *et al*., 2009). However, this 'cytokine storm' leads to immune activation in the acute and sometimes chronic phase, which can contribute to pathology, because its magnitude is excessive and greater than other infectious diseases (Stacey *et al*., 2009). Immune activation also induces apoptosis of B and T cells by upregulating apoptosis-inducing ligands such as TRAIL, TNFR2 and cytokine IFNα (Gasper-Smith *et al*., 2008). It has been shown that the cell numbers of many normal innate immune cells decrease and their functions are partially or fully impaired, largely due to the chronic inflammatory state (Donahue *et al*., 1987; Fantuzzi *et al*., 2004; Killian *et al*., 2006; Jiménez *et al*., 2016). Innate immune cells including some neutrophils (Olinger *et al*., 2000), monocytes (Ellery *et al*., 2007) and even bone marrow progenitor cells (McNamara and Collins, 2011) that express either CD4+, CCR5+, or CXCR4+ are directly permissive to HIV and some become viral reservoirs (Spiegel *et al*., 1992; McNamara and Collins, 2011). Many

cells are activated and recognise HIV upon TLR7/9 stimulation by viral RNA or DNA:RNA hybrid (Ito *et al*., 2006; Saitoh *et al*., 2012; Simmons *et al*., 2013). Interaction with the virus also results in an altered phenotype —immunosenescence and/or immune exhaustion— of cells and changed cytokine production (Deeks, 2011; Simmons *et al*., 2013). Taken together, the profile of innate immune response plays a crucial role in disease progression. Strategies for HIV cure targeting innate immune response must be cautious and must reach a fine balance of the immune system since the stimulation of cytokines acts as a double-edged sword in terms of the defence and assistance of the virus (Sandler *et al*., 2014).

Innate lymphoid cells (ILCs) including nature killer cells (NK cells), which are responsible for bridging innate and adaptive responses, are discussed in **Section 1.5.3**.

1.5 Adaptive Immune Response to HIV

1.5.1 HLA Class I

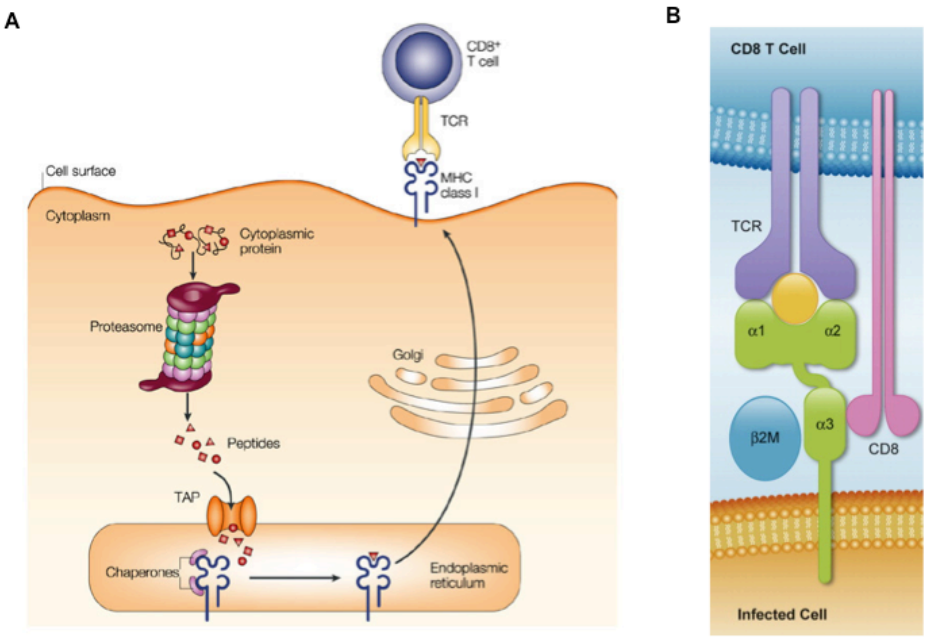
Adaptive immunity is mounted by antigen presentation. Professional antigen-presenting cells (APCs), such as macrophages and dendritic cells, are able to process endogenous and exogenous antigen to helper T cells; and display it on the major histocompatibility complex (MHC) I and II molecules, respectively (Murphy *et al*., 2007). MHC I, also known as human leukocyte antigen (HLA) class I, is expressed on the surface of all nucleated cells, and presents antigen to CD8+ T-cells. MHC I is the most polymorphic region in human genome (Robinson *et al*., 2011). The classical MHC I, located in the short arm of human chromosome 6 (6p21.3), contains the most diverse HLA-A, -B, -C loci (12,375 alleles in total) and the less variable HLA-E, -F, -G (107 alleles) (Horton *et al*., 2004; Robinson *et al*., 2011; HLA Nomenclature, hla.alleles.org). The extreme polymorphism of HLA class I is thought

to be driven by infectious diseases since HLA class I presentation of pathogen-derived epitopes is the basis for CD8+ T cell-mediated recognition (Doherty and Zinkernagel, 1975) and the control of cancers (Wherry and Kurachi, 2015) and infections caused by viruses (Goulder and Walker, 2012) and other intracellular organisms.

The endogenous pathway begins from digestion of intracellular pathogen by proteasome in the cytosol. Ubiquitinated proteins are bound to the proteasome, facilitating their degradation via proteolysis by the β-subunits of the proteasome 20S core particle (**Figure 1.10 A**; Rock *et al*., 1994). The peptides generated from this process can vary in size from 4 to 20 aa (at least *in vitro;* Kisselev *et al*., 1999). The peptides are then transported by transporter for antigen processing (TAP) to the ER through the hydrolysis of ATP (Neefjes *et al*., 1993; van Endert *et al*., 2002). Apart from trans-location of the peptide, TAP also facilitates peptide loading by recruiting the essential molecules such as chaperone tapasin, calreticulin and ERp60 to form a MHC class-I-loading complex (Ortmann *et al*., 1997; Garbi *et al*., 2000; Dick *et al*., 2002). Generally the HLA class I-restricted peptides are 9 (8-11) residues in length, which means only the peptides in the adequate length will fit into the peptide-binding groove of the MHC molecule (Falk *et al*., 1991), whereas the extended ones can be further cut by aminopeptidase in the cytoplasm or in the endoplasmic reticulum (ER) (Serwold *et al*., 2002; Reits *et al*., 2003). Once the peptide is loaded, the complex leaves ER, goes through a post-translational modification in the Golgi apparatus and finally reaches the cell surface, presented to CD8+ T cell triggering cytotoxic T-lymphocyte (CTL) responses (**Figure 1.10 A**; Hang *et al*., 2015).

The HLA class I molecule itself consists of three α domains and a soluble β2-microglobulin (**Figure 1.10 B**; Bjorkman *et al*., 1987a,b). The peptide-binding groove is formed by α1 and α2 domains comprising a sheet of eight β-pleated strand with

two flanking wall constituted by α-helix (**Figure 1.10 C**; Marsh *et al*., 2000). The transmembrane α3 domain together with β2-microglobulin supports the peptide-binding site (Marsh *et al*., 2000). The peptide-binding groove forms six A-F pockets, although not every pocket is necessarily occupied, depending on different HLA types (**Table 1.3**). Pocket B (position 2) and pocket F (carboxy-terminus) are the most selective positions for HLA-A, -B and -C allotypes (Marsh *et al*., 2000; Sidney *et al*., 2008). The features of the binding-motif define the supertype families for HLA-B alleles. In this thesis B07, B27, B44, B58, B62 supertypes are applied for classification according to Sidney *et al*., 2008. Other superfamilies clustering can be based on other features such as the structural interaction (Harjanto *et al*., 2014) HLA class I alleles also expressed certain public allotypes, such as Bw4 and Bw6 epitopes, based on the amino acids at residue 77, 80-83 in the F pocket of α1 helix. The Bw4 public epitopes formed by these residues are ligands for KIR3DL1 and the association between slow HIV disease progression and expression of high expressing KIR3DL1 allotypes in combination with Bw4-80I HLA-B alleles (Altfeld and Goulder, 2007; Martin *et al*., 2007) illustrates the important role HLA I molecules play in shaping the NK response.



C

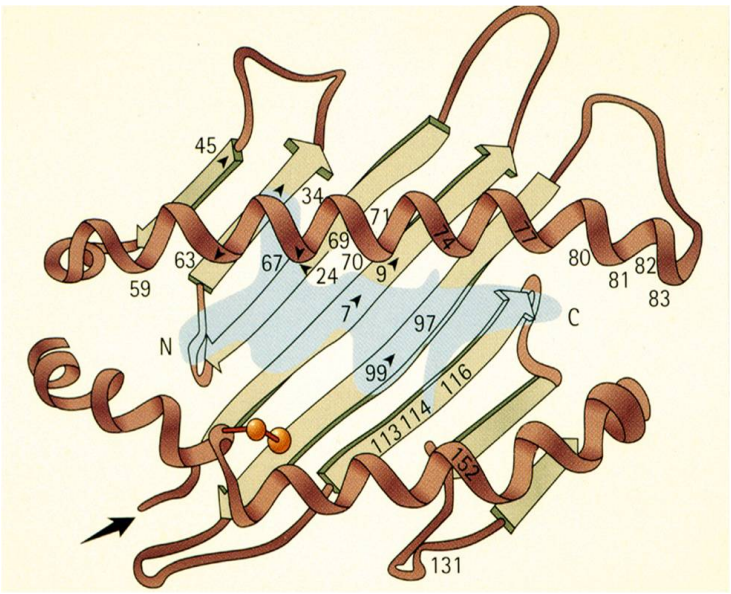


Figure 1.10. Schematic representation of MHC I antigen processing, presenting and the molecule structure. (A) Peptide processing. Proteins of intracellular pathogen are digested by the proteasome into peptides and transported to the ER by TAP. Once the peptide is loaded in MHC class I complex with the assistance of chaperone molecules, the whole complex is trans-located via the Golgi apparatus to the cell surface and presented to the T cell. **(B)** Antigen presentation by MHC class I molecules to CD8+ T cells. **(C)** Top view of MHC class I binding groove. β-sheet and α-helix are coloured in beige and brown, respectively. MHC residue number and N/C-terminus are stated. Blue shadow represents the peptide position. Figures adapted from Yewdell *et al*., 2003; Goulder and Walker, 2012; Khan, 2010 (http://www.hlab27.com/hlab27.aspx)

Table 1.3. Residues correspond to the pockets of HLA class I peptide-binding groove.

Pocket	Constituent residues	Peptide position accommodated
A	5 7 59 63 66 99 159 163 167 171	1
B	7 9 24 25 34 45 63 66 67 70 99	2
C	9 70 73 74 97	6
D	99 113 114 155 156 159 160	3
E	97 114 147 152 156	7
F	77 80 81 84 95 116 123 143 146 147	carboxy-terminus

Adapted from Marsh *et al*., 2000. The HLA FactsBook.

(**Figure 1.10 C**; **Table 1.3**; Goulder and Walker, 2012). Although there are several exceptions to this rule, HLA alleles carrying public allotype Bw4 are shown to be protective against HIV disease progression, whereas those expressing Bw6 tend to

be susceptible to HIV disease progression (Flores-Villanueva *et al*., 2001). The best examples of exceptions to this rule are from studies of African populations that have shown HLA-B*58:02 (Bw4) to be strongly associated with rapid HIV disease progression, and HLA-B*81:01 (Bw6) that is strongly associated with protection against HIV disease progression (Koehler *et al*., 2010; Tang *et al*., 2010; Carlson *et al*., 2012b; Goulder and Walker, 2012; Lazaryan *et al*., 2011; Matthews *et al*., 2012).

In HIV infection, the most decisive single nucleotide polymorphisms (SNPs) associated with delayed disease progression are located in HLA class I region, according to the genome-wide association study (GWAS) (Fellay *et al*., 2009; Limou *et al*., 2009; Pereyra *et al*., 2010). In particular, several studies have shown that the HLA-B alleles have the strongest association with better or worse disease outcomes (Emu *et al*., 2008; Pereyra *et al*., 2008). Well-characterised 'protective alleles' such as HLA-B*27 (HLA-B*27:05) (Kaslow *et al*., 1996; Fellay *et al*., 2009) and HLA-B*57:01 (Migueles *et al*., 2000; Pereyra *et al*., 2010) found enriched in the B-clade infected Caucasian LTNPs; HLA-B*81:01 found protective in C-clade infected African (Kiepiela *et al*., 2004; Wright *et al*., 2010; Carlson *et al*., 2012b). 'Disease-susceptible alleles' such as HLA-B*35 (HLA-B*35:01) are, in contrast, associated with rapid disease progression (Carrington *et al*., 1999; Pereyra *et al*., 2010). The precise mechanism of how the different motifs accommodating various viral peptides and perhaps the peptide-MHC stability contribute to the immune control of HIV still remains uncertain (Goulder and Walker, 2012). A consistent factor that has emerged is the impact of the HIV-specific response directed towards multiple conserved epitopes, from which viral escape may be constrained by effects on viral replicative capacity. For this reason a broad Gag-specific CD8+ T-cell response has been consistently linked with successful control of viraemia, and responses directed towards Env and Nef often linked with high viraemia. However there are clearly many other factors beyond breadth and protein specificity of the CD8+ T-cell response that

shape immune control, including epitope-specific effects and the avidity of the particular T cell receptor for a peptide-MHC complex (**Section 1.5.2**; Almeida *et al.*, 2007; Kiepiela et al., 2007; Sacha *et al*., 2007a,b; Kawashima *et al.*, 2009; Almeida *et al.*, 2009; Payne *et al*., 2010; Chen *et al*., 2012a; Goulder and Walker, 2012).

Although HLA-C is not as variable as HLA-A and -B alleles (Robinson *et al*., 2011), one of the privileges of HLA-C alleles is the exemption from Nef-mediated downregulation of HLA class I molecules, which enables the infected cells to remain susceptible to HLA-C-restricted immune response (Cohen *et al*., 1999; Goulder and Walker, 2012). Apart from eliciting a CTL response, the major role of HLA-C is to serve as a ligand to killer immunoglobulin receptor (KIRs) expressed on NK cells for either activation or inhibition (Blais *et al*., 2011). It is difficult to evaluate the specific contribution of HLA-C to immune-mediated control due to its often tight linkage disequilibrium (LD) with certain HLA-B alleles (Kiepiela *et al*., 2004; Leslie *et al*., 2010), for example HLA-C*12:02/HLA-B*52:01. This haplotype is found to be protective against HIV disease progression in Japanese cohorts (Naruto *et al*., 2012; Murakoshi *et al*., 2016). The protective effect of the CC variant within 3'UTR, 35kb upstream of the HLA-C locus, identified by GWAS, is responsible for regulating HLA-C expression: high expression to be associated with viral set-point (Fellay *et al*., 2009; Thomas *et al*., 2009; Kulkarni *et al*., 2011; Apps *et al.*, 2013).

1.5.2 CD8+ T cell response

After HLA class I presentation of the peptide, CD8+ T cell is primed with the assistance of CD4+ T cells and the professional antigen presenting cells with their co-stimulatory signals (Bevan, 2004). The priming of CD8+ T cell responses result in a clonal expansion of HLA-restricted CD8+ effector T cells, which then secrete cytotoxins (e.g. perforin, granulysins and granzymes) and cytokines (e.g. IFN-γ and TNF-α) to induce the apoptosis of the infected cells (CTL response) (Trapani and

Smyth, 2002; Murphy *et al*., 2007). Several studies have demonstrated the important role of the CTL response in suppressing HIV infection (Schmitz *et al*., 1999; Kiepiela *et al*., 2007; Saez-Cirion *et al*., 2007). During acute infection even prior to peak viraemia, the specific HIV-specific T cell responses that are elicited influences the viral set-point (Ndhlovu *et al*., 2015) and rate of subsequent disease progression (Lyles *et al*., 2000). In chronic infection, there are several features characteristic of an effective CTL responses in ECs compared to disease-progressors. First regarding the phenotype, the HIV-specific CD8+ T cells in ECs show lower expression of exhaustion and senescence markers (Quigley *et al*., 2010), high proliferative profile (Migueles *et al*., 2002) and cytotoxic activity (Zimmerli *et al*., 2005). Typically CTLs in ECs show a high production of cytotoxic granule components or IFN-γ, as a larger magnitude of CTL responses (Migueles *et al*., 2008). The HLA-class-I-restricted CTLs are more resistant to the inhibitory effect of regulatory T cells (Tregs) (Elahi *et al*., 2011). CD8+ T cells are typically polyfunctional in ECs (Betts *et al*., 2006). In addition, the distinct features of the T cell receptors (TCRs) in ECs are the high functional avidity and capability of cross-recognition of viral mutations (Almeida *et al*., 2007; Chen *et al*., 2012a). Last but not least, as described above, targeting the vulnerable parts (e.g. Gag and Pol) of viral proteome also contributes to slower disease progression, in particular Gag-specific CTL responses (Zuniga *et al*., 2006; Kiepiela *et al*., 2007). This might be due to the large amount of Gag present in infected cells and the more rapid digestion of these proteins in the cytosol (Sacha *et al*., 2007a; Chen *et al*., 2012a). Another reason might be the higher fitness cost for the virus to mutate within these conserved structural regions of HIV genome (Miura *et al*., 2009a; Dahirel *et al*., 2011). However, it is noted that immune responses targeting epitopes outside of Gag can also mount effective immune control and suppress viral infection (Maness *et al*., 2008; Valentine and Watkins, 2008).

To yield a competent CTL response, cooperation with HIV-specific CD4+ T cells (Lichterfeld *et al*., 2004) and DCs (Lichterfeld and Yu, 2011) are essential. Polyfunctionality of CD4+ T cell response, including the production of IL-2 and IL-21 to augment CD8+ antiviral activity, as well as the high functional avidity, are features of ECs (Chevalier *et al*., 2010; Vingert *et al*., 2010). HLA class I also serves as a binding site to leukocyte immunoglobulin-like receptors (LILRs) expressed on DCs (Lichterfeld and Yu, 2011). As DCs are the central APCs to CD4+ T cells, functional DCs therefore can modulate not only CTL but also the whole immune response.

1.5.3 Role of NK cells

NK cells are the first line of defence against tumours and viral infections (Murphy *et al*., 2007). In humans, as described above, HLA-B and -C molecules provide a diverse array of ligands for KIRs, whilst the more conserved HLA-E binds the CD94:NKG2A/C receptor (Parham *et al*., 2012). The KIR-binding site is located at the carboxy-terminus of the peptide-binding groove (residue 80 and 83) (Boyington and Sun, 2002; Sanjanwala *et al*., 2008). When encountering HIV, NK cells contribute to viral control through several pathways. The HLA-B-related pathway is the combination of KIR3DL1 and/or KIR3DS1 to HLA-Bw4 alleles expressing Isoleucine at residue 80; both KIR3DL1 and KIR3DS1 in combination with the relevant HLA-Bw4 alleles are associated with improved HIV disease outcomes (Martin et al., 2002; Martin *et al*., 2007). Normally, NK cells are inhibited by HLA class I, so that they identify the 'missing-self' infected cells whose HLA is downregulated, by Nef in the case of HIV (Ljunggren and Kärre, 1990). The exact mechanism by which KIR3DL1 (inhibitory KIR) and KIR3DS1 (activatory KIR) in combination with the same HLA-Bw4 ligand can both be protective remains unclear (Goulder and Altfeld, 2007; Carrington *et al*., 2008). The CD16+ expressing subset of NK cells that are capable of antibody-dependent cell-mediated cytotoxicity (ADCC) against HIV infection may involve KIR3D/HLA-Bw4 80I (Parsons *et al*., 2012;

Espíndola *et al*., 2016). Interestingly, although the role of HLA-E in immune control of HIV is still largely unknown, recent studies have pointed out that the binding of HLA-E with its restricted viral peptide aborts the interaction with inhibitory NKG2A/CD94 receptor (Davis *et al*., 2016). However, the principal peptide binding and therefore stablising HLA-E on the surface of the cell, is the leader sequence of HLA-A, -B and -C molecules. Most HLA haplotypes in human populations are such that they either bind CD94:NKG2A+ NK cells or KIR+ NK cells (Horowitz *et al*., 2016) and it has been proposed that subjects in whom NK cells are largely educated through the CD94:NKG2A+ pathway, and therefore whose antiviral NK activity is largely inhibited, will have less control of HIV than those whose NK cells are principally educated through the KIR pathway. This would also explain the observation that highly expressed HLA-A alleles are associated with higher HIV viral set-points, potentially via the education of NK cells through the CD94/NKG2A+ pathway (Naranbhai, Carrington *et al.*, unpublished data).

The other defences that NK cells are capable of involve more direct contact with HIV. For instance, NK cells are able to block the R5-tropic virus by producing chemokines CCL3, CCL4 and CCL5 which are the ligands for CCR5 (Scott-Algara *et al*., 2003). The ligand of cytotoxic receptor NKp44 can be induced by viral protein gp41 on infected CD4+ T cells, promoting the recognition and killing of NK cells (Vieillard *et al*., 2005).

1.5.4 Other host defence mechanisms

Humoral Immune Response. Establishment of the viral set-point, which strongly predicts time to AIDS, takes place at 6 weeks post-infection, and is coincident with the appearance of virus-specific CD8+ T-cell responses (Koup *et al*., 1994; Borrow *et al.*, 1994). Neutralising antibodies (NAbs), on the other hand, do not develop for months or years after infection, indicating that these do not contribute to viral set-

point. Broadly NAbs (bNAbs) are capable of binding many Env sequences across different subtypes (Sather *et al*., 2008), however their occurrence is correlated rather with high viral loads than with control of viraemia (Muenchhoff *et al*., 2016). Also, bNAbs are rare especially in the early course of infection (Scheid *et al*., 2009). Although bNAbs are greatly in vogue currently in order to prevent HIV transmission and potentially as a therapeutic agent, they do not typically neutralize autologous virus in the subject from whom they were derived (Baum, 2010). In addition to preventing viral entry to uninfected target cells, bnAbs also have a potential role in eliminating HIV-infected targets via ADCC. Abs mediating ADCC by binding to FcγIII on effectors such as NK cells, monocytes and neutrophils have generated considerable interest recently in macaques apparently cleared of established SIV infection via administration of a cocktail of two anti-HIV bnAbs (Hessell *et al*., 2016). Titres of ADCCs are higher in ECs and the vaccine-induced responses are able to lower viraemia (Lambotte *et al*., 2009; Brocca-Cofano *et al*., 2011).

CD4+ Cytotoxic T cells. Recent studies have revealed that there are epitopes targeted by CD4+ T cells, triggering an MHC II molecule-restricted immune response that play an important role in the immune control of HIV (Kaufmann *et al*., 2004; Ranasinghe *et al*., 2013). CD4+ cytotoxic T cells correlate with acute viral load and are associated with lower viral set-point (Soghoian *et al*., 2012). The number of CD4+ T cells exhibiting similar killing ability to CTL increases especially during chronic infection (Juno *et al*., 2017).

Non-protein Molecules. Other than cytokines and chemokines, non-protein molecules such as glycolipids and phosphoantigens are able to recruit γδ T cells to fight against infection (Hayday, 2009). Glycerol monolaurate, a fatty acid and its derivatives, have inhibitory activity against proinflammatory cytokines (Li *et al*., 2009).

1.6 Viral Mutations in Response to CTL Immune Control

1.6.1 Escape mutation

As previously described, CTL responses are the dominant means of immune control of HIV, especially in viraemic controllers who possess protective HLA class I alleles. This inevitably casts a strong selective pressure on the virus (Philips *et al*., 1991). Having a rapid mutation rate (**Section 1.2.5**), escape mutations from CTL can occur at a very early time-point soon after seroconversion (Price *et al*., 1997; Jones *et al*., 2004; Goonetilleke *et al*., 2009; Liu *et al*., 2012). It has been demonstrated that the virus is prone to mutate earlier and more frequently in individuals having protective HLA alleles (Matthews *et al*., 2008; Roberts *et al*., 2015). The kinetics governing the selection of an escape mutation varies between individuals and it is possibly affected by the strength and quality of the CTL responses (Goonetilleke *et al*., 2009; Liu *et al*., 2012).

Escape mutation can be classified into three genres based on the mechanism by which the mutation enables evasion of viral recognition. The first type of escape impairs proteosomal processing (Allen *et al.*, 2004; Milicic *et al.*, 2005), affects the binding of TAP on either C- or N-terminus (Schmid *et al*., 2009), or prevents cleavage from ER aminopeptidase (Draenert et al., 2004). This type of mutation is normally located outside of the epitope, but can occur within it (Jones *et al*., 2004; Yokomaku *et al*., 2004). It is more difficult to identify this type of escape since testing through endogenous system is required (Yokomaku *et al*., 2004). The second type of escape is the result of reduced peptide binding to the MHC. Mutation(s) at peptide position 2 and C-terminus, corresponding to HLA B and/or F pocket respectively, have the most obvious impact on binding of the peptide to the MHC; however many other residues, especially those at P1, P3 and, in certain cases, P5 and P7, may have substantial impact on peptide-MHC binding, via interaction with A, D, C and E

pockets, respectively. In addition to these effects on binding of the peptide to the MHC, these alterations often affect the peptide interaction with the TCR on the CD8+ T-cell – the third mechanism of escape, described further below – by altering the secondary structure of the epitope peptide (Reid *et al*., 1996). Among all HLA-driven mutations, this second mechanism of escape accounts for approximately 20% with a 10-fold or more decrease in peptide binding affinity (Carlson *et al*., 2012a,b; Bronke *et al*., 2013). Prominent examples are R264K of HLA-B*27-restricted Gag KK10 epitope (263-272, KRWIILGLNK; Kelleher *et al*., 2001) and A163X of HLA-B*5703-restricted KF11 epitope (162-172, KAFSPEVIPMF; Yu *et al*., 2007). This type of escape has been well studied and documented (Goulder *et al*., 1997; Geels *et al*., 2003; Jamieson *et al*., 2003; Furutsuki *et al*., 2004; Leslie *et al*., 2006). Finally, although the third mechanism of escape does not influence the peptide loading, it reduces or abrogates the recognition by TCR. Thus the mutant peptide is still efficiently presented on the cell surface. This might allow the development of a *de novo* CTL clone that is capable of recognising the variant (Oxenius et al., 2004; Allen *et al*., 2005; Feeney *et al.*, 2005). Collectively, mutations occurring within the epitope disrupting antigen presentation and TCR recognition are the main mechanisms by which HIV evades CTL surveillance. In addition, without certain HLA-restricted CTL pressure, the mutant might reverse back to wild-type amino acid, which has been observed especially in transmission pair studies (Leslie *et al*., 2004; Duda *et al*., 2008; Brener *et al*., 2015).

For HIV, the potential consequence of developing an escape mutation is the alteration of its VRC, also termed as viral fitness since it refers to the capability of an organism to adapt and reproduce in a defined environment (Geretti, 2006). An escape mutant can simultaneously reduce VRC but increase viral fitness. As mentioned previously (**Section 1.5.2**), CTL response targeting Gag, one of the most conserved regions, is associated with immune control of HIV. Therefore it is not

surprising that escape mutation in Gag epitope will result in a considerable fitness cost for the virus. Reduced VRC was observed in many case and cohort studies (Martinez-Picado *et al*., 2006; Schneidewind *et al*., 2007; Crawford *et al*., 2010; Wright *et al*., 2010). The mutant strain of the virus with lower VRC might persist as an attenuated virus (Crawford *et al*., 2010). In fact, escape mutations are also found in ECs and the great decrease of VRC is associated with viraemia control (Bailey *et al*., 2009; Miura *et al*., 2009a,b). Nevertheless, if the escape mutation has limited or moderate impact on VRC, there is also the possibility of losing CTL control, which is reflected by several parameters such as high viral load or progression to AIDS (Goulder *et al*., 1997; Kelleher *et al*., 2001; Feeney *et al*., 2004). Hence, it is noteworthy that there is no absolute direct correlation between reduced VRC and high CD4 count or low viral load. From the point of view of the virus, there appears to be a balance between evading the CTL recognition and diminishing its own replicative capacity. There are several factors besides viral fitness that play influential roles in maintaining the immune control, such as the timing of CTL targeting, T cell efficacy, immunodominance and TCR usage (Leslie *et al*., 2006).

1.6.2 Compensatory mutation

Another consequence of escape mutation might be the emergence of a compensatory mutation, by which reduced VRC is restored towards that of the wild-type (Crawford *et al*., 2007; Schneidewind *et al*., 2007; Schneidewind *et al*., 2008). In other words, it is a positive epistasis for the virus acquiring both escape and compensatory mutation to evade the immune surveillance and still restrain competent replicative capacity. The numbers of escape and compensatory mutations correlate with each other (Gijsbers *et al*., 2013). From the association of certain HLA allele with the frequency of these mutations, a compensatory mutation can be identified and it is located upstream or downstream proximate to the escape epitope especially when it is viewed in the secondary structure (Crawford *et al*., 2007;

Noviello *et al*., 2011; Wright *et al*., 2012). Another characteristic of compensatory mutation is that the mutation alone is able to reduce the VRC without the presence of an escape mutation (Schneidewind *et al*., 2007; Hinkley *et al*., 2011; Wright *et al*., 2012), although exceptions do exist which have influence on viral fitness (Liu *et al*., 2014). In spite of the beneficial aspects, the development and accumulation of two simultaneous and effective mutations may take years or even decades (Brockman *et al*., 2010; Goulder and Walker, 2012). This might partially offer an explanation for the lost of control in LTNPs or ECs who have protective HLA-B alleles such as HLA-B*27 (Schneidewind *et al*., 2007; Schneidewind *et al*., 2008). Due to the natural accumulation of mutation(s) in the circulating sequence, there might be preexisting compensatory mutation(s) prior to the corresponding escape mutation, which further allow the virus to evade CTL recognition (Chopera *et al*., 2011; Liu *et al*., 2014).

1.6.3 Co-evolution of HIV and HLA

As it has been mentioned previously, facing the escape mechanism form of the virus, the host is capable of expanding a *de novo* CTL clone to recognize the variant and mount a different dominant response (**Section 1.6.1**; Han *et al*., 2014). An increased sensitivity of human host restriction factor TRIM5α has also been found to be associated with escape mutations in epitopes targeted by protective HLA (Battivelli *et al*., 2011; Granier *et al*., 2013). These observations point to the evolving relationship between HIV and HLA-restricted immune response. Moreover, competitive evolution can also been seen in the longitudinal study where the adaptation of the virus weakens the protective effect of HLA. For instance, contradicting impacts on disease outcomes of carrying the HLA-B*51:01 allele have been reported based on geographic and evolutionary context. This allele lost its protective effect due to the accumulation of the escape mutations from TI8 epitope (RT 128-135, TAFTIPSI) at the population level in B-clade infection (Kawashima *et al*., 2009; Kawashima *et al*., 2010; Goulder and Walker, 2012). As for C-clade

infection, HLA-B*57 is gradually losing the protective effect in South Africa (Payne *et al*., 2014). Although dynamic analysis per individual shows all the CTL killing rate, the chance of the occurrence of escape mutations and viral fitness decrease and plateau in a longer time frame (Noviello *et al*., 2011), HLA-restricted CTL pressure is responsible for the population level immune evasion and ongoing subtype diversification (Tenzer *et al*., 2014). Different escape strategies were found to be clade-specific (McKinnon *et al*., 2009). Eventually even the best-defined epitope list will need to be revised. The HIV database Los Alamos 'A-list' will require the inclusion of novel epitopes and HLA-restriction (Frahm *et al*., 2004; Llano *et al*., 2013, LANL HIV Sequence Database, http://www.hiv.lanl.gov/). Thirty years since the first case of HIV was reported, there is a clear trend that HIV is gradually becoming pre-adapted to host immunity (Carlson *et al*., 2015). However, instead of being immune-compromised, many host defence mechanisms along with the fitness cost and other evolutionary pressures such as hardship of developing compensatory mutations limit the viral advancement such as increase of the viral load over time (Fraser *et al*., 2014; Payne *et al*., 2014).

1.7 Strategies of Vaccine Design: Conserved Regions and Viral Fitness

Despite of the success of antiretroviral therapy, developing an effective vaccine against HIV spread and achieving remission is still one of the priorities in the global scientific community. To date, strategies for eliciting T-cell response face many challenges. It remains undetermined what level of the magnitude and functionality of CD8+ T-cells should be mounted, in particular for eliminating viral reservoir (Buchbinder *et al*., 2008; Gray *et al*., 2011). Most of the T-cell human trials such as STEP and Phambili were unsuccessful due to the lack of efficacy in T cells, whereas in animal studies, the CTL responses were restricted by HLA-E (Buchbinder *et al*., 2008; Gray *et al*., 2011; Hansen *et al*., 2016). T-cell exhaustion and dysfunction are

another pressing problem (Wherry and Kurachi, 2015). Last of all, the diversity and high mutation rate of the virus will need to be overcome to reach durable and ART-free 'functional cure' of HIV infection (Richman *et al*., 2009; Deeks *et al*., 2012).

The fitness consequences of escape mutations yield an inspiration to conquer viral diversification in vaccine design. Although the clinical benefits arising from escape mutations remain debatable, at least the poor replicative virus is less transmittable and meant we are one step closer to the functional cure (Carlson *et al*., 2014). Conserved regions carrying the risk of significant fitness cost are the potential candidates for T-cell immunogen (Allen and Altfeld, 2008; Troyer *et al*., 2009). Regardless of choosing customised conserved sequences instead of sequences from natural strains, this approach has elicited proper T-cell responses both in breadth and magnitude in Phase I and is able to inhibit autologous CD4+ cell infection, compared to the STEP trial (McElrath *et al*., 2008; Borthwick *et al*., 2014). Yet it is noted that the HLA-associated CTL vaccine requires cellular responses across various HLA alleles in the population (Carlson *et al*., 2015). In the study of Pereyra *et al*., 2014, the viral control was reached by targeting the structurally conserved regions in an HLA-independent manner. The discovery of non-canonical epitope targeted by individuals who express different HLA type has opened the door to the rendering the possibility of this strategy (Frahm *et al*., 2007; Carlson *et al*., 2015).

Another CTL-based strategy is the mosaic vaccine. The mosaics provide the best coverage of natural variations since it consistes of recombination of T-cell (9-mer) epitopes in a given viral population (HIV-1 M group Gag sequences) (Fisher *et al*., 2007). The advantages of this approach are limiting the escape options by including both susceptible and common HLA-restricted escape forms and eliciting cross-reactivity and diversity of the TCR repertoire (Sunshine *et al*., 2014; Carlson *et al*.,

2015). The main challenges of mosaic vaccine are optimising the CTL response with the consideration of viral adaptation to HLA, incorporating the escape along with potential compensatory mutations (generating *de novo* CTL clone) (Allen *et al*., 2005; Mothe *et al*., 2011; Ondondo *et al*., 2016**;** Borthwick *et al*., 2017) and improving the population coverage (Carlson *et al*., 2015; Mann and Ndung'u, 2015).

Collectively, either targeting the conserved regions or taking the mosaic strategy, it requires more understanding of targeted epitopes and viral fitness to corner the virus. Further research needs to be conducted for enhancing high immunogenicity of cellular responses.

1.8 Aims and Hypotheses of this Thesis

Since the crippled VRC offers hope for reaching viral control and serves as an alternative strategy for vaccine design, there is a pressing demand to better define the viral fitness landscape. Recent studies used a computational model to predict the fitness cost and consequences of escape mutations according to viral sequences (Ferguson *et al*., 2013; Mann *et al*., 2014). However, the association of VRC with sequence conservation is not 100% established (Rolland *et al*., 2013; Liu *et al*., 2014).

In this thesis, I examine different HLA-B allele-restricted epitopes, mainly focusing on the conserved and immunogenic HIV p24 Gag sequence in both unique case study and regional cohorts from different epidemics. I use an *in vitro* viral fitness assay to evaluate the consequences of the escape mutations and putative compensatory mutations in both B- and C-clade viruses. The results are discussed relating to clinical parameters such as CD4 count and viral load in chronic adult/paediatric

infection to unveil more about protective HLA-B alleles and their best-defined epitopes. The aims and hypotheses are as follow:

(i) **Hypothesis:** The rare L188F mutation in protective HLA-B*81:01-restricted Gag TL9 epitope (180-188) contributes to paediatric slow disease progression.

Aim: To examine the viral fitness of L188F along with the potential compensatory mutation.

(ii) **Hypothesis:** The V280A in protective HLA-B*52:01-restricted Gag RI8 epitope (275-282) is associated with high CD4 count in a C-clade infected Indian cohort.

Aim: To define the consensus of circulating Indian sequences and assess the viral fitness of V280X.

(iii) **Hypothesis:** S165N along with other putative compensatory mutations outside of Gag KK10 epitope is capable of restoring the VRC of R264K in C-clade infection.

Aim: To identify the patterns of compensatory mutations for R264K in both B- and C-clade.

(iv) **Hypothesis:** HLA-B*40:01 has a moderate effect on disease progression since the escape mutations are common among a given population and do not result in better or worse disease outcomes.

Aim: To examine the HLA-B*40:01-restricted escape mutation frequency between the low and high viral load group in a B-clade infected Taiwanese cohort.

CHAPTER 2: MATERIALS AND METHODS

Chapter-specific methods are presented in the Materials and Methods sections of the relevant chapters. Components of media are listed in **Section 2.10**. Contributions are explicitly stated where relevant.

2.1 Study Subjects

Subjects from the following cohorts were studied:

(I) **HIV-infected Paediatric Slow Progressor Cohort** (Adland *et al*., 2015): chronically HIV C-clade infected subjects were recruited from Kimberley, South Africa, which included 47 treatment naïve mothers and 84 treatment naïve children. The mean absolute CD4 count of the mothers was 358 cell/mm^3 [Interquartile range (IQR) 235-449 cell/mm^3] and the median viral load 65,000 copies/ml (IQR 17,797–264,330 copies/ml). 55 out of the 84 children studied were defined as slow progressors based on failure to meet 2010 WHO criteria to initiate ART. These criteria may be summarised thus: (i) age (all those diagnosed at <1yr are recommended to initiate ART irrespective of CD4 count), (ii) CD4 criteria to initiate ART: a CD4% <25% in children aged 1–4yrs or an absolute CD4 <350 cell/mm^3 in children aged ≥5yrs, (iii) clinical criteria to start ART (WHO clinical disease stage 3 or 4). The mean age of the slow progressors (*n*=55) was 7.5yrs (IQR 66–112 months), mean CD4% 26% (IQR 21–29%), mean absolute CD4 counts 642 cell/mm^3 (IQR 422-824 cell/mm^3) and median viral load 30,000 copies/ml (IQR 6,800–79,000 copies/ml). The case study of a rare grand-mother-to-child transmission was selected from the slow progressor children and their corresponded mothers in this cohort (Chapter 3). Ethics approval was given by the University of the Free State Ethics Committee, Bloemfontein, South Africa; the Biomedical research Ethics Committee, University of KwaZulu-

Natal, Durban, South Africa; and the Oxford Research Ethics Committee, Oxford, UK.

(II) **North Indian Cohort** (Berner *et al.*, unpublished): 169 chronically HIV-infected ART-naïve adults were recruited from AIDS Division of National Centre for Disease Control (NCDC), Delhi, India. The mean CD4+ T cell count was 448 cell/mm^3 (IQR 311-584 cell/mm^3). The viral loads were not determined in this cohort. The CD4 counts and HLA data from this cohort were used to study the impact of HLA-B*52:01 on the immune control of HIV (Chapter 4). Sequences of HLA-B*27:05-positive subjects were analysed in the study of clade-specific immune control (Chapter 5). Ethical approval was obtained from National Centre for Disease Control (NCDC), Delhi, India and permission for carrying the proviral DNA samples required in this study to Oxford, UK, was obtained from Indian Council for Medical Research (ICMR), India.

(III) **Sinikithemba Cohort** (Kiepiela *et al*., 2004; Kiepiela *et al*., 2006): 578 chronically HIV-infected adults, naïve to antiretroviral therapy, were recruited from Durban, KwaZulu-Natal, South Africa. The median viral load of the entire cohort was 35,259 copies/ml (IQR 6,115-139,000 copies/ml). Samples of the HLA-B*27:05-positive individuals from this cohort were studied to identify the clade-specific fitness cost in HLA-B*27:05-mediated immune control (Chapter 3 and 5). The study was approved by institutional review boards at the University of KwaZulu-Natal, Durban, South Africa; Massachusetts General Hospital, Boston, USA; and the University of Oxford, Oxford, UK.

(IV) **Bloemfontein Cohort** (Huang *et al.*, 2009, Prince *et al*., 2012): 884 adult patients recruited at their first visit to government ARV clinics in the Free

State province of South Africa. In this cohort, HLA types were ascertained only for subjects with CD4+ T cell counts <100 or >500 cells/mm^3. The mean CD4+ T cell counts were 45.36 [±standard deviation (SD) 29, *n*=195] and 652 (±SD 179, *n*=110) cells/mm^3 respectively. The viral loads were not determined in this cohort. Samples of the HLA-B*27:05-positive individuals from this cohort were studied to identify the clade-specific fitness cost of HLA-B*27:05-restricted mutants (Chapter 5). The University of the Free State approved the study ethics (ETOVS 10/04 and ETOVS 206/05). The Department of Health of the Free State Province granted approval to conduct the study.

(V) **Zambia Emory HIV Research Project (ZEHRP) Discordant Couples Cohort** (Prince *et al*., 2012): The subjects selected from the cohort were initially HIV-1 serodiscordant partners in cohabiting heterosexual couples with subsequent intracouple (epidemiologically linked) HIV-1 transmission. All 149 patients in this cohort were antiretroviral therapy naïve. The mean CD4+ T cell count one year post seroconversion was 383 cell/mm^3 (IQR 308-503 cell/mm^3) and the median viral load of viral set-point was $\log_{10}$ 4.39 copies/ml (IQR 3.91-4.99 Log_{10} copies/ml). Samples of the HLA-B*27:05-positive individuals from this cohort were studied to identify the clade-specific fitness cost in HLA-B*27:05-mediated immune control (Chapter 5). This study was approved by both the University of Zambia Research Ethics Committee, Luska, Zambia; and the Emory University Institutional Review Board, Atlanta, USA.

(VI) **Gaborone Cohort** (Matthews *et al*., 2011; Kawashima *et al*., 2009): 514 treatment-naïve chronically HIV C-clade infected antenatal women were enrolled from Gaborone, Kimberly, South Africa. The mean CD4+ T cell count

was 342 cell/mm^3 (IQR 220-476 cell/mm^3) and the median viral load was 19,100 copies/ml (IQR 3,920-78,200 copies/ml). Together with Sinikithemba, Bloemfontein and ZEHRP cohorts, samples of the HLA-B*27:05-positive individuals from this cohort were studied to identify the clade-specific fitness cost in HLA-B*27:05-mediated immune control (Chapter 3 and 5). The study was approved by the Office of Human Research Administration, Harvard School of Public Health, Boston, USA; and the Health Research Development Committee, Botswana Ministry of Health, Gaborone, Botswana.

(VII) **Taiwanese Cohort**: 156 chronically HIV-infected treatment-naïve adults were enrolled from Taipei Veterans General Hospital in Taipei, Taiwan, with a mean CD4+ T cell count of 355 cell/mm^3 (IQR 239-466 cell/mm^3) and the median viral load was 27,600 copies/ml (IQR 8,945-28,400 copies/ml). Along with the Thames Valley Cohort, the CD4, viral load, HLA and ELISPOT data from this cohort were used to examine the differences in high and low vial load groups in HLA-B*40:01-positive individuals (Chapter 6). This study was approved by the Institutional Review Boards, Taipei Veterans General Hospital.

(VIII) **Thames Valley Cohort** (Payne *et al*., 2014): Chronically HIV-infected ART-naïve adults were recruited from the Royal Berkshire Hospital, Reading, UK; Northampton General Hospital, Northampton, UK; Churchill Hospital, Oxford, UK; and Wycombe Hospital, High Wycombe, UK. The samples, HLA and Elispot data from this cohort were used in the HLA-B*40:01 study (Chapter 3, 5 and 6). This is still an ongoing cohort. The study was approved by the institutional Review Board of the University of Oxford.

All individuals received counselling and signed a written informed consent form agreeing to participate.

2.2 Blood Processing

2.2.1 Separation of plasma and PBMCs

Fresh whole blood samples were processed within 24 hours of collection in heparin containers. Prior to the separation, 3ml blood was taken and preserved for subsequent DNA extraction (**Section 2.2.3**). The rest of blood was then centrifuged at 2,000rpm for 10 minutes. The plasma was collected after the spin and aliquoted in 1.5-ml cryovials stored at -80°C. Blood was diluted with phosphate buffered saline (PBS), together layered onto Lymphoprep (Stemcell Technologies) and centrifuged at 2,000rpm for 24 minutes with break off. The ratio of blood, PBS and Lymphoprep was 1:1:1. The extra plasma was discarded and the buffy coat layer containing peripheral blood mononuclear cells (PBMCs) were taken to one new 50-ml Falcon tube by using Pasteur pipette. The tube was then topped up with PBS to 45ml and spun at 2,000rpm for 15 minutes. After another two-times of wash with PBS at 2,000rpm for 5 minutes, the cells were resuspended in 20ml R10 medium made from RPMI-1640 (Sigma).

2.2.2 Cryopreservation of cells

All cells were spun at 2,000rpm for 5 minutes and resuspended in filtered FCS (0.45 μm, Sigma) containing 10% DMSO (Sigma) in a drop-wise manner, which was then aliquoted in pre-chilled 1.5-ml cryovials (1-10 million cells/vial). The cyovials were then placed into the prechilled 4°C 'Mr.Frosty' freezing container (Nalgene) that allows constant rate of cooling samples at 1°C per minute. The container was transferred to -80°C freezer for at least 2 hours. All the samples were then moved to liquid nitrogen for long-term storage for further use. To thaw, each cryovial was placed in 37°C water bath and was transferred into pre-warmed R10 medium again

in a drop-wise manner. Cells were further washed for three times in total at 2,000rpm for 5 minutes in pre-warmed R10.

2.2.3 DNA extraction

Genomic and proviral DNA were extracted from fresh whole blood (3ml, as mentioned above) using QIAmp DNA Blood reagents (QIAGEN). The blood was mixed well and incubated with 9ml red blood cell lysis for 10 minutes at room temperature. After the incubation, it was pelleted at 2,000rpm spin for 10 minutes, resuspended in 3ml of white blood cell lysis for more than 48 hours at room temperature. The sample at this stage can be stored up to 18 months. To proceed, 1ml protein precipitation solution was added and vortexed for 20 seconds. The sample was spun down at 2,000rpm for 10 minutes and the supernatant was transferred into 3ml isopropanol, inverted at least 50 times until threads of DNA can be observed. It was then spun down for 3 minutes followed by washing with 70% ethanol and a further spin for 1 minute. The DNA pellet was left to air-dry at room temperature overnight and re-hydrated in RNase/DNase free water (HyClone). The concentration and quality of DNA were measured on Thermo Scientific NanoDrop 1000 spectrophtometer (NanoDrop Products). DNA was kept at -20°C for long-term storage or 4°C for immediate use.

For samples where no fresh blood was available (Chapter 6), PBMCs were thawed and followed the DNA extraction procedure with adjusted initial sample and solution volume of 3-time volume of red blood cell lysis, 1-time volume of white blood cell lysis solution, 1/3 volume of protein precipitation solution, and 1-time volume of isopropanol.

2.2.4 RNA extraction

Plasma aliquots in 1.5-ml cryovial were thawed in 37°C water bath and spun down at 15,000rpm for 1.5 hours. 860 µl of supernatant was discarded, whereas viral RNA

was resuspended in the rest 140 µl and proceeded for extraction by using QIAmp Viral RNA Mini Kit (QIAGEN) following manufacturer's instructions. For samples with low viral load (<10,000 copies/ml), several aliquots of plasma were centrifuged and concentrated together through the same QIAmp column.

2.3 HLA-typing, CD4+ T Cell Counts and Viral Load Determinations

Four-digit HLA class I typing was performed from genomic DNA by sequence based typing at the CLIA/ASHI accredited laboratory of William Hildebrand, at the University of Oklahoma Health Sciences Centre, USA, using locus specific PCR amplification of class I exons 2 and 3 and heterozygous DNA sequencing. Resolution of ambiguities was undertaken according to the ASHI committee recommendations (Cano *et al.,* 2007). CD4+ T cell counts were determined by flow cytometry using standard clinical protocol. HIV plasma viral load was measured using the Roche Amplicor version 1.5 assay with COBAS Amplicor. Data of CD4+ count and viral load were provided by each centre providing clinical care.

2.4 IFN-γ Enzyme-Linked Immunospot Assay

2.4.1 ELISPOT procedure

Elispot assays were performed in sterile 96-well plates with a Polyvinylidene Fluoride membrane (Milipore). Plates were pre-coated with anti-human IFN-γ monoclonal antibody (Mabtech, Sweden), diluted with PBS at 1:2,000, and kept at 4°C until use. Plates can be stored up to one month after antibody coating. On Day 1, plates were washed for six times with blocking buffer containing 1% heat-inactivated FCS in PBS. 50 µl R10 media was added to each well, followed by peptide antigens at a final concentration of 1.25-2.5 µg/ml. 250 µg/ml of phytohaemagglutinin (PHA) (Roche) was used as positive controls in duplicate along with four negative controls without peptides or PHA. Rested PBMCs were added to each well containing $1x10^5$ cells in a

final volume of 160 µl. Plates were incubated at 37°C with 5% CO_2 for 14-16 hours. Cells were then discarded on Day 2 after the incubation and plates were washed with PBS for six times. Biotinylated anti-IFN-γ antibody (Mabtech, Sweden) was added at 1:2,000 dilutions in PBS and incubated at room temperature in the dark for 90 minutes. Plates were washed for six times with PBS, followed by the incubation of streptavidin-alkaline phosphatase enzyme conjugate (Mabtech, Sweden) in PBS at 1:2,000 at room temperature in the dark for 45 minutes. Plates were washed again as previously described and were developed with substrate BCIP/NBT colour solution (Bio-Rad Laboratories) until the positive control wells (PHA) became purple. The reaction was stopped by washing the plates with tap water for six times. Plates were then air-dried overnight. Spots were counted on Day 3 on an automated ELISPOT reader (AID ELISPOT v.4.0, Autoimmun Diagnostika, Germany). Background was defined as the mean of 4 negative controls + 3x SD, which was subtracted form the values of all wells. The results were expressed as number of spot-forming cells (SFC) per 10^6 PBMCs. Positive responses were counted as >50 SFC/10^6 PBMCs after background subtraction.

2.4.2 Screening for HIV-specific CD8+ T cell responses

A panel of 410 overlapping 18mer peptides (OLPs), spanning the entire HIV-1 proteome, was synthesized based on 2001 B- and C-clade consensus sequence (Massachusetts General Hospital Peptide Core Facility, Boston, MA, USA) (Kiepiela *et al*., 2004; Kiepiela *et al*., 2006) All the peptides were mixed in 72 pools with 9-12 OLPs/pool with each peptide present in two different pools, where a given peptide was identified by the positive responses in the two corresponding pools. **Figure 2.1** shows an example of the megamatrix layout. Peptides identified from this approach were then confirmed by performing another ELISPOT assay with optimal peptides, namely, a panel of well-characterised epitopes (LANL HIV Sequence Database, http://www.hiv.lanl.gov/) >80% purity as shown by HPLC profile.

A

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

B

	Pool B1	Pool B2	Pool B3	Pool B4	**Pool B5**	Pool B6	**Pool B7**	Pool B8	Pool B9	Pool B10	Pool B11	Pool B12
Pool A1	1	2	3	4	**5**	6	**7**	8	9	10	11	-
Pool A2	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	-	**12**
Pool A3	25	26	27	28	**29**	30	**31**	32	33	-	23	24
Pool A4	37	38	39	40	**41**	42	**43**	44	-	34	35	36

Figure 2.1. Example of megamatrix setup. (A) ELISPOT 96-well plate layout with 72 HIV-specific peptide pools shown as A1-A12, B1-B12, C1-C2, etc. "+" indicates positive control (PHA) wells, "-" indicates negative controls. Purple shading shows positive responses (pool A2, B5 and B7). **(B)** Partial of peptides constituting pool table is shown (only Pools A1-A4 and B1-B12 are shown for convenience). Purple shading indicates the three positive responses in (A). Responses to pools A2/B5 and A2/B7 identify reactivity to OLP 17 and 19, respectively. OLP 17 and 19 would then be used for subsequent confirmatory ELISPOT assay.

2.5 HIV Sequencing

2.5.1 Sanger sequencing of proviral Gag, Pol, Env and Nef

For all the studies focusing on Gag and B*40:01 study locating epitopes across the HIV proteome (Chapter 6), Gag, Nef, partial Pol and Env segments were amplified from proviral DNA using Bioline reagents by nested PCR as previously described (Leslie *et al*., 2004; Leslie *et al*., 2006). Mastermix reaction composition, amplification primers (Eurofin Genomics) and PCR conditions are shown in **Table 2.1-2.3**. Both B- and C-clade primers were used based on the dominant circulating virus in the cohorts. One negative control (water instead of template DNA) was run for each PCR reaction. After the amplification, 5µl of PCR product were mixed with 5x sample loading buffer (Bioline), and was run on 1% agarose gel (Sigma) in TBE buffer (Promega) containing ethidium bromide (5µl/250ml gel) at 180V for 40 minutes.

Hyperladder I (Bioline) was run along side as the indicator for the size of amplicons. Bands were visualised under UV light. The rest of PCR products were kept at 4°C for immediate use or at -20°C for long-term storage.

PCR products of the correct size were purified by using QIAquick PCR Purification Kit (QIAGEN) to remove leftover primers and dNTP. Sequencing PCR was performed using BigDye terminator v3.1 ready reaction mix (Applied Biosystems) in a thermocycler (**Table 2.4-2.5**). Sequencing primers (**Table 2.6**) were applied to generate bidirectional overlapping fragments for the corresponding proteins. Each Reaction product were precipitated using 2µl sodium acetate with 10µl ddH_2O and 50µl 100% ethanol, spun at 4°C for 1.5 hours at 3500rpm, washed twice with 70% ice-cold ethanol at 4°C for 10 minutes at 3500rpm, and dried in the thermocycler for 1 minute at 94°C with the lid open. Sequences were read by the Department of Zoology (University of Oxford), then manually edited for mixed or ambiguous bases using Sequencher 5.0 (Gene Codes corp.) translated in Geneious 6.1.3 and aligned in Clustal Omega.

Table 2.1. Nested PCR mastermix reaction.

Component	Volume (µl)	Final Concentration
10x NH_4 Reaction Buffer	5	1x
50mM $MgCl_2$	3	2 mM
10mM dNTP mix	0.5	0.2 mM each dNTP
Forward primer	1	0.2 mM
Reverse primer	1	0.2 mM
BioTaq^{TM} DNA Polymerase	0.5	1 Unit
Template DNA	2	Not applicable
ddH_2O	37	Not applicable

Table 2.2. Gag, Pol, Env and Nef amplification primers.

Region of genome	Clade	Primer name	Primer sequence (5'→3')	HXB2 coordinates	Tempe-rature 1 (°C)*	Tempe-rature 2 (°C)*
Gag	B	F1_5'OP	AAATCTCTAGCAGTGGCGCC	623-642	64	61
		F1_3'OP	TGTTGGCTCTGGTCTGCTCT	2138-2157		
		F1_5'IP	ACTAGCGGAGGCTAGAA	765-781	60	57
		F1_3'IP	GCCACAATTGAAACACTT	1960-1977		
	C	F1b_5'OP	CTCTAGCAGTGGCGCCCGAA	627-646	60	57
		F1b_3'OP	ACTCGGCTTGCTGAAGTGC	696-714		
		F1b_5'IP	CAATTTCTGGCTATGTGCCC	1984-2003	57	54
		F1b_3'IP	TCCTTTCCACATTTCCAACAGCC	2023-2045		
Pol	B	PolB_5'OP	CTGAAAACAGGAAAATATGCA	3594-3614	58	56
		PolB_3'OP	TACTTGCCACACAATCATCACC	5079-5058		
		PolB_5'IP	GAGGGGTGCCCACACTAATGA	3620-3640	60	58
		PolB_3'IP	CTTTTCTTCTTGGCACTACTTTTA	4999-5022		
	C	F3_5'OP	GAAGGACAGTACTAAGTGGAG	2744-2764	62	59
		F3_3'OP	ACTGCATTCACCATACCT	2931-2948		
		F3_5'IP	TTATGTGCTGGTACCCATGA	4149-4168	58	55
		F3_3'IP	CTGGCTACATGGACTGCTAC	4452-4471		
Env	B	EnvB_5'OP	TAGAGCCCTGGAAGCATCCAGGAAG	5853-5877	57	57
		EnvB_3'OP	TTGCTACCTGTGATTGCTCCATGT	8913-8936		
		EnvB_5'IP	CACCTTAGGCATCTCCTATGGCAGGAAGAAG	5953-5983	63	63
		EnvB_3'IP	GTCTCGAGATACTCGTCCCACCC	8882-8904		
	C	F7_5'OP	GGAGATATAAGACAAGCACATTG	7194-7216	60	57
		F7_3'OP	GTGGAGGAGAATTTTTCTATTG	7357-7378		
		F7_5'IP	CTATCTGTTCCTTCAGCTACTGC	8685-8707	63	60
		F7_3'IP	CCCTGTCTTATTCTTCTAGGT	8753-8773		
Nef	B	NefB_5'OP	GATCCATTCGATTAGTGAACGGA	8455-8477	60	57
		NefB_3'OP	CTTATATGCAGCATCTGAGG	9497-9517		
		NefB_5'IP	TTCAGCTACGACCGATTGAGA	8520-8540	60	57
		NefB_3'IP	TGAGGGTTGGCCACTCC	9502-9686		
	C	F8_5'OP	GATCCATGAGATTAGTGAGCGGA	8455-8477	61	58
		F8_3'OP	TTCAGCTACCACCGATTGAGA	8520-8540		
		F8_5'IP	TGAGGGCTGGCCACTCC	9486-9502	61	58
		F8_3'IP	CTTATATGCAGCATCTGAGG	9498-9517		

*Annealing temperature during PCR (see Table 2.3)

Table 2.3. PCR cycling conditions to amplify Gag, Pol, Env and Nef.

Step Number	Temperature (°C)	Time
1	94	2 minutes
2	94	15 seconds
3	Temperature 1*	30 seconds
4	72	1 minute
5	Go to step 2, 19 more times	
6	94	15 seconds
7	Temperature 2*	30 seconds
8	72	1 minute
9	Go to step 6, 19 more times	
10	72	7 minutes
11	4	Forever
12	End	

Table 2.4. Sequencing PCR mastermix reaction.

Component	Volume (µl)	Final Concentration
5x Buffer	1.5	0.75x
Primer	2	0.01 mM
BigDye	0.5	1 Unit
Template DNA	2	Not applicable
ddH_2O	4	Not applicable

Table 2.5. PCR cycling conditions for sequencing.

Step Number	Temperature (°C)	Time
1	96	10 seconds
2	50	5 seconds
3	60	4 minutes
4	Go to step 1, 29 more times	
5	4	Forever
6	End	

Table 2.6. Sequencing primers for Gag, Pol, Env and Nef.

Region of genome	Clade	Primer name	Primer sequence (5'→3')	HXB2 coordinates
Gag	B	F1_5'IP	ACTAGCGGAGGCTAGAA	765-781
		F1_2	CCTGGAGGTTCTGCAC	1192-1207
		F1_3	GACACCAAGGAAGCTTTAG	1075-1093
		F1_4	ATCCCAGGATTATCCATC	1580-1597
		F1_5	CCAAGGGAAGTGACAT	1480-1496
		F1_3'IP	GCCACAATTGAAACACTT	1960-1977
	C	F1b_5'IP	CAATTTCTGGCTATGTGCCC	1984-2003
		F1b_2	CTGCAGAATGGGATAGTTACAT	1415-1437
		F1b_3	GACACCAAGGAAGCCTTAG	1075-1093
		F1b_4	CTCCCACTGGAAtAGGTG	1550-1567
		F1b_5	GGAACAAATAGCATGGATGAC	1521-1541
		F1b_3'IP	TCCTTTCCACATTTCCAACAGCC	2023-2045
Pol	B	PolB_5'IP	GAGGGGTGCCCACACTAATGA	3620-3640
		PolB_2	GATTCACTTTTATCTGGTTGTG	4072-4093
		PolB_3	CAGACTCACAATATGCATTAGG	4039-4060
		PolB_4	ACTGGCCATCTTCCTGCTA	4540-4558
		PolB_5	AGTGGATATATAGAAGCAGAAGT	4470-4492
		PolB_3'IP	CTTTTCTTCTTGGCACTACTTTTA	4999-5022
	C	F3_5'IP	TTATGTGCTGGTACCCATGA	4149-4168
		F3_2	TGGGTCATAATATACTCCATGTA	3490-3512
		F3_3	CAGACATAGTACCACTAACTGAA	3418-3440
		F3_4	CATAGAAAGTTTCTGCTCCT	3854-3873
		F4_5	CTCCCCTAGTAAAATTATGGTA	3808-3829
		F3_3'IP	CTGGCTACATGGACTGCTAC	4452-4471
Env	B	EnvB_5'IP	CACCTTAGGCATCTCCTATGGCAGGAAGAAG	5953-5983
		EnvB_2	GTACCGTCAGCGTTATTG	7825-7842
		EnvB_3	TCCTTGGGTTCTTGGGAGC	7780-7798
		EnvB_4	ACCAATTCCACAGATTTTCC	8222-8241
		EnvB_5	TGGGGATAAAGAAATTAGTAATTA	8115-8137
		EnvB_6	TCTTTTTTAGTTCCAGACCCCAA	8630-8652
		EnvB_3'IP	GTCTCGAGATACTCGTCCCACCC	8904-8882
	C	F7_5'IP	CTATCTGTTCCTTCAGCTACTGC	8685-8707
		F7_2	GTACCGTCAGCGTTATTG	7825-7842
		F7_3	TCCTTGGGTTCTTGGGAGC	7780-7798
		F7_4	ACCAATTCCACAGATTTTCC	8222-8241
		F7_5	TGGGATAAAGAAATTAGTAATTA	8115-8137
		F7_6	TCTTTTTTAGTTCCAGACCCCAA	8630-8652
		F7_3'IP	CCCTGTCTTATTCTTCTAGGT	8753-8773
Nef	B	NefB_5'IP	TTCAGCTACGACCGATTGAGA	8520-8540
		NefB_2	TCAGGGAAGTAGCCTTGTG	9146-9162
		NefB_3	TTTTTAAAAGAAAAGGGGGGAC	9064-9085
		NefB_3'IP	TGAGGGTTGGCCACTCC	9686-9502
	C	F8_5'IP	TGAGGGCTGGCCACTCC	9486-9502
		F8_2	ACCTGAGGTCTGACTGGAAAG	8997-9017
		F8_3	AGATAAACATGGGGCACTTAC	8907-8927
		F8_4	AGCAGTCTTTGTAATACTCCG	9395-9415
		F8_5	ACCTCAGGTACCTTTAAGAC	9009-9028
		F8_3'IP	CTTATATGCAGCATCTGAGG	9498-9517

2.5.2 Amplification of Gag-protease viral RNA for Sanger sequencing

Amplification of Gag-protease segment is crucial for generating a recombinant virus of NL4-3 backbone in Chapter 3 and 4. It is also used for confirming viral sequences after mutagenesis and harvesting for all fitness studies in Chapter 3, 4 and 5. SuperScript III One-Step RT-PCR System with Platinum Taq High Fidelity (Invitrogen) was used for the amplification as previously described (Gall *et al.,* 2012). Mastermix reaction composition, and PCR conditions are shown in **Table 2.7-2.8**. Subsequently cDNA products from RT-PCR were amplified as the second round (Bioline), which is detailed in **Table 2.9-2.10**. Universal primers costumed by Eurofin Genomics for both RT-PCR (first round) and cDNA amplification (2nd round) were listed in **Table 2.11**. Products were confirmed, purified and sequenced as described above in **Section 2.9.1**, using corresponding primers in **Table 2.6**.

Table 2.7. RT-PCR mastermix reaction for Gag-protease amplification.

Component	Volume (µl)	Final Concentration
2x Buffer	25	1x
Forward primer	1	0.2 mM
Reverse primer	1	0.2 mM
Platinum Taq HiFi	1	1 Unit
Template RNA	10	Not applicable
ddH_2O	12	Not applicable

Table 2.8. PCR cycling conditions to amplify Gag-protease.

Step Number	Temperature (°C)	Time
1	55	30 minutes
2	94	2 minutes
3	94	15 seconds
4	55	30 seconds
5	68	2 minutes
6	Go to step 3, 39 more times	
7	68	5 minutes
11	4	Forever
12	End	

Table 2.9. Mastermix reaction for 2nd round protease amplification.

Component	Volume (µl)	Final Concentration
10x NH_4 Reaction Buffer	10	1x
50mM $MgCl_2$	3	1 mM
10mM dNTP mix	2	0.2 mM each dNTP
Forward primer	2	0.2 mM
Reverse primer	2	0.2 mM
SSIII RT/Platinum Taq HiFi Mix	0.4	0.4 Unit
RT-PCR product	8	Not applicable
ddH_2O	72.6	Not applicable

Table 2.10. 2nd round PCR cycling conditions to amplify Gag-protease.

Step Number	Temperature (°C)	Time
1	94	2 minutes
2	94	30 seconds
3	60	30 seconds
4	72	2 minute
5	Go to step 2, 19 more times	
6	94	30 seconds
7	57	30 seconds
8	72	2 minute
9	Go to step 6, 19 more times	
10	72	7 minutes
11	4	Forever
12	End	

Table 2.11. PCR primers for amplifying Gag-protease.

Round in nested PCR	Primer name	Primer sequence (5'→3')	HXB2 coordinates
1st	623Fi	AAATCTCTAGCAGTGGCGCCCGAACAG	623-649
	2cRx	TTTAACCCTGCTGGGTGTGGTATYCCT	2825-2851
2nd	GagPro chimeric recomb_F_Toshi	GACTCGGCTTGCTGAAGCGCGCACGGCAAG AGGCGAGGGGCGGCGACTGGTGAGTACGC CAAAAATTTTGACTAGCGGAGGCTAGAAGGA GAGAGATGGG	695-794
	GagPro chimeric recomb_R_new	GGCCCAATTTTTGAAATTTTTCCTTCCTTTTC CATTTCTGTACAAATTTCTACTAATGCTTTTA TTTTTTCTTCTGTCAATGGCCATTGTTTAACT TTTG	2605-2704

2.5.3 Amplification of full-length viral RNA for ultra-deep sequencing

The deep sequencing in the following section were performed by Dr Jacqueline R Brener, Department of Paediatrics, University of Oxford.

Due to the polymorphism in the viral genome, primers used in Sanger sequencing might not bind to patient's sequences, in particular for the proteins located at the highly variable 3'end. Therefore, deep sequencing with universal primers binding to conserved regions was used in Chapter 5. Amplification of the full HIV genome in four segments was performed using SuperScript III One-Step RT-PCR (Invitrogen) as mentioned in **Section 2.5.2**. Mastermix composition and primers (Eurofins Genomics) are detailed in **Table 2.12** and **2.13**. Mastermix (20µl/well) was aliquoted into a 96-well Low Profile PCR plate; 5µl/well RNA template (negative control – water) was added and PCR was run in a thermocycler (**Table 2.14**). To confirm, PCR products were run on 1% agarose gel with 240V, 300mA for 1.5 hours and bands were visualised under UV light. Amplicons were then combined into one pool (10µl of each amplicon) and ultra-deep sequenced at in collaboration with Dr Astrid Gall and Prof. Paul Kellam at the Wellcome Trust Sanger Institute and Dr Rebecca Batorsky and Dr Todd Allen at the Ragon Institue.

Table 2.12. RT-PCR for deep-sequencing mastermix composition.

Component	Volume (µl)	Final Concentration
2x Reaction Mix	12.5	1.25x
Pan-HIV-1-Mix 1, 2, 3 or 4*	2	10 pmol/µl
SSIII RT/Platinum Taq HiFi Mix	1	1 Unit
ddH_2O	4.5	Not applicable

*Pan-HIV-1 primer mixes for each of the 4 amplicons were prepared by mixing forward and reverse primers in T0.1E buffer (Sigma) at a final concentration 10 pmol/µl (Gall *et al.,* 2012)

Table 2.13. RT-PCR for deep-sequencing primers.

Primer name	Primer sequence (5'→3')	HXB2 coordinates
Pan-HIV-1_1F	AGCCYGGGAGCTCTCTG	480-496
Pan-HIV-1_1R	CCTCCAATTCCYCCTATCATTTT	2385-2407
Pan-HIV-1_2F	GGGAAGTGAYATAGCWGGAAC	1485-1505
Pan-HIV-1_2R	CTGCCATCTGTTTTCCATARTC	5037-5058
Pan-HIV-1_3F	TTAAAAGAAAAGGGGGGATTGGG	4783-4805
Pan-HIV-1_3R	TGGCYTGTACCGTCAGCG	7831-7848
Pan-HIV-1_4F	CCTATGGCAGGAAGAAGCG	5967-5985
Pan-HIV-1_4R	CTTWTATGCAGCWTCTGAGGG	9496-9517

Table 2.14. RT-PCR for deep-sequencing cycling conditions.

Step Number	Temperature (°C)	Time
1	50	30 minutes
2	94	2 minutes
3	94	15 seconds
4	58	30 seconds
5	68	4.5 minutes
6	Go to step 3, 39 more times	
7	68	10 minutes
8	10	Forever
9	End	

2.5.4 Ultra-deep sequencing and de novo assembly of genomes

Each pool of amplicons was given a unique bar code for Illumina library preparation and sequenced using MiSeq 300-bp paired-end technology. Consensus sequences were generated by *de novo* assembly, i.e. without a reference sequence, using Iterative Virus Assembler (Hunt *et al.,* 2015). This was performed by Dr Rebecca Batorsky at the Ragon Institue.

2.5.5 Phylogenetic analysis

Maximum-likelihood trees of cohort Gag/Gag-protease sequences in all Chapters were generated using the general time reversible model of nucleotide substitution, which was constructed with 1,000 bootstrap replicates to identify phylogeny of the sequences and to exclude the possibility of contamination with laboratory viral strains.

This was completed by using Mega 7.0.14 software and viewed using FigTree v1.4.0 software. HIV subtypes were further confirmed with NCBI (https://www.ncbi.nlm.nih.gov/projects/genotyping/formpage.cgi) and REGA (http://bioafrica.net/rega-genotype/html/subtypinghiv.html) HIV genotyping tools.

2.6 Site-Directed Mutagenesis

2.6.1 Plasmids preparation

Site-directed mutagenesis was performed in different plasmids for generating viruses carrying certain mutations which might affect viral fitness (Chapter 3, 4 and 5). In total, three plasmids were used (**Table 2.15**), whose details are listed in each Chapter-specific materials and methods sections. SK-254-TOPO plasmids and the p6.5 plasmid were kindly provided by Dr Jaclyn Wright and Dr Julia Prado respectively. All plasmids were amplified and purified by using either QIAGEN Plasmid Maxi Kit (QIAGEN) or QIAprep Spin Miniprep Kit (QIAGEN) depending on the amount and concentration needed, followed by the manufacturer's instructions. Plasmid maps are shown in Appendix.

Table 2.15. Plasmids used in site-directed mutagenesis.

Plasmid name	Clade	Size (kb)
pNL4-3	B	14.8
SK-254-TOPO*	C	12.3
p83-2	B	8.1
p6.5	C	8.1

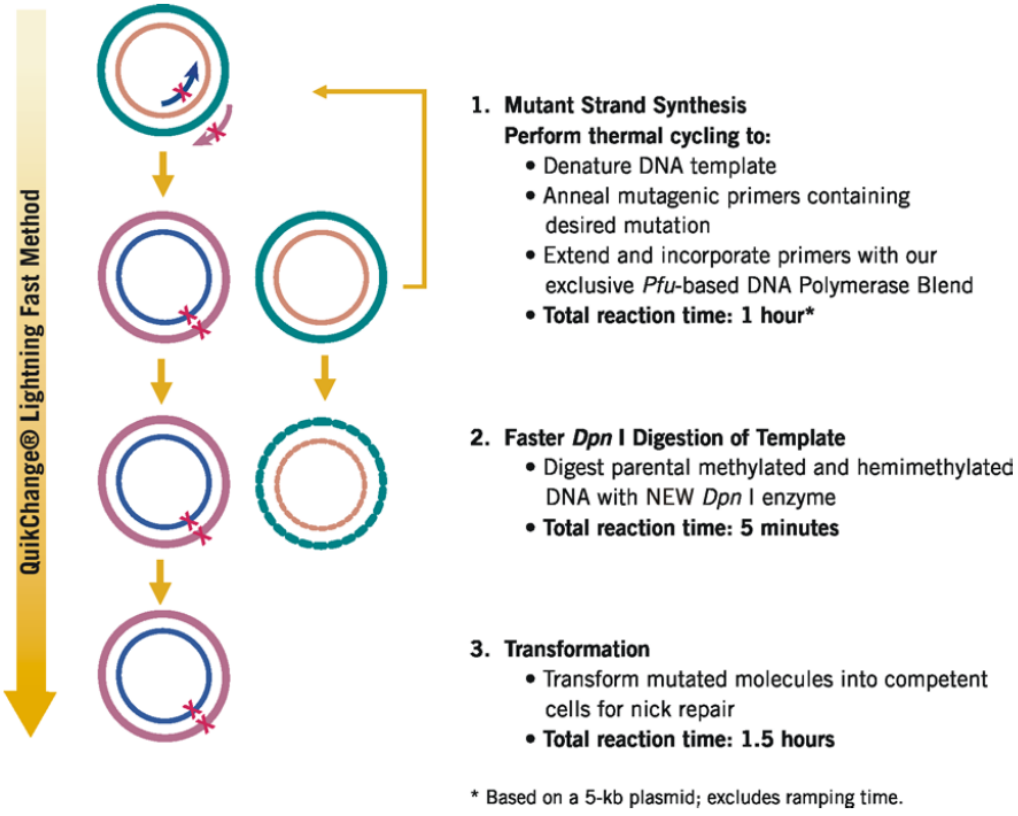
*SK-254-TOPO plasmid was only used as a vector for carrying HIV C-clade patient-derived sequence, therefore after the mutagenesis, the Gag-protease sequence was amplified by running 2nd Round PCR as described in section 2.5.2 (**Table 2.9-2.11**). The PCR products were together electroporated with ΔGag-protease pNL4-3.

2.6.2 Costumed primer design

All primer sequences and details are listed in each Chapter-specific materials and methods sections. In general, primers were costumed by using QuikChange Primer

Design online tool (http://www.genomics.agilent.com/primerDesignProgram.jsp). The primer sequences were visualised in SnapGene Viewer software v3.1.4 and checked for details on IDT OligoAnalyzer 3.1 (http://eu.idtdna.com/calc/analyzer) as well as Align Sequenced Nucleotide BLAST (NCBI). For QuikChange kit recommendations, primer length should be less than 45bp, GC content >40%, melting temperature up to 78°C, and avoid forming secondary structure. Since the QuikChange Primer Design does not support costuming primer that contains two or more intended mutations, primers were designed with Ms. Krystyna Nahlik from Agilent Technology under the circumstances when two mutations were located too close to each other (<10 aa).

A



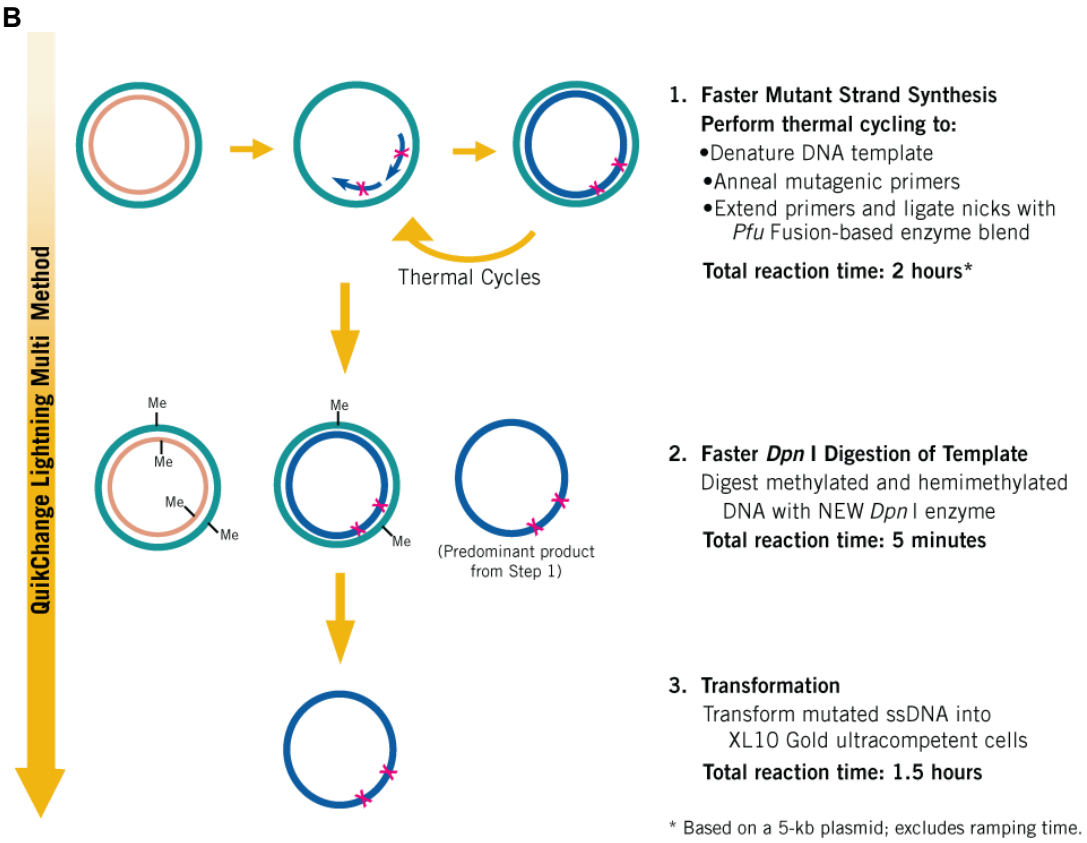


Figure 2.2. Overview of the QuikChange lightning multi site-directed mutagenesis method (directed from Agilent Technology user manuals). (A) QuikChange Site-Directed Mutagenesis **(B)** QuikChange Multi Site-Directed Mutagenesis.

2.6.3 Site-Directed mutagenesis procedure

For introduction of a single mutation and multiple mutations, QuikChange Lightning Site-Directed Mutagenesis Kit and QuikChange Lightning Multi Site-Directed Mutagenesis Kit (Agilent Technologies) were used according to manufacturer's instructions. **Figure 2.2** shows the overview of the mutagenesis process. Main difference between the two kits is that both forward and reverse primers were required for the single mutation, whereas when introducing multiple mutations, only forward primer was used. On Day 1, mutant strand synthesis PCR was run in a thermocycler. Mastermix compositions and cycling conditions were listed in **Table 2.16 and 2.17**. A positive control (positive control plasmid) and a negative control

(water) were included in each run. The methylated original template was then digested by adding 1µl or 2µl DpnI enzyme for single or multi site-directed mutagenesis respectively to the PCR products at 37°C for 5 minutes.

Table 2.16. Mutagenesis PCR mastermix composition.

Number of mutation(s)	Component	Volume of sample (µl)
Single	10x Reaction Buffer	5
	QuikSolution reagent	1.5
	ds-DNA template (100 ng)*	1
	Forward mutagenic primer (100 ng/µl)	1.25
	Reverse mutagenic primer (100 ng/µl)	1.25
	dNTP mix	1
	QuikChange Lightning Enzyme	1
	ddH_2O	top up to 51
Multi	10x QuikChange Lightning Multi reaction buffer	2.5
	QuikSolution*	0.75
	ds-DNA template (100 ng)*	1
	Mutagenic primers (100 ng each)*	1-3
	dNTP mix	1
	QuikChange Lightning Multi enzyme blend	1
	ddH_2O	top up to 25

*When DNA template was >5kb, QuikSolution and 100ng template were added. To maximise mutagenesis effeciency, no more than 3 mutations were introduced for each run. 100ng of each primer was added in this condition.

For the transformation step, XL 10-Gold ultracompetent cells were thawed on ice and were aliquoted into 14-ml Falcon polypropylene round-bottom tube (BD Bioscience) as 45µl each. Cells were incubated with 2µl of the β-ME mix on ice for 10 minutes and swirled gently every 2 minutes. The Dpn I-treated DNA from both controls and samples was transferred to separate aliquots of the cells at the amount of 2µl and 1.5µl for single and multi mutagenesis. The transformation reactions were swirled gently to mix and incubated on ice for 30 minutes. Tubes were then heat-pulsed in a 42°C water bath for 30 seconds and incubated again on ice for 2 minutes. 1x LB broth (Sigma) was made, autoclaved and preheated (42°C) beforehand. Each tube was added with 500µl broth and put into a shaking incubator at 37°C for 1 hour at 225 rpm. For positive control, 10µl of the reaction mix was plated on a pre-made

autoclaved 1x LB agar (Sigma) containing 50μg/ml ampicillin (Sigma), while 10μl/plate and 100μl/plate were added for each sample. All the plates were incubated at 37°C with 5% CO_2 overnight (>16 hours). On Day 2, five clearly isolated colonies were picked per sample, and were added into 5ml LB broth in the 14ml round-bottom tube. All tubes were incubated in shaking incubator at 37°C overnight at 225rpm. Plasmid DNA was extracted on Day 3 by using QIAprep Spin Miniprep Kit (QIAGEN) and dissolved in 50μl DNase/RNase-free water. The Gag-protease sequence of each sample was checked by Sanger sequencing as previously described in **Section 2.1**.

Table 2.17. PCR for mutagenesis cycling conditions.

Number of mutation(s)	Step Number	Temperature (°C)	Time
Single	1	95	2 minutes
	2	95	20 seconds
	3	60	10 seconds
	4	68	30 seconds/kb of plasmid*
	5	Go to step 2, 17 more times	
	6	68	5 minutes
	7	4	Forever
	8	End	
Multi	1	95	2 minutes
	2	95	20 seconds
	3	55	30 seconds
	4	65	30 seconds/kb of plasmid*
	5	Go to step 2, 29 more times	
	6	65	5 minutes
	7	4	Forever
	8	End	

*For example, a 5-kb plasmid requires 2.5 minutes per cycle at 68°C

2.7 Plasmid Digestion

2.7.1 pNL4-3 linearisation

In Chapter 3 and 4, to generate chimeric viruses containing SK-254-TOPO plasmid (**Table 2.15**) or patient's Gag-protease sequence, ΔGag-protease pNL4-3 was

required as backbone. ΔGag-protease pNL4-3 is a deleted version of pNL4-3, constructed that lacks the entire Gag and Protease region (Stratagene Quick-Change XL Kit), replacing this region with a BstE II restriction site. Before used, this plasmid must be linearised using the BstE II enzyme (New England Biolabs). To confirm the linearisation, both pNL4-3 and ΔGag-protease pNL4-3 were digested with BestE II and/or Hind III enzyme (New England Biolabs). Conditions were detailed in **Table 2.18**. Plasmids were incubated at 60°C for 2 hours. 5µl of each reaction was used to run with Hyperladder I on 1% agarose gel at 180V for 50 minutes. Bands were visualised under UV light. BstEII-linearised and Hind III-digested ΔGag-protease pNL4-3 should have one and four distinctive bands at the correct sizes respectively, whereas linearised pNL4-3 will have one band sized equally as undigested and Hind III-digested pNL4-3 has five bands (**Figure 2.3**). The full reaction mix of linearized plasmid band then was run on a gel again and extracted using QIAquick Gel Extraction Kit (QIAGEN).

Table 2.18. Plasmid digestion reaction (for 50 co-transfections)

Component	Volume (µl)	Final Concentration
Plasmid DNA (1,000 µg/ml)	500	0.75 µg/µl (500 µg DNA)
10x NEB Buffer	65	1x
100x BSA	6	1x
Enzyme (10 U/µl)	100	2 U/µg DNA

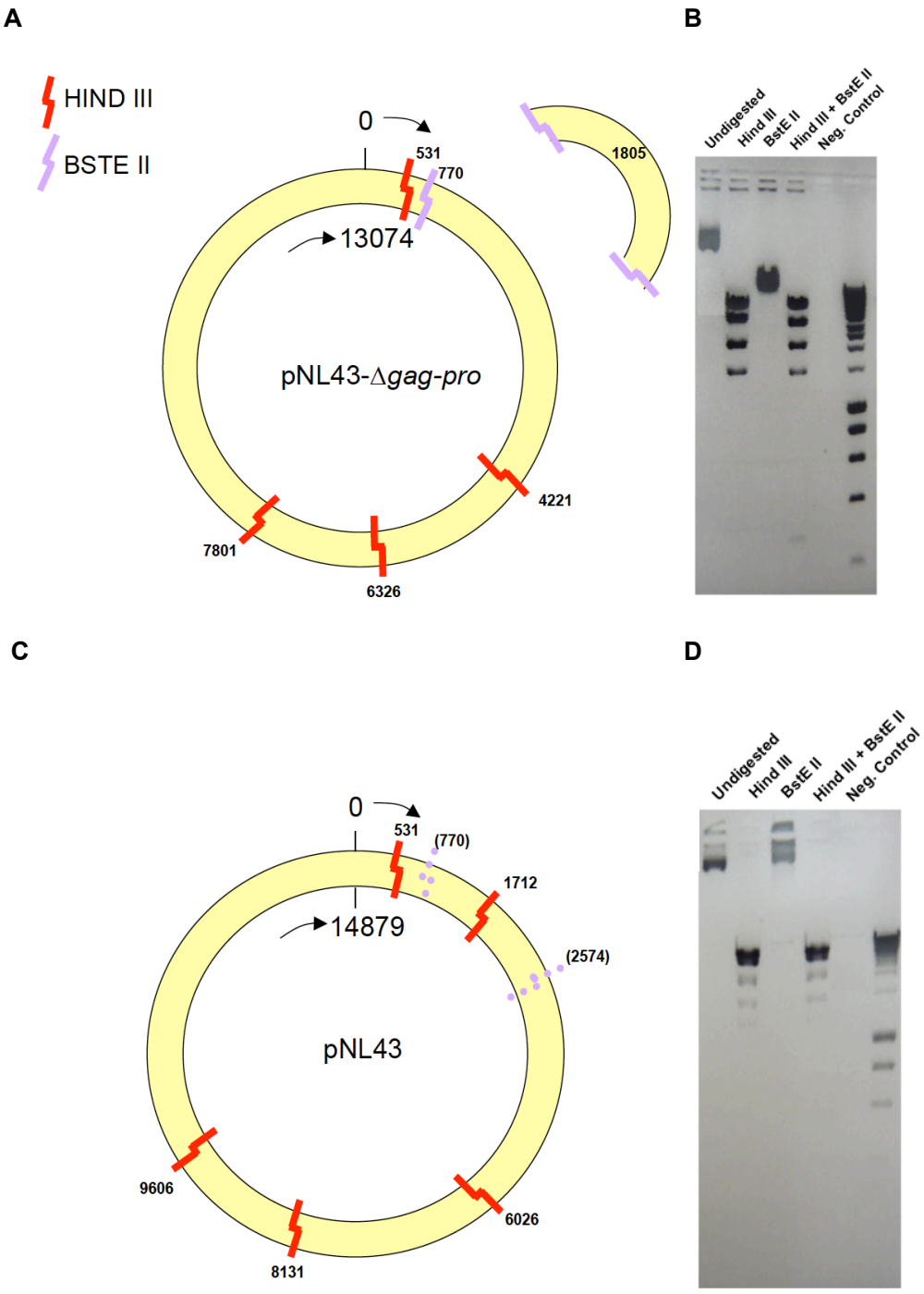


Figure 2.3. Plasmid map and electrophoresis of Hind III and BstE II enzyme digestion. **(A)** ΔGag-protease pNL4-3 enzyme restriction sites. **(B)** Electrophoresis bands of ΔGag-protease pNL4-3. **(C)** pNL4-3 enzyme restriction sites. **(D)** Electrophoresis bands of pNL4-3.

2.7.2 p83-2/p6.5 and p83-10$_{GFP}$ linearisation

In Chapter 5, a different system of fitness assay was performed. Two plasmids (p83-2/p6.5 and p83-10$_{GFP}$) recombine to form a single template for the whole virus

production. Unlike, pNL4-3, p83-10_{GFP} encodes GFP, hence MT-4 cell line was used instead of GXR cells. To linearise, 30µl of each plasmid (~6-7 µg) was incubated with 2µl EcoRI at 37°C for 2 hours. The full reaction mix was run on a gel. The digested band was extracted using QIAquick Gel Extraction Kit (QIAGEN).

2.8 Immortal Cell Line

HIV-permissive cell lines GXR (CEM-GXR25) and MT-4 (**Table 2.19**) were grown in suspension in R10 and split every 2-3 days and 24 hours prior to infection or transfection.

Table 2.19. HIV-permissive cell lines used for *in vitro* studies.

Cell line	Description	Encoding GFP	Reference
GXR	Derived from CEM. CEM cells are lymphoblasts isolated from the peripheral blood buffy coat collected from a female child with acute leukemia. Permissive to both CXCR4- and CCR5-tropic HIV-1 strains.	Yes	Gervaix *et al.,* 1997 and Brockman *et al.,* 2006
MT-4	T-cell line established from cord blood lymphocytes cocultured with leukemic cells from patients with adult T-cell leukemia. Permissive to CXCR4-tropic HIV-1 strains.	No	Harada, *et al.,* 1985; Pauwels *et al.,* 1987 and Larder *et al.,* 1989

2.9 Viral Fitness Assay

2.9.1 Electroporation

To generate chimeric virus, plasmids were prepared for transfection, detailed in **Table 2.20**. GXR and MT-4 cells were suspended in R10 media and Electroportation media (EM) respectively. Together with plasmids, co-transfection mixture was added in a 4 mm cuvette without creating bubbles. For each transfection, a negative control containing no plasmids was included. After setting the electroporation in 'exponential programme' in a BioRad GenePulsar II (BioRad), each cuvette was inserted into the electroporator with the metal parts attached and pulsed. The sample should appear

frothy after the pulsing. GXR and MT-4 cells were allowed to stand for 45 and 5 minutes at the room temperature, then transferred to a T25 flask containing 5ml R10 and 1 million healthy MT-4 cells in 10 ml, respectively. All the cultures were incubated at 37°C with 5% CO_2. Another 5ml of R10 was added into each T25 flask GXR culture on Day 4 after the electroporation.

Table 2.20. Conditions of transfection for each sample

Co-transfection mixture	Amount	Media	GenePulsar setting	Resting time
pNL4-3 / ΔGag-protease pNL4-3	10 µg	R10	300 V 500 µF infinite resistance	45 minutes
Gag-protease PCR product	2.5 µg			
GXR cells	800 (2 million cells)			
p83-2/p6.5	5	EM	250 V 950 µF infinite resistance	5 minutes
p83-10_{GFP}	5			
MT-4 cells	500 (5 million cells)			

2.9.2 Monitoring viral growth

Viruses were measured according to the GFP+ percentage from Day 3 and Day 9 for GXR and MT-4 cells respectively, then followed by the frequency of every 2 days. Each time, culture was mixed; 2ml was taken into a FACS tube (BD Biosciences) and replaced by fresh R10. Cells were pelleted by spinning at 2,000rpm for 5 minute. Supernatant was aspirated, while cells were fixed in 250 µl of 2% paraformaldehyde solution (Sigma) at 4°C for 20 minutes. Cells were subsequently pelleted at 2,000rpm for 5 minutes; resuspended in 200 µl PBS, and run on a MACSQuant Analyzer 10 flow cytometer (Miltenyi Biotec) B1 channel. Once the GFP+ percentage was reached 30% and 60-80% for GXR and MT-4 respectively, supernatant was collected by spinning down the culture at 2,000rpm for 10 minutes. After filtering through a 0.22µm filter (Sigma), the supernatant was 1ml-aliquoted in 1.5ml cryovial and stored at -80°C until use.

2.9.3 Monitoring viral growth titration of cultured virus using an endpoint dilution assay

To titre the harvested viruses, an endpoint dilution assay was run based on $TCID_{50}$ determination. A 96-well U-bottomed plate (Fisher Scientific) was set up as **Figure 2.4** for each virus. 200µl PBS was added to the outside wells and 100µl of R10 containing 30,000 GXR or MT-4 cells were applied to all other wells. 25µl of viral stock was added and mixed in the sample wells to the first column in six replicates. A 1:5 serial dilution was carried on by removing 25µl from one column of wells into the adjacent column. In the final sample column wells, 25µl was discarded. Negative wells contained no virus but cells. The plate was incubated at 37°C with 5% CO_2 for five days. All the viruses run for fitness assay were titred at the same time using same batch of cell line. For the GFP detection, cells were pelleted by spinning down at 2,000 rpm for 5 minutes, followed by one wash of 200µl PBS spun at 2,000rpm for 5 minutes. Cells were fixed in 100µl 2% paraformaldehyde solution at 4°C for 20 minutes. The plate was spun again at same condition and 100µl of PBS was added to each sample and negative wells. GFP expression was detected using a MACSQuant Analyzer 10 flow cytometer.

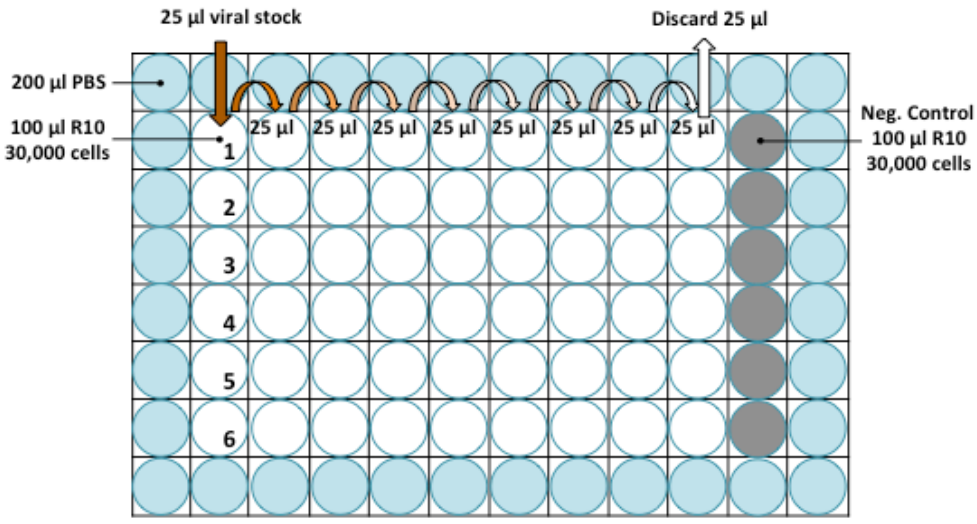


Figure.2.4. Layout of 96-well plate for endpoint dilution assay. Blue wells represent 200 µl PBS; white wells are the sample in six replicates; and grey wells are the negative controls without adding viruses.

The $TCID_{50}$ (tissue culture infective dose at which 50% of cells are infected, log U/ml) was determined by the Excel sheet developed by Lindenbach (Lindenbach, 2009) on the basis of Reed-Muench Calculation (Reed and Muench, 1938). The dilution factor that gives higher than 50% infectivity and the dilution factor that gives lower than 50% infectivity are used to determine the $TCID_{50}$ (which lies between these 2 dilutions.

1. Calculate the interpolated value/proportionate distance of the 50% endpoint (I)

   from the dilution factor above 50% and the dilution factor below 50%

$$I = \left\{ \frac{\text{(\% of wells infected at dilution above 50\% - 50\%)}}{\text{(\% of wells infected at dilution above 50\% - \% of wells infected at dilution below)}} \right\}$$

2. Where h=dilution factor:

$$\text{50\% endpoint titer} = 10^{\text{log total dilution above 50\% - (I x log h)}}$$

3. $TCID_{50}$ (U/volume of virus dilution tested) = 1/50% endpoint titre. This value can then be converted to log units/ml.

The multiplicity of infection (MOI), ratio of infectious particles to target cells, was then calculated: MOI = volume of virus stock × $TCID_{50}$ (log U/ml) / number of cells to be infected.

2.9.4 Viral replicative capacity assay

To measure the viral replicative capacity (VRC), a final 10-day assay was run to calculate the exponential phase of viral growing curve. A negative control without virus was included and all the viruses were set in triplicate. One million of either GXR or MT-4 cells were prepared in 100µl R10 in a FACS tube. Virus at MOI 0.001 was

added and made up to 1ml with R10 per tube. All the samples were incubated at 37°C with 5% CO_2 for 4 hours, vortexed gently every 30 minutes to ensure cells remain in suspension. Cells were washed 3 times in PBS to remove cell-free virus, resuspended in 2ml R10, and transferred to 24-well plate. Viral spread was measured at the same time each day, from Day 2 to Day 10. Cells were mixed well before transferring 200µl (~100,000 cells) to a 96-well U-bottom plate and washed once in 200µl PBS. Cell culture was replaced with 200µl fresh R10. After the wash, Live/Dead marker (Invitrogen) was used for staining at 1:50 in 50µl, 20 minutes at room temperature. Another 150µl PBS was added subsequently and the plate was spun at 2,000rpm for 5 minutes. Cells were resuspended in 100µl 2% formaldehyde, fixed for 20 minutes at room temperature, and spun down again in the same condition. 100µl PBS was added in each well after the spin and the plate was ready for running on MACSQuant Analyzer 10 flow cytometer. Data were analysed using Flowjo version 9.9. Gating strategy on GFP+ cells is shown in **Figure 2.5**. Further analyses were specified in Chapter 3, 4 and 5.

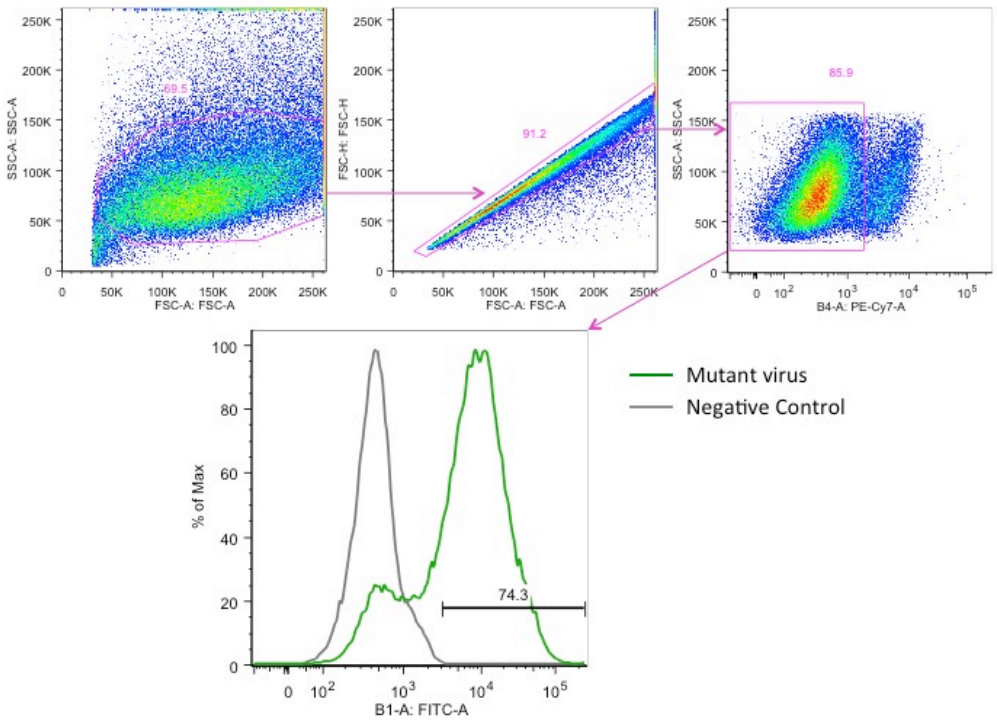


Figure.2.5. Representative example of the gating strategy used in fitness assay. Gated on live cells (SSC/FSC), singlets (FSC-H/FSC-A), live/dead (SSC/PE-Cy7), FITC+ cells. Negative as <0.1% GFP+. Compensation was not necessary due to the non-overlapping emission spectra of FITC (B1) and PE-Cy7 (B4).

2.10 Media Composition

- **R10:** RPMI-1640, 10% heat-inactivated FCS (Sigma) filtered through 0.45μm (Sigma), 1% L-glutamine (Sigma) and 1% penicillin-strptomycin (Sigma).
- **Electroporation Media:** RPMI-1640 and 1% L-glutamine.

2.11 Statistical Analysis

All analyses were performed in Prism (v6.0c, GraphPad). Data were first run to determine the distribution was parametric or nonparametric (D'Agostino and Pearson's omnibus normality test). For two-group analyses, student's *t*-test (parametric) or Mann-Whitney U-test (nonparametric) was performed; for ≥3 group amalyses, one-way ANOVA (parametric) or Kruskal-Wallis (non-parametric) followed by post-hoc test was performed. Nominal data categorized in two distinct ways were compared by Fisher's exact test. *p* values <0.05 were considered significant. Post-hoc analysis including Bonferroni's correction and Steel test for corrected *p* value was performed, depending on the sample characteristics, sample size and *p* value distribution. Strength of association between two variables was analysed by Spearman (parametric) and Pearson (nonparametric) correlation. The viral replication capacity was defined by the log calculation of the mean slope of exponential growth in Microsoft Excel (Mac 2011). In the association between CD4 counts/viral loads and HLA-B alleles, only the alleles that are expressed by more than three subjects in the cohort were included in the analysis. HLA Further and detailed analyses were specified in each Chapter.

CHAPTER 3: PAEDIATRIC NON-PROGRESSION FOLLOWING GRANDMOTHER-TO-GRANDCHILD HIV TRANSMISSION

[HLA-B*81-related work of this chapter has been published in **Tsai M,** Muenchhoff M, Adland E, Carlqvist A, Roider J, Cole D, Sewell A, Carlson J, Ndung'u T, Goulder P. 2016. Paediatric non-progression following grandmother-to-child HIV transmission. Retrovirology 13(1):65.]

3.1 Introduction

In the absence of antiretroviral therapy (ART), HIV infected children normally progress to disease rapidly, approximately 50% developing AIDS by 1 year of life, and by 2 years of life 50% have died (Prendergast *et al*., 2012). In the absence of ART, the median times to AIDS and death in adult infection are, respectively, 10 and 11 years (Prendergast *et al*., 2012). For these reasons, in 2013 the WHO have recommended that ART should be initiated in all HIV-infected children aged 5 years or above, and also in HIV-infected children whose absolute CD4 counts are <500 cells/mm^3 (WHO.2013. Consolidated ART guidelines) There remain, however, a number of ART-naïve, HIV-infected children, aged >5 years whose absolute CD4 counts are >500 cells/mm^3 and who represent approximately 10-15% of infected children, who are relative slow progressors (Mphatswe *et al*., 2007; Blanche *et al*., 1997; Jourdain *et al*., 2007). Approximately one-third of the children within this group, therefore around 5% of all infected children, have absolute CD4 counts of >750 cells/mm^3, that are therefore indistinguishable from those of age-matched HIV-uninfected children (10^{th}-90^{th} centiles; Shearer *et al*., 2003; Lugada *et al*., 2004). Study of these slow-progressor children provides the opportunity to gain insights into mechanisms of non-pathogenicity in HIV infection that appear to differ substantially from those operating in adult infection (Muenchhoff *et al*., 2014; Goulder *et al*., 2016).

Long-term non-progressing HIV-infected adults tend to maintain high CD4 counts in the setting of low viral loads, so-called 'elite' controllers having undetectable viraemia (Pereyra *et al*., 2009). Protective HLA class I molecules such as HLA-B*27, HLA-B*57, HLA-B*58:01 and HLA-B*81:01 are highly enriched amongst adult elite controllers (Kiepiela *et al*., 2004; Pereyra *et al*., 2009; Leslie *et al*., 2010). The mechanisms proposed for HLA-associated control of HIV include the ability of these 'protective' HLA molecules to present Gag-specific epitopes for recognition by CD8+ T-cells to facilitate rapid killing of virus-infected target cells (Goulder and Walker, 2012; Kløverpris *et al*., 2016). In these cases, the location of the epitopes, often in highly conserved regions of the capsid protein, may compromise the ability of the virus to select escape mutants that do not incur substantial costs to viral replicative capacity as a consequence (Miura *et al*., 2009a). In contrast, these 'protective' HLA-B molecules are not expressed at high frequency in long-term non-progressing children (Adland *et al*., 2015). The reasons for this apparent lack of impact of 'protective' HLA in children may relate to immune ontogeny (Prendergast *et al*., 2012; Muenchhoff *et al*., 2014) and the inability of young children to generate especially strong Th1 type CD4 T-cell responses. Although HIV-specific CD8+ T-cell responses are detectable from birth in *in utero* infected children (Thobakgale *et al*., 2007; Luzuriaga *et al*., 1995), these responses usually do not have a significant impact on viral replication, at least for the first years of life (Adland *et al*., 2015).

In contrast to HLA type, the viral replicative capacity of the transmitted virus has a significant impact on disease progression in both paediatric and adult infection (Goepfert *et al*., 2008; Prince *et al*., 2012; Adland *et al*., 2015). We here studied the role of HLA and viral replicative capacity in a family including three HIV-infected individuals, all expressing HLA-B*81:01, an HLA allele that is protective in adult infection (Kiepiela *et al*., 2004; Pereyra *et al*., 2009; Leslie *et al*., 2010). HLA-B*81:01 belongs to the B07 supertype family; the features of its peptide-binding groove are

listed in **Table 3.1**. The B pocket prefers Proline only and the F pocket can accommodate aromatic, aliphatic and hydrophobic molecules (F, Q, Y, L, I, V, M, W, A) (Sidney *et al.,* 2008). The dominant epitopes restricted by this allele in the HIV genome are Gag TL9 (180-188, TPQDLNTML), Pol SM9 (311-319, SPAIFQSSM), Vpr FL9 (34-42, FPRIWLHGL) and Nef TM9 (71-79, TPQVPLRPM) (Carlson *et al*., 2012a).

Table 3.1. Characteristics of B*81:01-restricted epitope.

	Peptide-Binding Motif				
Position	**P1**	**P2**	**P3**	**P5-7**	**PC**
Binding pocket	A	B	D	C	F
B*81:01	-	P	-	-	L/M/I/F

Table is based on The HLA Factsbook (Marsh *et al*., 2000).

Two of these HLA-B*81:01-positive family members were slow progressor children, the third was the mother of one child and grandmother of the other (**Figure 3.1 B**). In this case, it appears that the latter child was infected by breast milk transmission from the grandmother.

3.2 Chapter-Specific Materials and Methods

3.2.1 Study subjects

The family trio was selected from HIV-infected Paediatric Slow Progressor Cohort (**Section 2.1**). In total, 47 treatment naïve mothers and 84 treatment naïve children chronically infected with HIV C-clade were recruited from Kimberley, South Africa.

Additional viral sequence analyses were performed on previously described, multi-centre cohorts (**Section 2.1**): 1470 African clade C Gag sequences from cohorts based in Durban (Matthews *et al*., 2008), Bloemfontein (Huang *et al*., 2009) and Kimberley (Matthews *et al*., 2011), South Africa, Zambia and the Thames Valley area of the United Kingdom (Matthews *et al*., 2011).

3.2.2 Site-directed mutagenesis of NL4-3, SK-254 and SK-254(M)

T186S, L188F, L188F/T190I mutations of HIV-1 Gag sequence were introduced into the HIV-1 subtype B NL4-3 plasmid (pNL4-3) and a patient-derived subtype C HIV-1 Gag-protease sequence (SK-254, GenBank accession number HM593258) respectively by using QuikChange Lightning site-directed mutagenesis kit (Agilent technologies) along with custom-designed mutagenesis forward and reverse primers listed in **Table 3.2**. In addition, considering the six SK-254 variants in p24 Gag might affect viral fitness, we used QuikChange Lightning Multi site-directed mutagenesis kit (Agilent technologies) to retro-mutate these variants to Concensus C. Hence, the p24 Gag in this modified SK-254 sequence [SK-254(M)] is the same as Consensus C sequence. Note that some primers contain two mutations due to the proximity of the sites of mutation.

3.2.3 Phylogenetic analysis

Forty-three Gag-protease C-clade reference sequences from Durban, South Africa, from the same locality as the study subjects GM (the grand-mother), GD (the grand-daughter) and D2 (the second, HIV-infected daughter of the grand-mother: **Figure 3.1 B**), were included as reference sequences in a maximum likelihood phylogenetic tree.

3.2.4 Structural modelling

The structural modeling and analysis were undertaken by Dr David K Cole, Sewell Group, Cardiff University School of Medicine.

The p24-Gag protein (from PDB: 1GWP) was used to model known mutations in the virus. Sequences were adjusted with COOT (Emsley and Cowtan, 2004) and graphical representations were prepared with PYMOL (The PyMol Molecular Graphics System, Schrodinger, LLC).

Table 3.2. HLA-B*81:01 custom-designed mutagenesis forward and reverse primers.

HXB2 position in Gag	Clade	Mutant	Primer direction	Primer sequence (5'→3')
174	C	T→I	Forward	G CAA ATG GTA CAC CAA GCC **ATA** TCA CCT AGA ACT TTG AAT G
			Reverse	C ATT CAA AGT TCT AGG TGA **TAT** GGC TTG GTG TAC CAT TTG C
186	B	T→S	Forward	GCC ACC CCA CAA GAT TTA AAT **AGC** ATG CTA AAC ACA GT
			Reverse	AC TGT GTT TAG CAT **GCT** ATT TAA ATC TTG TGG GGT GGC
	C	T→S	Forward	ACC CCA CAA GAT TTA AAC **AGC** ATG TTA AAT ACA GTG GGG
			Reverse	CCC CAC TGT ATT TAA CAT **GCT** GTT TAA ATC TTG TGG GGT
188	B	L→F	Forward	CC CCA CAA GAT TTA AAT ACC ATG **TTC** AAC ACA GTG GGG GGA CAT C
			Reverse	G ATG TCC CCC CAC TGT GTT **GAA** CAT GGT ATT TAA ATC TTG TGG GG
	C	L→F	Forward	CAA GAT TTA AAC ACC ATG **TTC** AAT ACA GTG GGG GGA CAT C
			Reverse	G ATG TCC CCC CAC TGT ATT **GAA** CAT GGT GTT TAA ATC TTG
188/190	B	L→F/ T→I	Forward	CC CCA CAA GAT TTA AAT ACC ATG **TTC** AAC **ATA** GTG GGG GGA CAT CAA GCA
			Reverse	TGC TTG ATG TCC CCC CAC **TAT** GTT **GAA** CAT GGT ATT TAA ATC TTG TGG GG
	C	L→F/ T→I	Forward	ACC CCA CAA GAT TTA AAC ACC ATG **TTC** AAT **ATA** GTG GGG GGA CAT C
			Reverse	G ATG TCC CCC CAC **TAT** ATT **GAA** CAT GGT GTT TAA ATC TTG TGG GGT
252/256	C	G→S/ I→V	Forward	AA ATA GCA TGG ATG ACT **AGT** AAC CCA CCT **ATC** CCA GTG GGA G
			Reverse	C TCC CAC TGG **GAT** AGG TGG GTT **ACT** AGT CAT CCA TGC TAT TT
319	C	D→E	Forward	ACA CAA GAT GTA AAA AAT TGG ATG ACA **GAT** ACC TTG TTG GTC CAA
			Reverse	TTG GAC CAA CAA GGT **ATC** TGT TAC CCA ATT TTT TAC ATC TTG TGT
340/342	C	A→G/ S→T	Forward	C ATT TTA AGG GCA TTA GGA CCA **GGA** GCT **ACG** TTA GAA GAA ATG ATG ACA GCA TG
			Reverse	CA TGC TGT CAT CAT TTC TTC TAA **CGT** AGC **TCC** TGG TCC TAA TGC CCT TAA AAT G

*Mutant amino acids are shown in bold in the primer sequences.

3.3 Results

3.3.1 Identification of a paediatric slow progressor infected via grandmother-to-grandchild transmission

We have previously defined 'paediatric slow progressors' as ART-naïve, HIV-infected children who have not progressed to meet clinical or CD4 count criteria for ART initiation (Adland *et al*., 2015). The CD4 count criteria for ART initiation in children

aged >5 years in South Africa until 2013 was an absolute CD4 count of ≤350 cells/mm^3 (Adland *et al.*, 2015). Within a study cohort of paediatric slow progressors in Durban, South Africa, a female ('GD') was enrolled at age 9.1 years, with an absolute CD4 count of 830 cells/mm^3 and viral load of 42,000 copies/ml (**Figure 3.1 A**). The mother ('D1') of GD, however, tested HIV-uninfected. It emerged that the HIV-infected grandmother ('GM') of GD had a second daughter ('D2') who was born 5 weeks after GD (**Figure 3.1 B**), and that GM had breast-fed both her daughter (D2) and her grand-daughter (GD). Phylogenetic analysis (**Figure 3.1 C**) of HIV Gag sequences obtained from GM, D2 and GD were consistent with mother-to-child transmission of HIV from GM to D2, and with grandmother-to-grandchild transmission of HIV from GM to GD.

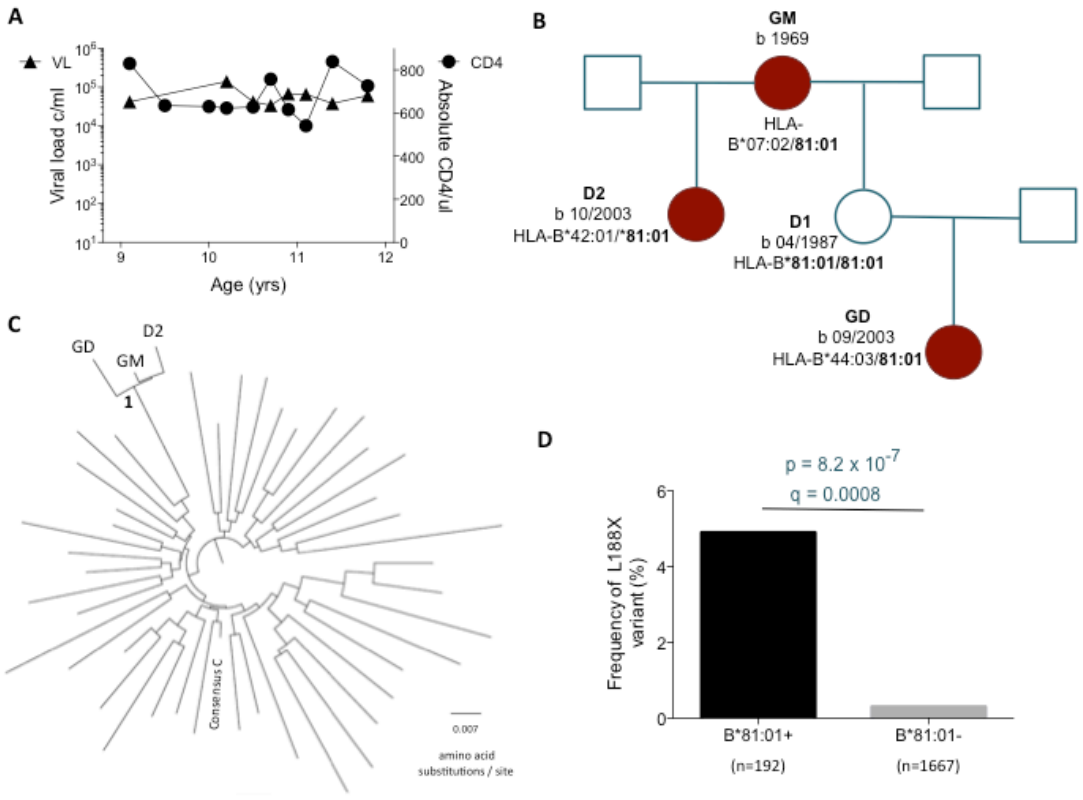


Figure 3.1. Grandmother-to-grandchild transmission in a non-progressing child. (A) Viral load and absolute CD4 count in non-progressing, ART-naive child GD. **(B)** Family tree including HIV-infected family members GM (grand-mother), GD (grand-daughter), D2 (second daughter of GM) and uninfected mother of GD, D1. **(C)** Phylogenetic tree including GD, GM,

D2 and 43 HIV-infected adults from the same locality in Durban, South Africa. **(D)** Frequency of the Gag variant L188F present in all three family members in chronically infected adults with C-clade infection (data Ref. from Carlson *et al.* 2014). Fisher's exact test was performed.

*3.3.2 Rare mutation Gag-L188F within HLA-B*81:01-TL9 epitope shared by transmission trio*

Both GD and D2 were paediatric slow progressors, as defined by the criteria of reaching 5 years of age, ART-naive, with an absolute CD4 count of >350 cells/mm^3 (**Table 3.3**). In case of D2, ART was initiated at age 6.7 years despite a relatively high CD4 count of 625 cells/mm^3, but due to tuberculosis (TB) infection. Factors that might contribute to slow disease progression in related family members would include both viral and shared host genetic factors. We first sought to test the hypothesis that slow disease progression in GD and D2 might be related to common infection by transmission from GM of a virus with low replicative capacity, since viral replicative capacity has been shown to play an important role in disease progression in both paediatric and adult HIV infection (Prince *et al*., 2012; Adland *et al*., 2015). We focused on the Gag region, since previous studies have demonstrated a strong correlation between Gag mutants, viral replicative capacity and disease outcome in both adults and paediatric infection (Prince *et al*., 2012; Adland *et al*., 2015; Claiborne *et al*., 2015). The virus in all family members GM, D2 and GD expressed the mutation Gag-L188F (**Table 3.4**). This mutant arises at the carboxy-terminal anchor residue of the HLA-B*81:01-TL9 epitope (TPQDLNTML, Gag 180-188). This variant is rare, being observed in <0.2 % of published sequences (LANL HIV Sequence Database, http://www.hiv.lanl.gov/) but is strongly associated with expression of HLA-B*81:01 from previous large cohort studies of C-clade infection (Carlson *et al*., 2014) (**Figure 3.1 D**, $p=6.3 \times 10^{-7}$, $q=0.0008$). These data are consistent either with L188F being an escape mutant driven by the HLA-B*81:01-TL9 response in GM and transmitted to both GD and D2; or being selected independently in each of the three family members, all of whom express HLA-B*81:01 (**Table 3.3**).

Table 3.3 HLA types and clinical data available for Grand-mother, Grand-daughter, Daughter-1 and Daughter-2.

Subject	ID	Sex	dob	HIV status	ART initiation	Age* (yrs)	Pre-ART CD4	Pre-ART CD4%	Pre-ART viral load	HLA-B type
Grand-mother	GM	F	12/68	Pos	05/04	35.4	329†	32%†	<50†	07:02/ 81:01
Daughter-1	D1	F	04/87	Neg	n/a	n/a	nd	nd	nd	81:01/ 81:01
Daughter-2	D2	F	10/03	Pos	06/10	6.7	625	27%	n/a	42:01/ 81:01
Grand-daughter	GD	F	09/03	Pos	ART-naïve	9.7	830	33%	42,000	44:03/ 81:01

* Age at ART initiation, or in grand-daughter (GD) age at enrollment.

† pre-ART CD4 counts and viral load not available for Grand-mother GM; data shown are CD4 counts and viral loads after 23m on ART in 04/2006 when GM was aged 37.3 yrs.

nd stands for not-determined; n/a stands for not applicable.

Table 3.4. Differences in autologous Gag sequences between Grandmother, Grand-daughter and Daughter-2, TL9 variants tested and their frequency in HIV-infected adults.

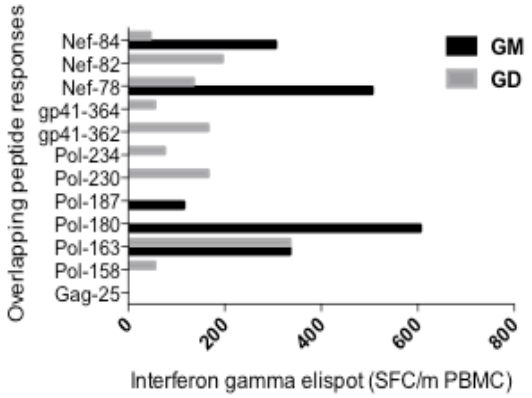
Subject	Gag sequence
Consensus (Gag residue)	I…VHQAISPRTL…KAFSPEVIPMF…**TPQDLNTML**NT…HPVHAGPIAPGQMR……A……T……T…… 134 146-7 167 188 218-9 228 248 332 347
Grand-mother	V…---PL-----…------I----…**--------F**--…--AQ--------I-……T……-……I……
Daughter-1	N/A
Daughter-2	V…---PL-----…------I----…**--------F**--…--VQ--------I-……T……-……I……
Grand-daughter	V…---PL-----…------I----…**--------F**--…--AQ--------I-……A……-……I……

		B*81:01-TL9 **TPQDLNTML**NT	Frequency in B*81+ves†	Frequency in B*81-ves
Variants tested	**WT**	**----------**--	34.2%	85.6%
	T186S	**------S----**	17.1%	0.6%
	Q182S	**--S--------**--	0.0%	1.9%
	T190A	**----------**-A	0.0%	0.3%
	T186S/Q182S	**--S---S----**	12.8%	0.6%
	T186S/T190A	**------S---**-A	0.9%	0.5%
	T186S/T190I	**------S---**-I	1.8%	0.8%
	L188F	**--------F**--	1.8% †	0.1%
	L188F/T190I	**--------F**-I	0.0%	0.0%
Variants not tested			31.4%	10.6%

† the frequency is of the variant shown only; for example other variants in combination with L188F such as Q182S/L188F are found in the study cohort but these variants were not tested.

3.3.3 Impact of L188F on CTL recognition of TL9 peptide

After identifying the rare mutation L188F in all three family members, the PBMCs of GM and GD were tested for *ex vivo* ELISPOT to examine epitope-specific responses (Kiepiela *et al.*, 2006; Altfeld *et al.*, 2000). Both GM and GD showed a range of HIV-specific CD8+ T-cell responses but in neither case did these include the overlapping peptide Gag-25 that contains TL9 epitope. Since >75% of HIV-infected subjects expressing HLA-B*81:01 make a response to this TL9 peptide (Leslie *et al.*, 2006), and in most cases the selection of TL9 variants reduces or abrogates recognition of this epitope (Leslie *et al.*, 2006), these data suggest that L188F is an efficient escape mutation that benefits the virus to avoid immune surveillance from specific CD8+ T cell recognition. This further indicated that it is unlikely that the wild-type virus was transmitted from the grandmother to the grandchild.



GM	**A*43:01**/nd B*07:02/**81:01 C*04:01**/07:02
D-1	**A*43:01**/nd **B*81:01**/81:01 **C*04:01**/nd
D-2	A*30:01/**43:01** B*42:01/**81:01 C*04:01**/17:01
GD	A*29:02/**43:01** B*44:03/**81:01 C*04:01**/08:04

18mer	18mer sequence	Optimal epitope	HLA restriction	Responses in GM	Responses in GD
Nef-84	NYTPGPGVRYPLTFGWCF	TPGPGVRYPL	B*81:01	Y	Y
Nef-82	LWVYHTQGYFPDWQNY	YFPDWQNY	A*29:02	N	Y
Nef-78	FKGAFDLSFFLKEKGGL	Not known	Not known	Y	N
gp41-362	IVQQQSNLLRAIEAQQHM	RAIEAQQHM	C*08	N	Y
gp41-364	HMQLTVWGIKQLQTRVL	Not known	Not known	N	N
Pol-234	AQPDKSESELVNQIIEQL	Not known	Not known	N	Y
Pol-230	LALQDSGSEVNIVTDSQY	SEVNIVTDSQY	B*44:03	N	Y
Pol-187	QGWKGSPAIFQSSMTKIL	SPAIFQSSM	B*81:01	Y	N
Pol-180	KKKSVTVLDVGDAYFSV	Not known	C*04:01	Y	N
Pol-163	LVGPTPVNIIGRNMLTQL	TPVNIIGRNML	B*81:01	Y	Y
Pol-158	NLPGKWKPKMIGGIGGFI	Not known	B*44:03	N	Y
Gag-25	GATPQDLNTMLNTVGGH	TPQDLNTML	B*81:01	N	N

Figure 3.2. IFN-γ ELISPOT responses of PBMC in Grand-mother and Grand-daughter to a panel of 410 overlapping peptides spanning the C-clade proteome. Responses to the indicated overlapping peptides are shown. The immunodominant HLA-B*81:01 epitope TL9 is contained within overlapping peptide-25, there was no response to this in either GM or GD.

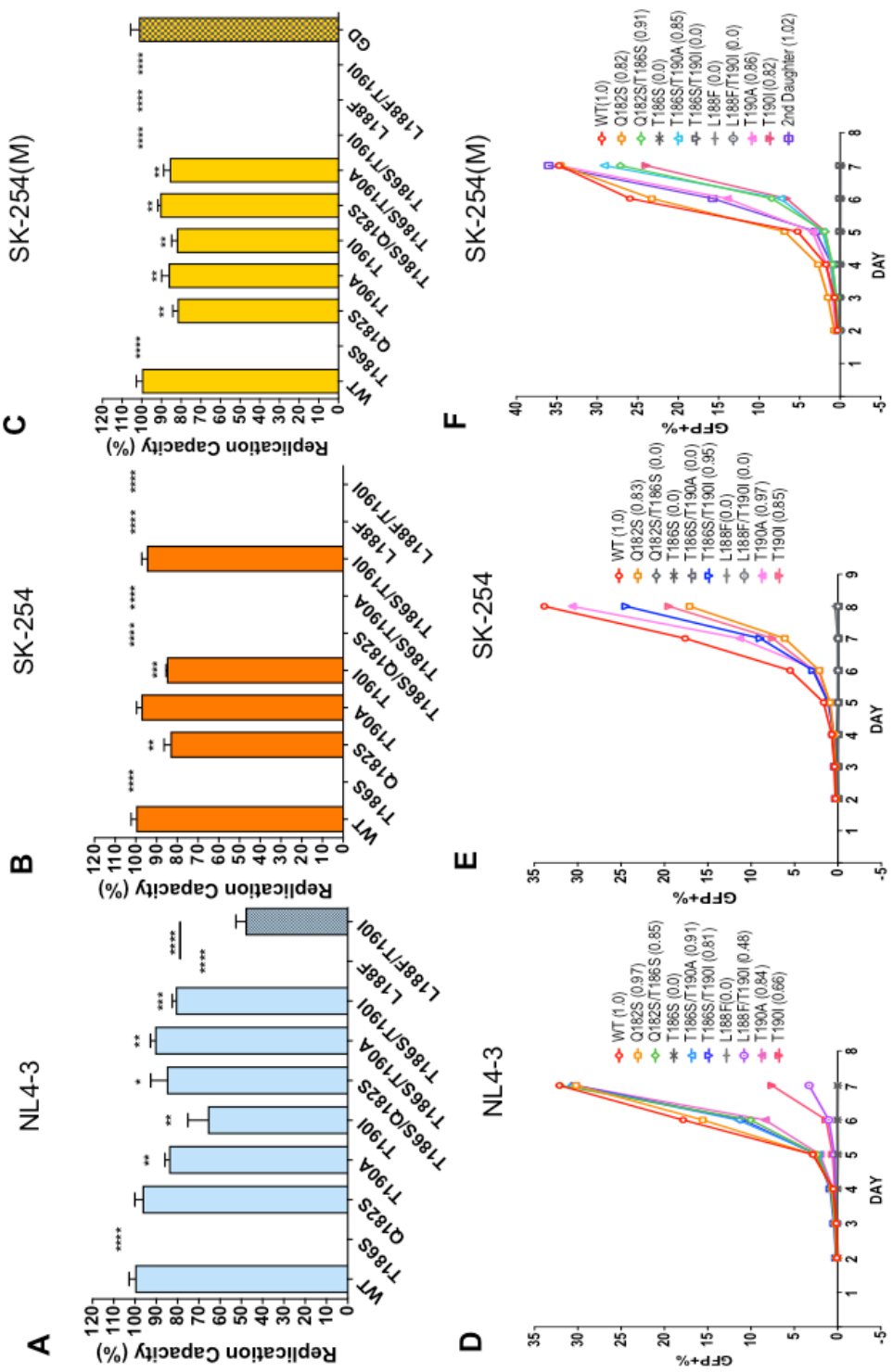
Table 3.5. p24 Gag sequences for NL4-3, SK-254, SK-254(M) and Consensus C-clade.

Plasmid	HXB2 Gag position
	147 159
NL4-3	P I V Q N L Q G Q M V H Q A I S P R T L N A W V K V V E E K A F S P E V I P M F
SK-254	- - - - - - - - - - - - - - T - - - - - - - - - - - I - - - - - - - - - - - - -
SK-254(M)	- I - - - - - - - - - - - - -
Consensus C	- I - - - - - - - - - - - - -
	173 203
NL4-3	S A L S E G A T P Q D L N T M L N T V G G H Q A A M Q M L K E T I N E E A A E W
SK-254	T - D - - - - - - - - -
SK-254(M)	T - D - - - - - - - - -
Consensus C	T - D - - - - - - - - -
	248 252
NL4-3	D R L H P V H A G P I A P G Q M R E P R G S D I A G T T S T L Q E Q I G W M T H
SK-254	- A - - - G
SK-254(M)	- A - - - S
Consensus C	- A - - - S
	256 260 280 286
NL4-3	N P P I P V G E I Y K R W I I L G L N K I V R M Y S P T S I L D I R Q G P K E P
SK-254	- - - V - - - D - - - - - - - - - - - - - - - - - - - V - - - - - K - - - - - -
SK-254(M)	- - - - - - - D - - - - - - - - - - - - - - - - - - - V - - - - - K - - - - - -
Consensus C	- - - - - - - D - - - - - - - - - - - - - - - - - - - V - - - - - K - - - - - -
	301 310 312 319
NL4-3	F R D Y V D R F Y K T L R A E Q A S Q E V K N W M T E T L L V Q N A N P D C K T
SK-254	- - - - - - - - F - - - - - - - - T - D -
SK-254(M)	- - - - - - - - F - - - - - - - - T - D - - - - - - D - - - - - - - - - - - - -
Consensus C	- - - - - - - - F - - - - - - - - T - D - - - - - - D - - - - - - - - - - - - -
	335 340 342 357
NL4-3	I L K A L G P G A T L E E M M T A C Q G V G G P G H K A R V L
SK-254	- - R - - - - A - S - - - - - - - - - - - - - - S - - - - - -
SK-254(M)	- - R - S - - - - - -
Consensus C	- - R - S - - - - - -

3.3.4 Impact of L188F on viral replicative capacity

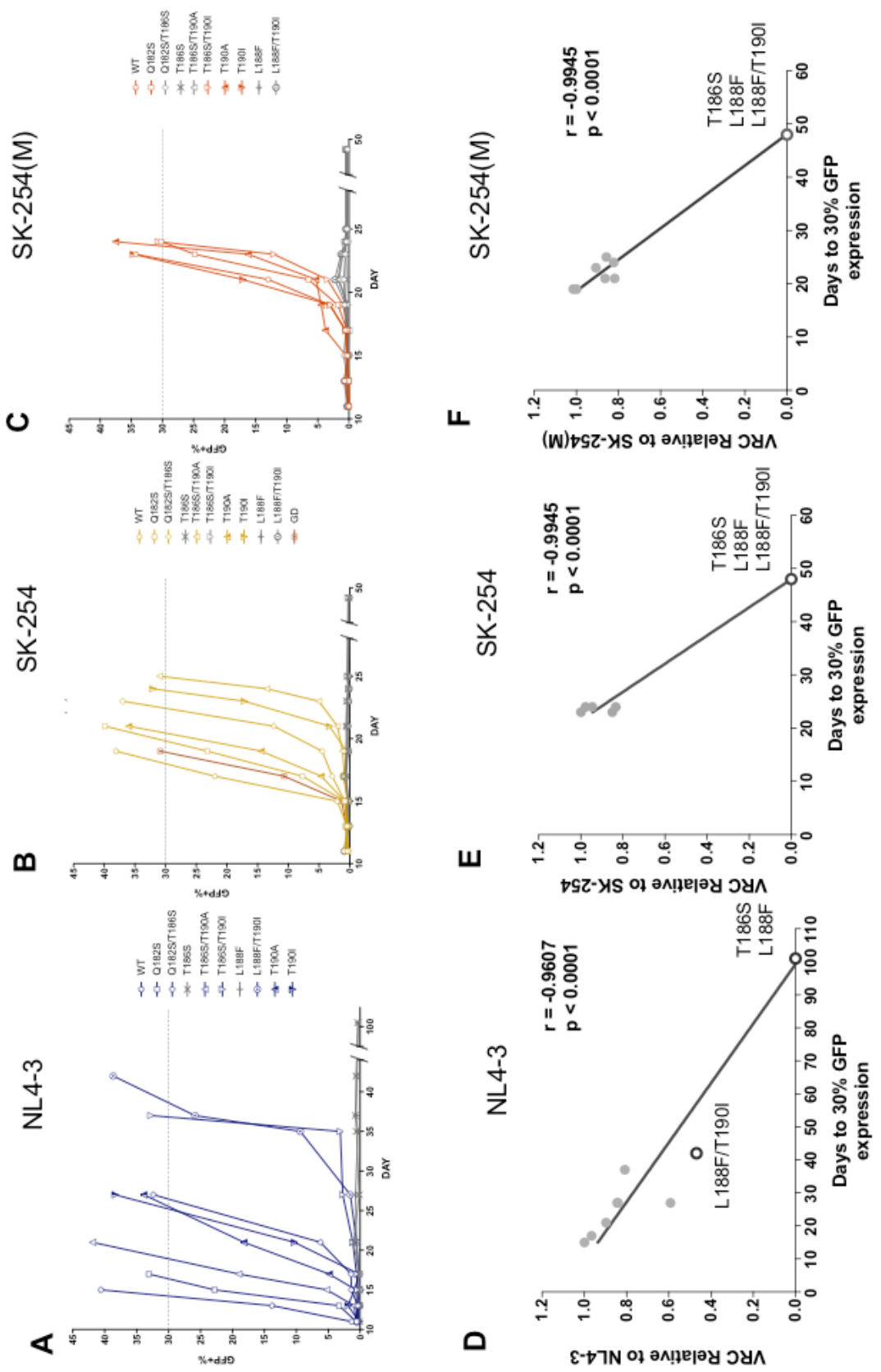
Previous studies (Wright *et al*., 2010; Wright *et al*., 2012) have demonstrated that the most common escape variants within the HLA-B*81:01-TL9 epitope, T186S and Q182S/T186S (**Table 3.4**), can have a substantial impact on viral replicative capacity (VRC). To investigate the possibility that this rare L188F mutant might similarly affect VRC, thereby contributing to slow progression observed in subjects GD and D2, we generated this variant via site-directed mutagenesis and the replicative capacity was compared with the more common and well-described HLA-B*81:01-associated TL9 escape mutants; T186S, the combination of T186S/Q182S and the putative compensatory mutants T190A and T190I (Wright *et al*., 2012; **Table 3.4**). These comparisons were undertaken using three different viral backbones: the laboratory-adapted B-clade virus NL4-3; a chimeric virus comprising NL4-3 with Gag-protease inserted from a C-clade infected South African subject SK-254, whose viral sequence is similar to the C-clade consensus (Wright *et al*., 2012); and a chimeric virus, referred to here as SK-254(M), comprising NL4-3 with the same Gag-protease inserted from the C-clade infected South African subject SK-254, but with the necessary retro-mutations made to enable the viral amino acid sequence to be identical with the C-clade consensus (**Table 3.5**).

Figure 3.3. (Next page) Impact of variants within the HLA-B*81:01-TL9 epitope on viral replicative capacity (VRC). (A-C) Impact of viral variants studied on replication capacity compared to NL4-3 in three distinct viral backbones. **(A)** NL4-3 viruses. **(B)** SK-254 viruses. **(C)** SK-254(M) viruses. (D-F) Percentage GXR cells expressing GFP following infection with an MOI of 0.03 using virus variants shown, from Day 2 to Day 7 or Day 8. **(D)** NL4-3 viruses. **(E)** SK-254 viruses. **(F)** SK-254(M) viruses. All bars are shown as mean with SD. Patterned bars represent L188F/T190I and GD, respectively. Colour blue stands for B-clade viruses (NL4-3), orange represents C-clade SK-254 and yellow shows C-clade SK-254(M). One-way ANOVA multiple comparison and Dunnett's correction were used for calculating *p* value. Asterisks marked on top of each bar are the *p* value when each mutant virus was compared to the wild-type. (* $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$)


A
NL4-3
B
SK-254
C
SK-254(M)
Replication Capacity (%)
WT
T186S
Q182S
T190A
T190I
T186S/Q182S
T186S/T190A
T186S/T190I
L188F
L188F/T190I
GD
D
NL4-3
E
SK-254
F
SK-254(M)
GFP+%
DAY
WT (1.0)
Q182S (0.83)
Q182S/T186S (0.0)
T186S (0.0)
T186S/T190A (0.0)
T186S/T190I (0.95)
L188F(0.0)
L188F/T190I (0.0)
T190A (0.97)
T190I (0.85)
WT(1.0)
Q182S (0.82)
Q182S/T186S (0.91)
T186S (0.0)
T186S/T190A (0.85)
T186S/T190I (0.0)
L188F (0.0)
L188F/T190I (0.0)
T190A (0.86)
T190I (0.82)
2nd Daughter (1.02)
WT (1.0)
Q182S (0.97)
Q182S/T186S (0.85)
T186S (0.0)
T186S/T190A (0.91)
T186S/T190I (0.81)
L188F(0.0)
L188F/T190I (0.48)
T190A (0.84)
T190I (0.66)

The VRC determined for each mutant virus is shown relative to NL4-3, SK-254 and SK-254(M) respectively (**Figure 3.3**). Irrespective of the backbone, any viruses encoding L188F or T186S were incapable of replicating at a sufficient level *in vitro* to measure VRC. Although GFP+ expression in target cells remained negative up to >100 days in culture, viral RNA was successfully extracted from the supernatant and the respective L188F and T186S mutations were confirmed by sequencing. This would indicate that the viruses were generated and replicating, although insufficiently to assay VRC. The second observation here is that the viral backbone has a considerable influence on the impact that individual variants may make on VRC. For example, T190I shows a modest compensatory effect in combination with L188F in NL4-3, but not in SK-254 or SK-254(M). In relation to T186S, T190A and T190I are compensatory in the setting of NL4-3, but only T190I in the context of SK254 and only T190A in the context of SK254(M). Of note, however, the entire Gag-protease sequence taken from GD to form a chimeric virus with NL4-3 replicated as well as SK-254(M) WT, despite containing the L188F variant. This indicates that one or more of the additional variants within the Gag-protease sequence of the GD were capable of compensating for L188F. We were not able to generate chimeric viruses from the GM or D2 samples available but these differed in sequence by only two and one residue from that of GD, respectively (**Table 3.4**).

Figure 3.4. (Next page) Impact of variants within the HLA-B*81:01-TL9 epitope on time till viral expansion. (A-C) Percentage GXR cells expressing GFP before reaching 30% harvesting point in three distinct viral backbones. **(A)** NL4-3 viruses. **(B)** SK-254 viruses. **(C)** SK-254(M) viruses. (D-F) Correlation between growing days and viral replicative capacities (VRC). **(D)** NL4-3 viruses. **(E)** SK-254 viruses. **(F)** SK-254(M) viruses. Colour blue stands for B-clade viruses (NL4-3), orange represents C-clade SK-254 and yellow shows C-clade SK-254(M). Non-growing viruses are shown in grey. Pearson's correlation *t*-test was used for calculating *r* and *p* value.


A
NL4-3
B
SK-254
C
SK-254(M)
GFP+%
DAY
WT
Q182S
Q182S/T186S
T186S
T186S/T190A
T186S/T190I
L188F
L188F/T190I
T190A
T190I
GID
D
NL4-3
E
SK-254
F
SK-254(M)
r = -0.9607
p < 0.0001
r = -0.9945
p < 0.0001
T186S
L188F
L188F/T190I
VRC Relative to NL4-3
VRC Relative to SK-254
VRC Relative to SK-254(M)
Days to 30% GFP expression

3.3.5 Time till viral expansion correlates with VRC

Our fitness assay results were consistent with the previous study of the impact of TL9 mutants on VRC (Wright *et al*., 2012). In particular, viruses carrying certain mutations such as T186S had a sufficiently low VRC that these detectable in the VRC assay. In the current studies, in our attempt to grow these mutant viruses, the transfection was set up at least in three independent benches and the viral growth was monitored for up to 100 days. The GFP+% of these mutant viruses remained negative, meaning the VRC was too low to meet the limit of detection of the assay (**Figure 3.4 A** to **C**). As might be anticipated, there was a strong correlation between the time to detection of the viruses by GFP expression and the VRC (**Figure 3.4 D** to **F**).

3.3.6 Structural analysis indicates likely mechanism of L188F and T186S impact on VRC

The structural modeling and analysis in the following section was undertaken by Dr David K Cole, Sewell Group, Cardiff University School of Medicine.

In order to better understand why the HLA-B*81:01-driven escape mutants L188F and T186S have such a marked impact on VRC, we analysed these mutants using previously solved crystal structures of the capsid protein (Momany *et al*., 1996; Gamble *et al*., 1996; Berthet-Colominas, 1999; Tang *et al*., 2002). The capsid protein comprises an amino-terminal domain (NTD: Gag 133-283) linked to a carboxy-terminal domain (CTD: Gag 284-363); the main features of the NTD are an N-terminal β-hairpin loop, seven α-helices and the cyclophilin binding loop linking helices-4 and-5 (**Figure 3.5 A**). The TL9 epitope (Gag 180-188) is located almost entirely within helix-3 (Gag residues 181-189) (**Figure 3.5**).

To consider the impact of the L188F escape mutant, the contacts between 188-Leu and neighbouring residues 184L, 198M, 201L and 266I were identified, showing the

formation of seven van der Waals interactions, in addition to the hydrogen bond between the backbone amine group at residue 188 and the backbone carbonyl group at residue 184 (**Figure 3.5 B**). Modelling of the L188F substitution would indicate that all but one of these seven van der Waals interactions are disrupted (**Figure 3.5 C**), thereby likely substantially destabilizing the interface between helices-3 and -4. Similarly, modelling suggests that the consequence of the T186S escape mutant would be to abrogate three of four van der Waals interactions between 186-Threonine on helix-3 and 152-Leucine and 147-Isoleucine on helix-1 (**Figure 3.5 D**). The destabilizing impact of this would, however, be largely mitigated by the compensatory T190I mutation, in providing three new van der Waals interactions between 190-Isoleucine and 152-Leucine (**Figure 3.5 E** and **F**). However, the variable ability of variants at 190-Threonine to compensate successfully for the fitness cost of the T186S escape mutant according to viral backbone (NL4-3, SK-254 or SK-254(M)) is not explained by these models, with the majority of the differences between these viruses located within the carboxy-terminal domain (**Table 3.5**).

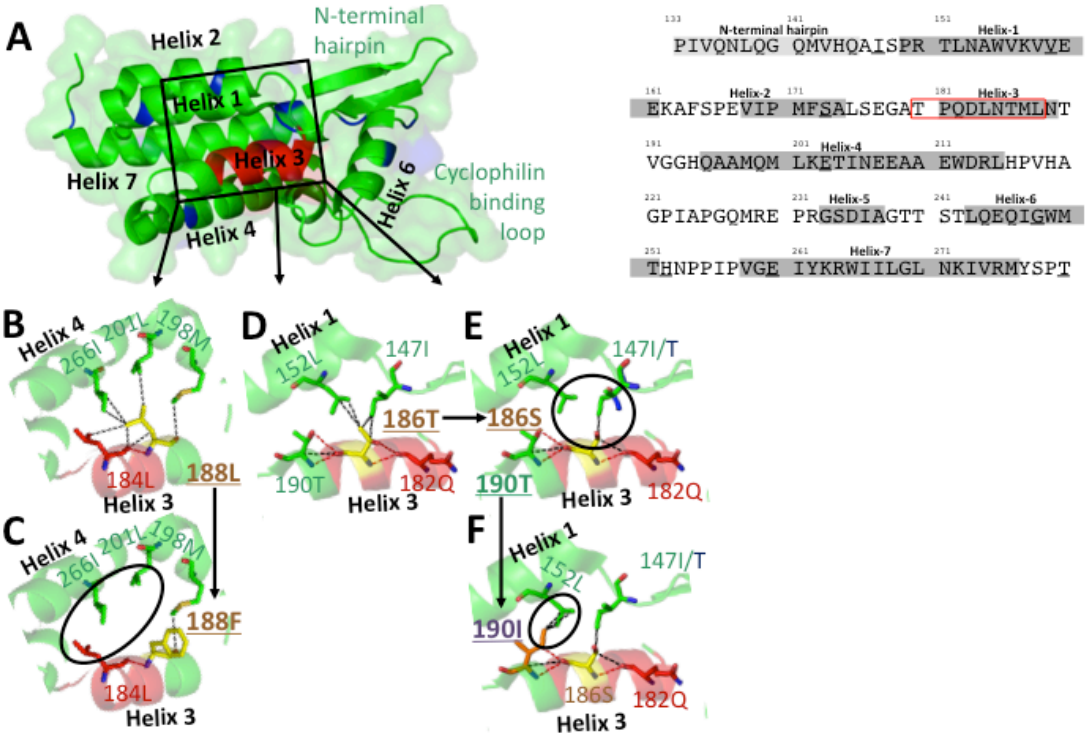


Figure 3.5. Structural modelling of the impact of escape mutations around the p24-Gag B81:01-TL9 epitope. (A) Overall structure of p24-Gag protein residues 133–283 demonstrating the position of the B*81:01-TL9 epitope (*red cartoon*) and the polymorphisms between the NL4-3 and SK-254 viruses (*blue cartoon*). The environment that is likely to be affected by mutations around the B*81:01-TL9 epitope, in helix 3, is highlighted in the *black box*. **(B)** Binding network (*green* and *red sticks*) around residue 188L (*yellow sticks*) in the NL4-3 virus, likely to be important for maintaining the protein fold between helix 3 and 4. **(C)** Modelling of the rare L188F mutation, shown to reduce viral replication capacity. This mutation could abrogate interactions between residue 188 with 184L, 198 M, 201L and 266I, possibly destabilizing the helix 3–helix 4 interface. **(D)** Binding network (*green* and *red sticks*) around residue 186T (*yellow sticks*) in the NL4-3 virus, likely to be important for maintaining the protein fold between helix 1 and 3. **(E)** Modelling of the T186S mutation (*black arrow*), shown to be detrimental to viral health. This mutation abrogates interactions between residue 186 with 152L and 147I/T (*black circle*), possibly destabilizing the helix 1–helix 3 interface. **(F)** Structural modeling of the T190I mutation (*black arrow*) that can rescue the T186S mutation, restoring viral replication capacity. Modelling suggests that new interactions can form between residue 190I and 152L (*black circle*), potentially restoring helix 1–helix 3 interface stability.

3.4 Discussion

These studies describe two HIV-infected, genetically related children, born within 5 weeks of one another, both of whom progressed slowly to disease. One child was infected via mother-to-child transmission and the other, more unusually, was infected by the same donor, but via grandmother-to-grandchild transmission. This scenario raised the question of whether slow progression in the two genetically related children could be related to the virus transmitted, or to shared host factors. The donor, ‘GM’, the grand-daughter, ‘GD’ and the HIV-infected daughter, ‘D2’, all expressed HLA-B*81:01, which is normally associated with protection against disease progression in adult infection (Kiepiela *et al*., 2004; Leslie *et al*., 2010). Autologous virus in each of the three family members also shared the same, rare, L188F mutation, which is strongly associated with HLA-B*81:01, and lies at the carboxy-terminus of the immunodominant HLA-B*81:01-restricted Gag epitope TL9 (TPQDLNTML, Gag 180-188). In common with the more frequent HLA-B*81:01-TL9

escape mutation T186S, the L188F variant almost completely prevented viral replication when introduced into heterologous HIV strains. However, the replicative capacity of the virus encoding the entire grand-daughter (GD) Gag sequence did not differ from wild-type virus, suggesting VRC was nor responsible for the children's slow progressing phenotype. It is entirely possible (as discussed further below) that either the uncompensated L188F variant was transmitted to both daughters, or that wild-type virus was transmitted and L188F plus the near-identical compensatory mutants were selected independently post-transmission in GM, GD and D2. However, on the basis that low fitness viruses are less likely to be transmitted (Carlson *et al*., 2014) and no response toward wild-type OLP containing TL9 epitope in ELISPOT assay was observed (**Figure 3.2**), the most parsimonious interpretation of these data may be that the fully compensated L188F virus was transmitted in both cases and that slow progression in the two children was not related to viral factors but to undefined host-specific factors.

Previous studies of the mutations selected within the HLA-B*81-TL9 epitope have similarly shown that the most frequent escape variant, T186S, can virtually abrogate viral replication (Wright *et al*., 2012). Similar to the analyses reported here, the impact of T186S on viral replicative capacity varied considerably according to the nature of the viral backbone, as did the ability of variants such as T190I to rescue the virus from the fitness cost resulting from T186S (Wright *et al*., 2012). The studies described here are highly consistent with these findings, in demonstrating, first, the dramatic negative impact of L188F, as well as T186S, on viral replicative capacity; and, second, the striking influence of viral sequence context. For example, T190I and not T190A was able to compensate for T186S in the SK-254 virus but the reverse was the case in the SK-254(M) virus. The p24 Gag residues that might contribute to the substantial influence of sequence context on the impact of the TL9 escape mutants on viral replicative capacity observed here include several that have been

previously implicated as compensatory mutants (Brockman *et al*., 2007; Leslie *et al*., 2004; Crawford *et al*., 2012; Carlson *et al*., 2012a), such as those within the cyclophilin A binding loop (V218A, H219Q and M228T) and the Gag tropism-determining loops (Gag residues 137-147; and Gag 248-254) (Hatziioannou *et al*., 2004).

The lack of samples at the time of transmission means that the sequences of the viruses that were transmitted are unknown. It is possible therefore that the same L188F variant within the B*81:01-TL9 epitope was selected independently from the wild-type TL9 sequence in all three family members subsequent to transmission, since all three individuals expressed HLA-B*81:01. If this were the case, it is likely that the virus transmitted by GD without L188F would have had normal viral replicative capacity, at least on the basis of the otherwise unremarkable Gag sequence in these three family members. In the more likely scenario of the virus encoding L188F being transmitted both to GD and D2, the fact that this virus was fully compensated in GD suggests that low viral replicative capacity is unlikely to have contributed to slow progression in the two children.

These observations would suggest therefore that undefined host factors are more likely to have contributed to slow progression in the two genetically related children. Although in this case both children expressed HLA-B*81:01 that is associated with slow progression in adult infection, large cohort studies indicate that ‘protective’ HLA alleles, such as HLA-B*57, HLA-B*58:01 and HLA-B*81:01, do not significantly influence paediatric disease progression (Adland *et al*., 2015). Furthermore, in a small series of HLA-B*27-positive HIV-infected children, any benefit to the child of expressing HLA-B*27 was negated in instances where the HLA-B*27-mother transmitted an escape mutant in the critical HLA-B*27 Gag epitope (Goulder *et al*., 2001). In the current study, no HLA-B*81:01-TL9 response was detectable in the

grandchild GD (**Figure 3.2**), suggesting that this response is not critical to non-progression. Furthermore, in keeping with other non-progressor children (Adland *et al.*, 2014; Goulder *et al.*, 2016), normal-for-age CD4 counts were maintained in GD through childhood despite relatively high viral loads (between 38,000 and 140,000 copies/ml between 9 and 12 years of age). HLA-mediated control of viraemia that is characteristic of adult non-progression (Goulder *et al.*, 2012) is not a typical feature of non-progressive paediatric infection. In common with the natural hosts of SIV in whom normal CD4 counts and low systemic immune activation are also maintained despite persistent high level viraemia, the mechanisms underlying non-progression in paediatric infection differ substantially from those operating among adult elite controllers (Pereyra *et al.*, 2009) and remain incompletely defined.

In unpublished studies of HLA-B*81:01-positive paediatric slow progressors who have undetectable viral load at some time points – an extremely unusual phenomenon in ART-naïve HIV-infected children – viral sequences showed other rare mutations in the TL9 epitope (Adland *et al.*, unpublished data). Mixed base variants (e.g. T/P) appeared at position 3 (Gag 182) of this epitope. All the autologous mutants Q182T, Q182G and Q182P had a diminished VRC. In particular, Q182P aborted VRC completely as T186S and L188F in both B- and C-clade viruses. In these paediatric cases studied, Q182P had not been preserved in chronic infection. This suggests that certain changes in the highly conserved TL9 epitope are very unstable, as a result of the negative impact on viral replicative capacity. In contrast to our family trio, whose viral loads were high, these slow progressors had relatively low viral loads, which more closely resembles the situation observed in adult viraemic controllers. However, whilst HLA-specific CD8+ T-cell responses have a significant impact on viral load in both paediatric and adult infection (Muenchhoff *et al.*, 2016; Kiepiela *et al.*, 2007), disease progression itself is less affected by viral load in paediatric infection, and hence the impact of HLA molecules such as HLA-

B*81:01 is less among infected children than adults.

One additional notable feature in this study is the occurrence of paediatric HIV infection in the absence of mother-to-child transmission. The particular scenario of grandmother-to-grandchild transmission has not, to our knowledge, been described previously, but transmission via surrogate breast-feeding is a well-recognized cause of non-vertical, non-sexual HIV infection in children (Cotton *et al.*, 2012; Goepfert *et al.*, 2012). The prevalence of paediatric transmission via surrogate breast-feeding is unknown but, from data accumulated within large studies of >250 infected children, this would likely represent approximately 1-2 % of paediatric infections (van Kooten Niekerk *et al*., 2006; Shisana *et al*., 2008; Vaz *et al*., 2010). The relevance of grandmother-to-grandchild transmission to non-progression is open to speculation, but certainly timing of transmission affects outcome in paediatric infection, *in utero* infected children progressing faster than those infected *intra-partum* and children infected *post-partum* via breast milk progressing the slowest (Marinda *et al*., 2007).

Overall, these data are consistent with previous studies indicating that, compared to adult elite controllers, distinct mechanisms underlie slow HIV paediatric disease progression, compared to adult elite controllers. The principal host factors responsible for disease non-progression in children appear not to include HLA class I and remain yet to be defined. Further evaluation of this group of paediatric slow progressors is therefore warranted.

CHAPTER 4: IMPACT OF HLA-B*52:01-DRIVEN ESCAPE MUTATIONS ON VIRAL REPLICATIVE CAPACITY

4.1 Introduction

In HIV infection, CTL responses play a central role in suppressing viral replication, and HLA class I allele variation has consistently been the strongest genetic associated factor in disease outcome, in particular HLA-B alleles (Kiepiela *et al*., 2004; Pereyra *et al*., 2010). Previous studies have shown HLA-B*52:01 is protective both in the early time points and against disease progression (Fabio *et al*., 1992; Antoni *et al*., 2013; Vijaya Lakshmi *et al*., 2006; Teixeira *et al*., 2014). Yet our understanding of the underlying mechanism of HLA-B*52:01-mediated immune control of HIV is still very limited due to the low prevalence of the allele in the populations most commonly studied. The analyses undertaken to date are based on B-clade cohorts in Japan where the HLA-B*52:01 is relatively prevalent (Yagita *et al*., 2012; Chikata *et al*., 2014; Sakai *et al*., 2015; Murakoshi *et al*., 2015). In the current study, we focus on a north Indian population in which HLA-B*52:01 is expressed in 20% of subjects (The Allele Frequency Net Database, www.allelefrequencies.net), and where C-clade is the predominant subtype, accounting for 96% infection So far there has been no study examining the HLA-associated impact on HIV in India, which has the world's third highest burden with 2.7 million infections.

HLA-B*52:01, as with other members of B62 supertype family, accommodate aliphatic amino acids (L, I, V, M, Q) in the B pocket, whereas in the F pocket, aromatic, aliphatic and hydrophobic molecules are preferred (F, Q, Y, L, I, V, M, W, A) (Sidney *et al.,* 2008). More specifically, the B*52:01-restricted amino acids in the binding-motif are listed in **Table 4.1**, Glutamine at position 2 and Isoleucine or Valine at the carboxyl-terminus (Marsh *et al.*, 2000; Goulder and Walker, 2012).

Table 4.1. Characteristics of B*52:01-restricted epitope.

	Peptide-Binding Motif				
Position	**P1**	**P2**	**P3**	**P5-7**	**PC**
Binding pocket	A	B	D	C	F
B*52:01	-	Q	F/Y/W	L/I/V	I/V

Table is based on The HLA Factsbook (Marsh *et al*., 2000).

Although there have been various peptides reported as HLA-B*52:01-restricted in different parts of HIV genome, in most cases these have not been characterized in such a way that unequivocally defines optimal epitopes presented by HLA-B*52:01 (Hadida *et al*., 1995; Brander *et al*., 1999; Chikata *et al*., 2014). To date, there has been only one optimally defined HLA-B*52:01-restricted epitope RI8 (275-282, RMYSPVSI) in the conserved capsid protein p24 Gag (Brumme *et al*., 2008; Goulder and Walker, 2012). In a comprehensive analysis identifying associations between HIV amino acid polymorphisms and particular HLA class I allele expression, the T280S mutation is selected in subjects expressing HLA-B*52:01 (n=1,888 treatment-naïve, chronically B-clade infected, q value <0.05; Carlson *et al*., 2012a).

Successful viral control mediated by protective HLA alleles typically is associated with targeting highly conserved regions of the HIV genome (Miura *et al.*, 2009a; Goulder and Walker, 2012). Population specific genetic variation is largely due to the strong selection pressure mediated by existing HLA alleles (Kawashima *et al*., 2009; Kløverpris *et al*., 2016). In this case, the strong selection pressure of HLA-B*52:01-restricted-p24-Gag-RI8 CTL might be expected to drive the selection escape mutations that lower the viral replicative capacity, since as illustrated in Chapter 3, the conserved capsid protein does not generally tolerate amino acid sequence changes without some negative impact on viral fitness (Miura *et al.*, 2009a). Several studies have shown that VRC plays a critical role in disease progression (Goepfert *et al*., 2008; Prince *et al*., 2012). Subsequent work to aforementioned the Carlson study

(Carlson *et al*., 2012a) not only confirmed the association of the T280S mutation in HLA-B*52:01-positive individuals in B-clade infection, but also showed that T280S, along with T280V and T280A, are escape mutations which lower the viral fitness (Murakoshi *et al*., 2015).

Of note, six amino acids downstream of the RI8 epitope, there is an HLA-B*52:01-associated Arginine to Lysine substitution R286K which is observed in linkage disequilibrium with, and dependent on, the T280S escape mutation (Carlson *et al*., 2012a). This occurrence of a mutation that is selected subsequent to an escape mutation, it is highly suggestive of a compensatory mutation that can in part restore VRC towards wild-type (Crawford *et al*, 2007; Wright *et al*, 2012). The K286R mutation is therefore potentially a compensatory mutation for T280S.

The focus of this study was to investigate the impact of HLA-B*52:01-mediated control of HIV in both B- and C- clade infection by examining the escape and potential compensatory mutants within the p24 Gag-RI8 epitope and to determine whether these mutations affect VRC.

4.2 Chapter-specific Materials and Methods

4.2.1 Study Subjects

In total, 169 chronically HIV infected individuals were recruited in a North India Cohort. Characteristics of the studied participants are provided in **Table 4.2**. CD4+ T cell counts data was obtained from the AIDS Division of the National Centre for Disease Control (NCDC), Delhi. Viral load data were unavailable in this cohort. Sequences that belonged to clade other than C were excluded from further analysis; including CD4 data analysis and identification of HLA associated polymorphisms.

Table 4.2. Characteristics of studied participants from New Delhi, India (*n*=169).

Characteristic	
Recruitment year	2011-2014
Ethnicity	
Indian	169 (100)
Sex (*n*=169)	
Male	92 (54.4)
Female	76 (45.0)
Transgender	1 (0.6)
CD4+ T cell count, cells/mm^3 (*n*=169)[a]	447.7 [311-584]
HLA-B typing (*n*=169)	
HLA-B*52:01+ve	38 (22.5)
HLA-B*52:01-ve	131 (77.5)
HIV subtype (*n*=158)[b]	
C	140 (88.6)
B	17 (10.8)
A1	1 (0.6)

*Data are no. (%) of patients, otherwise indicated. IQR is shown in brackets. [a]CD4+ T-cell measurements were performed at enrolment. [b]HIV subtype was determined by p24 Gag sequences.

4.2.2 p24 Gag PCR amplification from proviral DNA

The p24 Gag amplification and sequencing were performed by Dr Supriya Singh, Department of Biotechnology, National Centre for Disease Control Delhi, India.

Nested PCR was carried out for the p24 region of Gag by using different sets of primers: G00 (5'-GACTAGCGGAGGCTAGAAG-3') and G01 (5'-AGGGGTCGTTGCCAAAGA-3') for the first round; G60 (5'-CAGCCAAAATTACCCTATAGTGCAG- 3') and G25 (5'-ATTGCTTCAGCCAAAACTCTTGC -3') for the second round. The PCR products were purified and subjected to automated nucleotide sequencing using the dideoxy terminator sequencing chemistry on an automated DNA sequencer (ABI Genetic Analyzer 3730xl).

4.2.3 Site-directed mutagenesis of NL4-3, SK-254(M) and Indian Consensus (IC)

All the mutations were introduced by using the QuikChange Lightning site-directed mutagenesis kit (Agilent technologies). The fitness assay performed in this study was based on ΔGag-protease pNL4-3 with three different insertions of Gag-protease sequences as the backbones: (i) pNL4-3, a HIV-1 subtype B NL4-3; (ii) SK-254(M), a modified version of patient-derived HIV-1 Gag-protease sequence (SK-254, GenBank accession number HM593258) with p24 Gag identical to Consensus C as stated in Chapter 3 and (iii) Indian Consensus (IC), determined based on the cohort subjects (*n*=140) by using Geneious ver.7.0.4. Custom-designed mutagenesis forward and reverse primers of I147L, I256V, D260E; 280X; and 286X are listed in **Table 4.3**. Note that due to the proximity of I256V and D260E in IC, the mutagenesis was performed twice, one mutation each time.

Table 4.3. HLA-B*52:01 Custom-designed mutagenesis forward and reverse primers.

HXB2 position in Gag	Clade	Mutant	Primer direction	Primer sequence (5'→3')
147	C	I→L	Forward	CAA ATG GTA CAC CAA GCC **TTA** TCA CCT AGA ACT TTG AA
			Reverse	TT CAA AGT TCT AGG TGA **TAA** GGC TTG GTG TAC CAT TTG C
256	C	I→V	Forward	G ATG ACT AGT AAC CCA CCT **GTC** CCA GTG GGA G
			Reverse	C TCC CAC TGG **GAC** AGG TGG GTT ACT AGT CAT C
260	C	D→E	Forward	CA CCT ATC CCA GTG GGA **GAG** ATC TAT AAA AGA TGG ATA AT
			Reverse	AT TAT CCA TCT TTT ATA GAT **CTC** TCC CAC TGG GAT AGG TG
280	B	T→V	Forward	AAA ATA GTA AGA ATG TAT AGC CCT **GTC** AGC ATT CTG GAC ATA AGA CAA GG
			Reverse	CC TTG TCT TAT GTC CAG AAT GCT **GAC** AGG GCT ATA CAT TCT TAC TAT TTT
		T→S	Forward	A GTA AGA ATG TAT AGC CCT **AGC** AGC ATT CTG GAC ATA AG
			Reverse	CT TAT GTC CAG AAT GCT **GCT** AGG GCT ATA CAT TCT TAC T
		T→A	Forward	ATA GTA AGA ATG TAT AGC CCT **GCC** AGC ATT CTG GAC ATA AG
			Reverse	CT TAT GTC CAG AAT GCT **GGC** AGG GCT ATA CAT TCT TAC TAT
	C	V→T	Forward	T AAA ATA GTA AGA ATG TAT AGC CCT **ACC** AGC ATC TTG GAC ATA AAA CAA GGG
			Reverse	CCC TTG TTT TAT GTC CAA GAT GCT **GGT** AGG GCT ATA CAT TCT TAC TAT TTT A
		V→S	Forward	AAA ATA GTA AGA ATG TAT AGC CCT **AGC** AGC ATC TTG GAC ATA AAA CAA GGG
			Reverse	CCC TTG TTT TAT GTC CAA GAT GCT **GCT** AGG GCT ATA CAT TCT TAC TAT TTT A
		V→A	Forward	A GTA AGA ATG TAT AGC CCT **GCC** AGC ATC TTG GAC ATA AAA C
			Reverse	G TTT TAT GTC CAA GAT GCT **GGC** AGG GCT ATA CAT TCT TAC T
286	B	R→K	Forward	C AGC ATT CTG GAC ATA **AAA** CAA GGA CCA AAG GAA CCC
			Reverse	GGG TTC CTT TGG TCC TTG **TTT** TAT GTC CAG AAT GCT G
	C	K→R	Forward	C AGC ATC TTG GAC ATA **AGA** CAA GGG CCA AAG GAA CC
			Reverse	GG TTC CTT TGG CCC TTG **TCT** TAT GTC CAA GAT GCT G

*Mutant amino acids are shown in bold in the primer sequences.

4.3 Results

4.3.1 Identification of Indian Consensus (IC)

The phylogenetic analysis in the following section was kindly performed by Dr Supriya Singh, Department of Biotechnology, National Centre for Disease Control Delhi, India.

Indian C-clade viral sequences (*n*=138) in this cohort clustered closely to the other reference C-clade sequences from LANL HIV Sequence Database, http://www.hiv.lanl.gov/ (**Figure 4.1**). Clustered sequences represent epidemiologically linked infections with short durations of time between transmissions and thus clusters represent the leading edge of the epidemic (Ragonnet-Cronin *et al*., 2016). In order to closely evaluate VRC, the Indian Consensus (IC) sequence was generated as the representative C-clade p24 Gag sequence for the cohort. IC has only three amino acids different from the C-clade Consensus in p24 Gag: I147L, I256V and D260E. Comparison of IC, SK-254(M) and NL4-3 is shown in **Table 4.4**. These sequences were the three backbones of the chimeric viruses used in the following fitness assay.

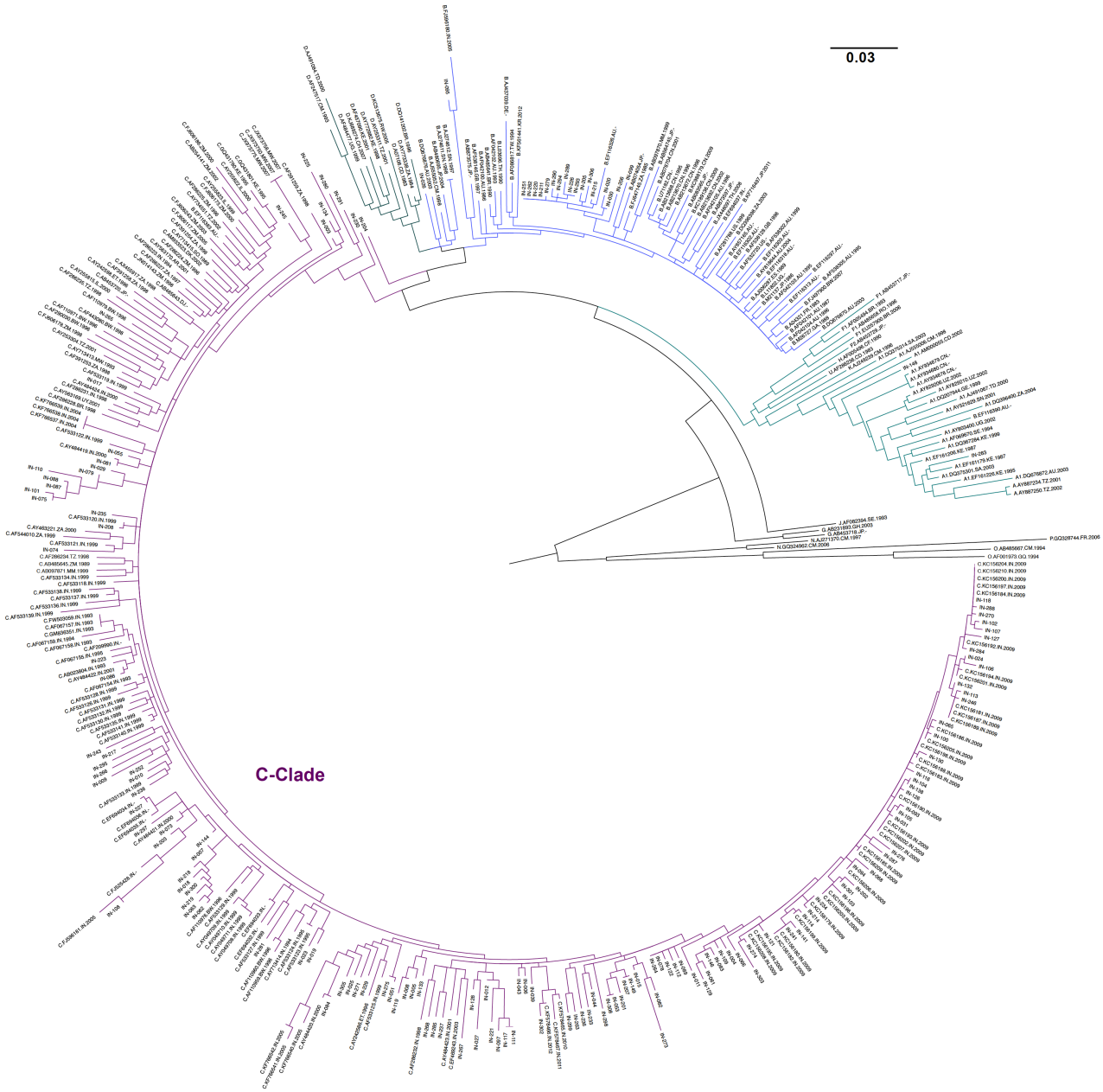


Figure 4.1. Maximum-likelihood phylogenetic tree of p24 Gag sequences of C-clade-infected individuals. The C-clade sequences are shown in purple (*n*=138). Subjects recruited in our cohort are numbered 'IN-.' Other reference sequences are shown in different colours. All the reference sequences are from the Los Alamos database (LANL HIV Sequence Database, http://www.hiv.lanl.gov/).

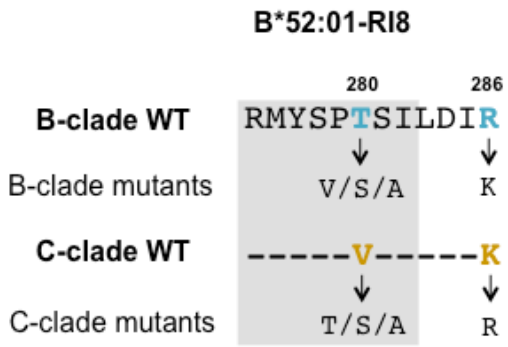
Table 4.4. p24 Gag sequences of NL4-3, SK254(M) and IC.

Plasmid	HXB2 Gag position
	147 159
NL4-3	P I V Q N L Q G Q M V H Q A I S P R T L N A W V K V V E E K A F S P E V I P M F
SK-254(M)	- I - - - - - - - - - - - - -
IC	- - - - - - - - - - - - - - L - - - - - - - - - - - I - - - - - - - - - - - - -
	173 203
NL4-3	S A L S E G A T P Q D L N T M L N T V G G H Q A A M Q M L K E T I N E E A A E W
SK-254(M)	T - D - - - - - - - - -
IC	T - D - - - - - - - - -
	248 252
NL4-3	D R L H P V H A G P I A P G Q M R E P R G S D I A G T T S T L Q E Q I G W M T H
SK-254(M)	- A - - - S
IC	- A - - - S
	256 260 280 286
NL4-3	N P P I P V G E I Y K R W I I L G L N K I V R M Y S P T S I L D I R Q G P K E P
SK-254(M)	- - - - - - - D - - - - - - - - - - - - - - - - - - - V - - - - - K - - - - - -
IC	- - - V - V - - - - - K - - - - - -
	301 310 312 319
NL4-3	F R D Y V D R F Y K T L R A E Q A S Q E V K N W M T E T L L V Q N A N P D C K T
SK-254(M)	- - - - - - - - F - - - - - - - - T - D - - - - - - D - - - - - - - - - - - - -
IC	- - - - - - - - F - - - - - - - - T - D - - - - - - D - - - - - - - - - - - - -
	335 357
NL4-3	I L K A L G P G A T L E E M M T A C Q G V G G P G H K A R V L
SK-254(M)	- - R - S - - - - - -
IC	- - R - S - - - - - -

*4.3.2 HLA-B*52:01-positive individuals had higher frequencies of the V280X mutation and higher CD4 counts*

Similar to the findings in B-clade cohorts, the distinctive HLA-B*52:01 footprint at position 6 in its p24 Gag RI8 epitope could be seen in our C-clade infected individuals (**Figure 4.2 A**). However, it is important to note that the consensus residue at Gag-280 in C-clade virus is Valine (as opposed to Threonine in B-clade) and the consensus residue at Gag-286 in C-clade virus is Lysine (as opposed to Arginine in B-clade: **Table 4.5**).

Table 4.5. Amino acids at Gag 280 and 286 of B- and C-clade viruses.



The frequency of the V280X mutation was significantly higher in HLA-B*52:01-positive individuals when compared with HLA-B*52:01-negative individuals (**Figure 4.2 A** and **B**; 30% versus 71% in the respective groups, $p<0.0001$ by Fisher's exact test, $q=0.02$). These suggest that changes from the wild-type/consensus sequence at this position are driven by HLA-B*52:01, just as is in B clade infection (Carlson *et al*., 2012a, Murakoshi *et al*., 2015). Furthermore, our cohort showed that HLA-B*52:01 is a protective allele, strongly associated with higher CD4 count, with corrected *p* value <0.0001 (Bonferroni's test, *p* value threshold = 0.0038; **Figure 4.2 C**). Differences among HLA-B alleles had a significant impact on CD4 count ($p=0.004$). The previously identified 'protective alleles' such as HLA-B*57:01, HLA-B*27 and HLA-B*58:01 were associated with higher CD4 counts, whilst HLA-B*35 alleles had

significantly lower CD4 counts (**Figure 4.2 C**). In particular, HLA-B*35:03 was significantly correlated with lower CD4 count and this allele has been previously reported to associate with worse disease outcomes (Gao *et al*., 2001; Juarez-Molina *et al*., 2014). **Figure 4.2 D** further shows that the V280A mutant in HLA-B*52:01-positive individuals was strongly associated with a significantly higher CD4 count (p=0.004). Unexpectedly, any V280X mutant trended towards high CD4 count; this is highly unusual, since in the great majority of cases escape mutations are associated with higher viral loads and therefore lower CD4 counts (Carlson *et al*., 2016). In contrast, the V280X mutations appeared to make no difference among the HLA-B*52:01-negative subjects. Taken together, HLA-B*52:01 was highly protective in this C-clade infected cohort; and, among the HLA-B*52:01-positive subjects studied, the V280X escape mutants were associated with higher CD4 counts, and in particular the V280A mutant in the viral sequence was associated with especially favourable CD4 counts.

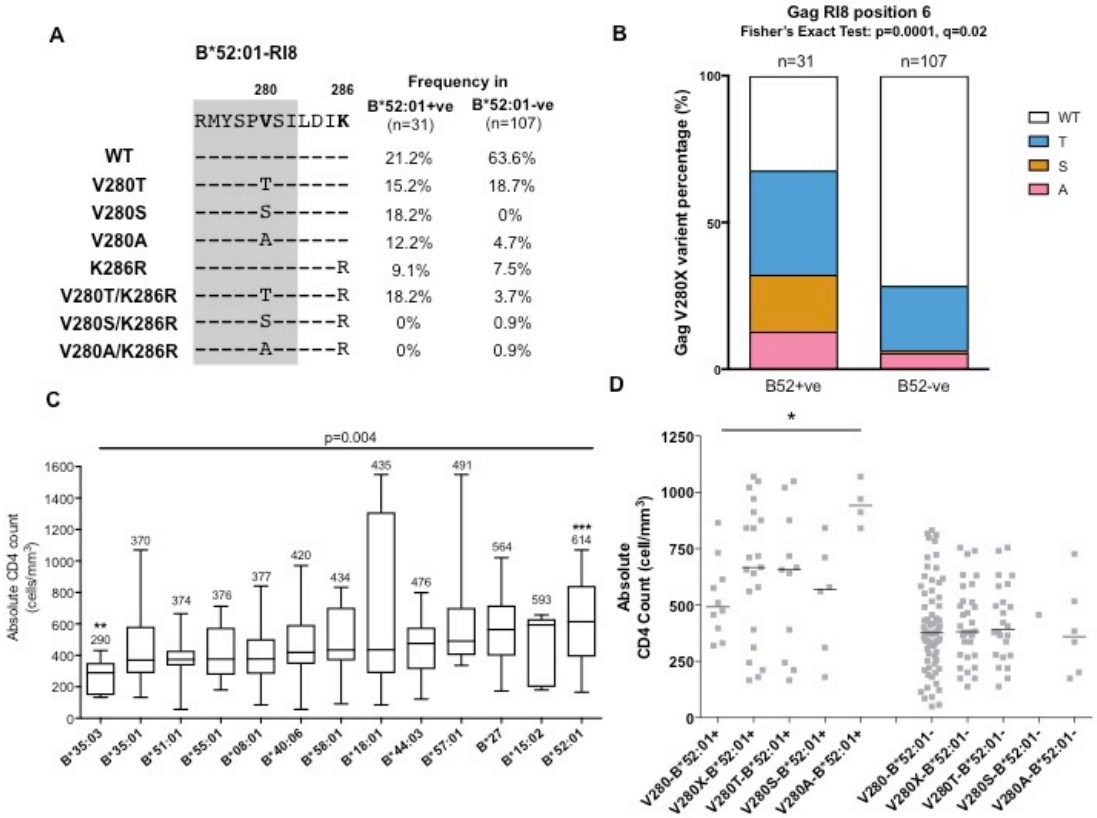


Figure 4.2. Frequency of B*52:01-restriced mutations and the association between CD4 count and HLA-B alleles (*n*=138). (A) Frequency of Gag V280X with/without K286R in HLA-B*52:01-positive and -negative C-clade infected individuals. **(B)** Comparison of the percentages of wild-type and variants at Gag 280 in HLA-B*52:01-positive and -negative individuals. $p<0.0001$ (Fisher's exact test); $q=0.006$. **(C)** Association between HLA-B alleles and CD4 counts. Vertical bar represents the median CD4 counts, boxes indicate the IQR and brackets are the maximum and minimum values. Kruskal-Wallis test was applied ($p=0.004$). Here the individual *p* value was corrected by Bonferroni's correction (*p* value threshold=0.0038). **(D)** CD4 counts of each V280X mutant in HLA-B*52:01 B*52:01-positive and -negative individuals. Bar represents the median of CD4 counts. Kruskall-Wallis test was applied ($p=0.03$). Mann-Whitney U-test and Steel-Dwass test was performed for post-hoc analysis ($p<0.05$ of V280A when compared to WT). (* $p<0.05$, ** $p<0.01$, *** $p<0.001$)

*4.3.3 None of HLA-B*52:01-positive individuals carried V280S and V280A with K286R*

In our study, the Gag 286 change subsequent to the polymorphism at 280 was also observed in HLA-B*52:01-positive subjects (**Figure 4.3 A**). The frequency of K286R alone was not significantly different between the HLA-B*52:01-positive and -negative group, despite the fact that the percentage acquiring this mutation was still higher in the positive group. When K286R and V280X were present together, they were significantly more likely to be found in HLA-B*52:01-positive individuals ($p= 0.027$). These data indicate that the K286R association with HLA-B*52:01 is dependent upon the V280X mutant, and arises subsequent to it. However, the K to R substitution only occurred with the presence of wild-type V280 and V280T, not in the subjects carrying V280S and V280A ($p=0.012$) (**Figure 4.2 A** and **Figure 4.3 B**). In the HLA-B*52:01-negative group, however, the frequencies of K286R together with V280X were not significantly different according to the specific residue at Gag-280 (**Figure 4.3 B**). No difference could be found in the association of CD4 count and V280X with or without K286R in either HLA-B*52:01-positive or -negative individuals (**Figure 4.3 C**). Therefore, whether K286R serves as a compensatory mutation for V280X is not clear,

based on these CD4 data. However, given that the association between K286R and V280X only applies in the case of V280T, this suggests an effect which may be only apparent when confounding sequence variables are removed in in vitro fitness experiments.

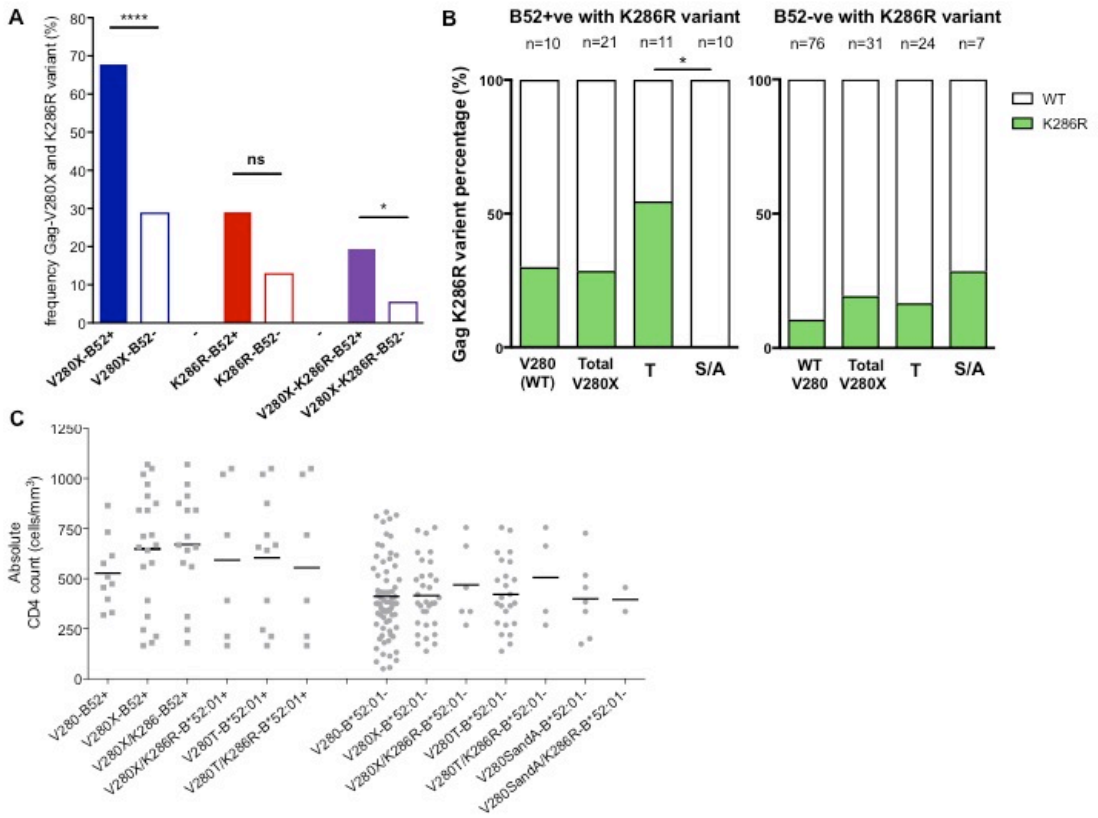


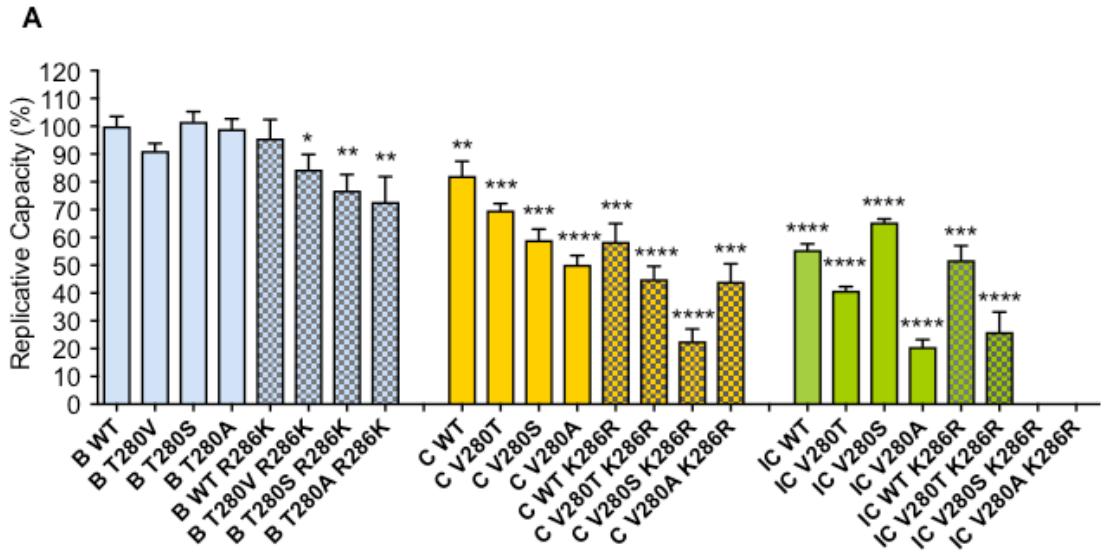
Figure 4.3. Frequency of Gag K286R and the CD4 count of each mutant group (*n*=138). **(A)** Frequency of Gag V280X, K286R, V280X/K286R in HLA-B*52:01-positive and -negative individuals. **(B)** Comparison of the percentages of K286R in wild-type and V280X in HLA-B*52:01-positive and -negative individuals. **(C)** CD4 count of wild-type and V280X with/without K286R in HLA-B*52:01-positive and -negative individuals. Bar represents the median of CD4 count. Fisher's exact test for frequency difference; and Mann-Whitney U-test and Steel-Dwass test post-hoc analysis for CD4 counts were applied. (* $p<0.05$, **** $p<0.0001$)

4.3.4 V280S and V280A with K286R in IC backbone crippled VRC

To examine whether the mutations at Gag 280 and 286 affect VRC, viral fitness assays were performed. Based on **Table 4.4**, The T280X with or without R286K mutants was introduced into B-clade backbone (NL4-3), whereas mutations of V280X with or without K286R were situated in C-clade backbone including both SK-254(M) and IC (Indian Consensus). The exponential growth of the viruses was measured and the results are shown as the relative GFP+% compared to the wild-types. **Figure 4.4 A** displays the relative VRC values of all the tested viruses compared to NL4-3. Overall, as expected, VRCs of C-clade viruses were lower than those of B-clade, similar to our previous findings in the Gag-protease-GXR-cell system fitness assay (Juarez-Molina *et al*., 2014; Payne *et al*., 2014; Adland *et al*., 2015). The IC viruses were the least fit of all, which means I147L, I256V and D260E (the residues unique to the IC virus of those studied here) contributed to lowering VRC in C-clade viruses. However, acquiring these three mutations did not necessarily result in a higher CD4 count, which suggests having a low fitness virus is not the sole factor responsible for better disease outcomes (**Figure 4.5 A**). HLA-restricted CTL pressure, escape or compensatory mutation(s), viral fitness, and the patient's immune profiles, all play a role in maintaining high CD4 counts.

Figure 4.4 B, C and D shows each virus's relative VRC value to their own template wild-type. Changes at position 280 had the least impact on B-clade viruses (**Figure 4.4 B**). None of T280X mutant viruses had significantly different VRCs compared to the wild-type. In C-clade SK-254(M) viruses, V280S and V280A mutants lowered the VRC significantly (**Figure 4.4 C**). T280V and V280T did not have a great impact on the VRC in B- and C- clade, respectively, whereas in IC, VRC of V280S was similar to the IC-wild-type. V280A in both SK-254(M) and IC was the most destructive variant (**Figure 4.4 C** and **D**).

Unexpectedly, changes at Gag 286 subsequent to Gag 280 significantly *lowered* the VRC across all backbones (**Figure 4.4 A**-**D**). In the absence of a consensus variant at Gag-280, changes at Gag-280 made little difference to VRC, only in C-clade SK-254(M) did the Gag-286 variant significantly reduce viral fitness (p=0.024; **Figure 4.4 C**). However in the presence of Gag-280 variation from consensus, changes at Gag-286 not only failed to restore VRC in the Gag 280 mutants, but also further lowered the viral fitness. This argues against Gag-286 variation being a compensatory change to escape mutants selected at Gag-280. In IC, particularly, the K286R mutant actually crippled viruses carrying either the V280S and V280A mutation to such an extent that the VRCs were too low to quantify. The impairment of VRC is consistent with the absence of K286R with either V280S or V280A in HLA-B*52:01-positive individuals (**Figure 4.3 B**).



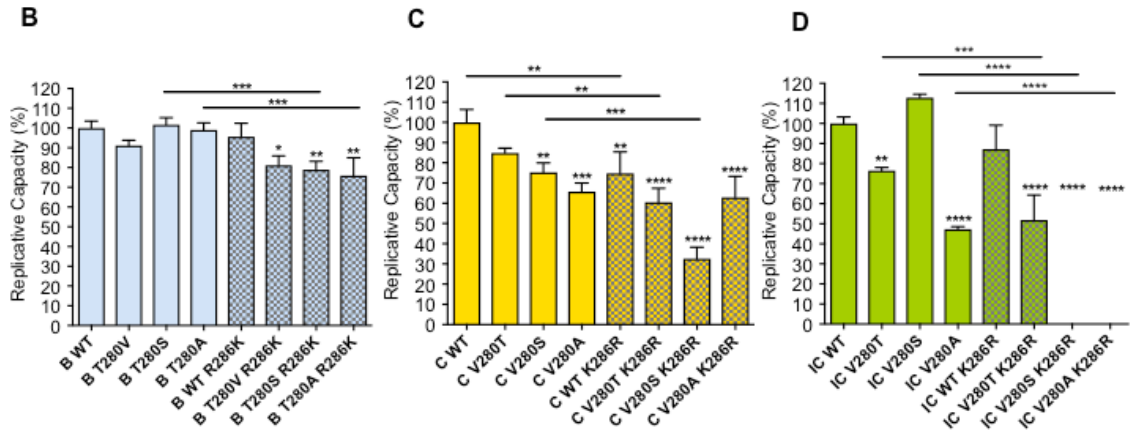


Figure 4.4. Viral replicative capacity (VRC) of T280X/R286K B-clade and V280X/K286R C-clade mutant viruses. (A) Relative VRC (%) of all tested viruses to B-clade wild-type (NL4-3). **(B)** Relative VRC (%) of B-clade viruses to B-clade wild-type (NL4-3). **(C)** Relative VRC (%) of C-clade viruses to C-clade wild-type [SK-254(M)]. **(D)** Relative VRC (%) of IC viruses to IC wild-type. All bars are shown as mean with SD. Patterned bars represent R286K/K286R mutant. Colour blue stands for B-clade viruses (NL4-3), yellow shows C-clade SK-254(M), and green represents C-clade IC. One-way ANOVA multiple comparison and Dunnett's correction were used for calculating *p* value. Asterisks marked on top of each bar are the *p* value when each mutant virus was compared to the wild-type. (* $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$)

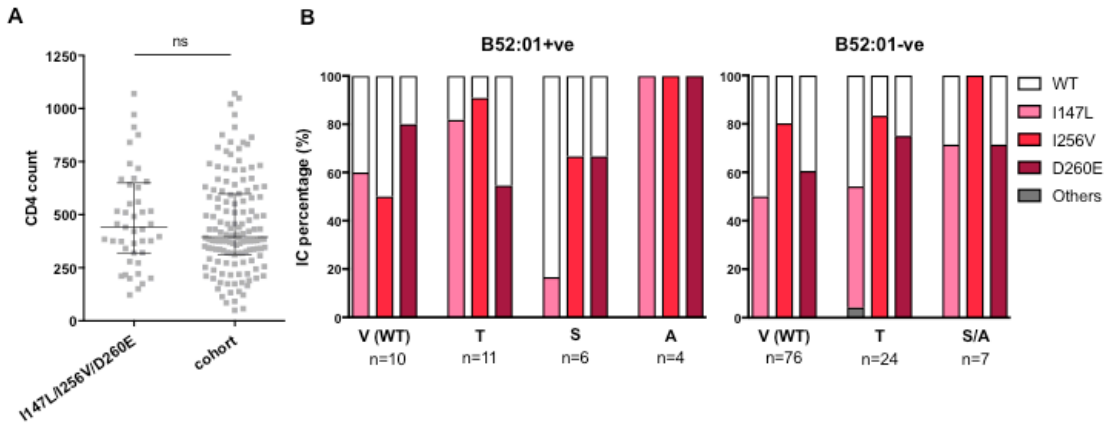


Figure 4.5. CD4 count and the percentages of subjects carrying IC sequence. (A) Comparison of CD4 count between all subjects (*n*=138) and ones carrying all I147L/I256V/D260E variants (*n*=45). Mann-Whitney U-test was applied. **(B)** Percentages of acquiring I147L, I256V and D260E variant in HLA-B*52:01-positive and -negative individuals. V280S and V280A subjects were grouped together in HLA-B*62:01-negatives due to the small size of V280S group (*n*=1). (ns, $p\geq0.05$)

In the Indian study cohort, 95% of individuals carried at least one of three mutations I147L/I256V/D260E observed in the IC sequence, and 100% of the HLA-B*52:01-positive subjects whose autologous virus encoded V280A had all three of these IC-specific mutations (**Figure 4.5 B**). This collectively suggests that the IC sequence well represented the actual circulating viruses, and the low fitness due to V280A might indeed be one of the factors responsible for patients' high CD4 count (**Figure 4.2 D**).

4.3.5 Emergence of compensatory mutations L135F and I223T for V280S/K286R and V280A/K286R during viral inoculation

For some less fit viruses, the inoculation time for the viral growth could take up to 60 days. The crippled viruses such as V280S/K286R and V280A/K286R were set up in at least in three different experiments up to 100 days; GFP+% remained negative indicating the low VRC was below the limit of detection. Viral sequences were confirmed by RNA extraction from the harvested supernatant. It was possible and indeed likely that viruses would develop *in vitro* mutations during this long culture time. For the first time, we observed the developed variants in late C-clade culture. L50F in p17 Gag appeared in V280A SK-254(M) sequence, while L135F and I223T emerged in IC V280S/K286R and V280A/K286R, respectively (**Figure 4.6 A**). In C-clade backbone, L50F further lowered the VRC of V280A by 20%. This was not beneficial for the virus in terms of restoring replicative capacity (**Figure 4.6 B**). In contrast, L135F in IC effectively reinstated the VRC of V280S/K286R to almost 80%, not significantly different from the wild-type (**Figure 4.6 C**). Although L135F was unable to fully restore V280S/K286R to the same extent as the VRC of V280S alone, it was a competent compensatory mutation for V280S/K286R. I223T served as a compensatory mutation which only partially recovered the VRC to a low level. The VRC did not reach significance but I223T was at least competent enough in the

detection range of the assay. Both L135F and I223T were harvested from cell culture on Day 73 and 71, respectively (data not shown). It should be noted that this culture time for L135F and I223T was much longer than L50F, which was harvested on Day 43. All of the mutations were confirmed by sequencing.

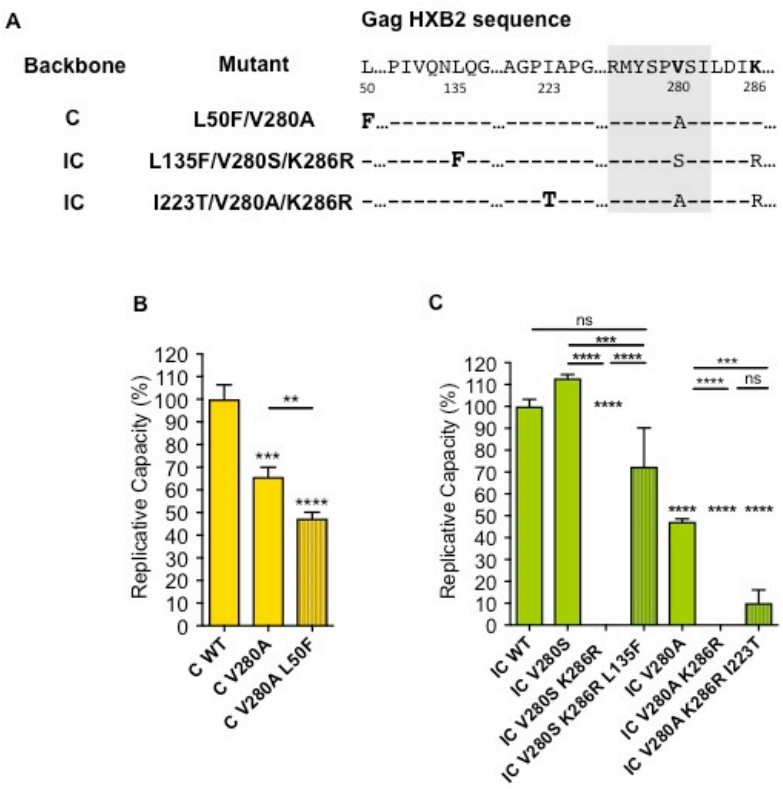


Figure 4.6. Identification and VRC of the developed variants during viral inoculation. **(A)** Positions of the developed variants in Gag HXB2. **(B)** VRC(%) of L50F/V280A in C-clade SK-254(M). **(C)** VRC(%) of L135F/V280S/K286R and I223T/V280A/K286R in IC. All bars are shown as mean with SD. Striped bars represent developed mutants. Colour yellow stands for C-clade SK-254(M), and green represents C-clade IC. One-way ANOVA multiple comparison and Dunnett's correction were used for calculating *p* value. Asterisks marked on top of each bar are the *p* value when each mutant virus was compared to the wild-type. (ns, p≥0.05; ** *p*<0.01, *** *p*<0.001, **** *p*<0.0001)

4.4 Discussion

This cohort study is the first one to evaluate HLA-B*52:01-associated immune control of HIV in C-clade infection. As with previous studies in B-clade, our cohort shows that HLA-B*52:01 was strongly associated with a better disease outcome in terms of having an average higher CD4 counts. In parallel to Murakoshi *et al*., 2015 research, HLA-B*52:01 restricted position 6 in the RI8-epitope in p24 Gag and mutations from the wild-type sequence led to a fitness cost to the virus. However, we found in B-clade viruses that all T280X mutations did not lower VRC significantly in our fitness assay.

In this C-clade infected cohort, the mean CD4 count (448 cell/mm^3) was not only higher than the rest of chronically infected adult cohorts in this thesis (**Section 2.1**), but was also very close to healthy CD4 count range (500-1,400 cell/mm^3). We were also able to compare the CD4 count of each V280X escape mutation. Strong CTL targeting Gag exerts a selection pressure on the virus. However, an escape mutation in the highly conserved capsid protein often comes with a cost to viral replicative capacity, and the benefit to the virus overall of evading the CTL response may be almost cancelled out by the reduced replicative efficiency. Nonetheless there must be some short-term benefit to the virus for the mutant to be selected at all. It is surprising therefore that the V280X mutants are all associated with higher CD4 counts than the wild-type V280. In particular, the V280A in the Gag RI8 epitope was associated with an especially high CD4 count. One of the possible explanations might be that the VRC was not competent enough (only 47.3% to IC wild-type and 20.6% to NL4-3 wild-type) to result in high viral load and hence low CD4 count for the time period studied. Analysis was limited by the absence of viral load data from this cohort, but the HLA-B associations observed here between HLA-B*57/58:01 and B*52:01 with high CD4 counts and therefore protection against disease progression,

and between HLA-B*35 and in particular HLA-B*35:03 with low CD4 counts and therefore more rapid disease progression are remarkably consistent with other studies showing HLA associations with low and high viral loads, respectively (Goulder and Walker, 2012).

Studies have shown low VRC contributes to suppression of viraemia (Friedrich *et al*., 2004; Leslie *et al*., 2004; Fernandez *et al*., 2005; Kobayashi *et al*., 2005). However, low VRC alone does not prevent HIV disease progression and the contribution of all components of the immune system to control of HIV, apart from the single RI8-specific HLA-B*52:01-restricted response and Gag-280 variant being considered here, are critically important to disease outcome.

In addition to the immune response, the viral backbone as a whole has a major influence on the impact of Gag-280 variants on VRC, as illustrated clearly by use of 3 backbones in the current experiments. The putative compensatory mutation at Gag-286 is a single residue that we have considered here that alters the VRC. Intriguingly, R286K in B-clade and K286R in C-clade viruses not only failed to rescue the VRC of T280A and V280X, but also served as another potential escape mutation that was co-selected with Gag-280. It is still unclear why the mutant was developed that lowered VRC. It is also not a random independent attempt either since the R286K/K286R was subsequent to T280/V280. The RI8 epitope covers partial capsid helix 7 of the N-terminal domain (NTD; CA 143-145, Gag 275-277), and the whole interdomain region (CA 146-150, Gag 278-282). Despite the proximity, 280 is located in the middle of interdomain linker structure, while 286 is responsible for the construction of Helix 8. It is very unlikely that these two mutations have a compensatory relationship because they are responsible for different structures of the capsid protein (CA). Studies in B-clade modelling showed that 70-80% of mutants in the interdomain region still retain some viability (Rihn *et al*., 2013). In one

study, T280A was still able to function as wild-type (Jiang *et al*., 2011). Changes in the helix that composes the C-terminal domain (CTD) are more fragile. In particular, K286R/R286K is located in the highly conserved major homology region (MHR; CA 153-172, Gag 285-304). Changes in this region can affect the structural stability since MHR aids helix 2 of NTD for capsid dimerization (Gamble *et al*., 1997). MHR is known to be crucial for viral assembly and virion maturation (Patarca and Haseltine, 1985; Wills and Craven, 1991).

These modelling data based on the B-clade sequence supported our viral fitness findings. In B-clade, T280X seem to be the perfect candidate for the virus to develop an escape mutation since it had mild influence on VRC, but together with R286K, the fitness cost was significant for T280S and T280A. It is likely that mutants in the interdomain linker, in spite of the weaker influence on VRC itself, might change the interaction between NTD and CTD, putting pressure on position 286 to change into a different amino acid for increased stability. The interaction between NTD and CTD was crucial for proper capsid formation (Byeon *et al*., 2009; Zhao *et al*., 2013). However, V280X in C-clade viruses did not necessarily diminish the VRC (e.g. IC V280S) and the influence on VRC was more substantial, which suggests that the structural relationship is more complicated than it appears. This might be due to the various interactions with host cellular factors in the interdomain binding pocket (Gamble *et al*., 1997; Bhattacharya *et al*., 2014). With regards to IC V280S and V280A, the virus can hardly afford to develop K286R as an approach to adjust either its structural change or altered interacting interface since it completely devastated VRC and there was no HLA-B*52:01-positive individual carrying these sequences.

Another possible explanation for the selection of K286R is its potential to impair proteosomal processing as first type of escape mechanism (section 1.6.1; Allen *et al.*, 2004; Milicic *et al*., 2005). One characteristic of this type of escape is that the

mutation is normally located outside of the epitope (Jones *et al*., 2004; Yokomaku *et al*., 2004). K286R is only downstream of RI8 epitope by four amino acids, which can affect the binding of TAP on either C- or N-terminus (Schmid *et al*., 2009), or prevents cleavage from ER aminopeptidase (Draenert et al., 2004).

Of note, mutants T280V in B-clade and V280T in C-clade were tested along with S and A at this position was only due to the difference between B- and C-clade consensuses. It is unlikely that Valine or Threonine were naturally selected as an escape mutation as Serine or Alanine. In fact, T280V and V280T made either no or relatively mild impact on VRC compared to S or A in different backbone (**Figure 4.4 B**, **C** and **D**), and the combination with the presence of K286R in C-clade infection was only observed in HLA-B*52:01 positive individuals carrying V280T (**Figure 4.3 B**). These data indicated that T280V or V280T hardly served as the same role as the change to Serine or Alanine.

Unexpectedly, we observed two valid compensatory mutations, L135F for V280S/K286R and I223V for V280A/K286R, during the long viral inoculation (~100 days). The accumulation of compensatory mutation(s) is not unheard of after several passages through the host cell (González-Ortega *et al*., 2011). As the evolutionary speed of HIV is very fast, the virus will be able to cope with the effective CTL killing and rescue itself from having impaired VRC, although whether these selected mutations last throughout the course of infection is uncertain. The long-term compensatory mutations that restore VRC and maintain the escape from CTL epitopes are more likely to remain in the sequence. L50F in SK-254(M) C-clade virus was possibly one attempt the virus tried to make but the mutation did not successfully raise the VRC. The time for developing this mutation was much shorter than the other effective compensatory mutations L135F and I223V, 43 days compared to 73 and 71 days, respectively. It is reasonable that the latter two viruses

required more time to develop a mutation that was capable of coordinating the interaction between its structural domains while at the same time retaining or even improving VRC. The mechanism of how these two mutations located in β-hairpin (CA 3, L135F) and Cycophilin A loop (CA 91, I223V) compensate the V280S/K286R and V280A/K286R, respectively, remains unclear.

Apart from VRC, there are other various factors that might play roles in modulating the CD4 counts. For instance, the low VRC of having I147L, I256V and D260E as IC alone had limited impact on CD4 counts (**Figure 4.6 A**). In the study of Sakai *et al*., 2015, there was no correlation between VRC and viral load or CD4 count, even in the subjects expressing the protective alleles. This indicates the complexity of the evolving relationship between HLA-restricted CTL and VRC.

Other than HLA-B-restricted CTL pressure, HLA-C alleles might also play a role in the immune control of HLA-B*52:01-positive individuals. The reason for this is that HLA-B*52:01 is in tight linkage disequilibrium (LD) with HLA-C*12:02 (Naruto *et al*., 2012; Murakoshi *et al*., 2016). This LD was also observed in our B-clade infected Thames Valley cohort where HLA-B*52:01 is strongly associated with HLA-C*12:02. The HLA-B*52:01- and HLA-C*12:02-haplotype was the most prevalent and protective in the Japanese cohorts. The HLA-C-restricted CTL responses and the escape mutations were associated with a lower plasma viral load (Murakoshi *et al*., 2016). In fact, HLA-C*12 is one of the highest expressed HLA-C molecules and like other highly expressed HLA-C molecules, itself is therefore associated with improved control of HIV (App *et al*., 2013). Like HLA-B*52:01, which is a Bw4-80I molecule and therefore a ligand for KIR3DL1 (Martin *et al*., 2007) HLA-C alleles are also the ligands for killer cell immunoglobulin-like receptors (KIR). By reducing the binding affinity of KIR2DL2 to the respective HLA complexes, HLA-C*12:02 induces the KIR2DL2+ NK cell activity to achieve effective viral control (Lin *et al*., 2016). The

relationship between certain HLA class I molecules such as HLA-B*52:01, and immune control of HIV is therefore a complex one, requiring more than the interaction between the dominant Gag CD8+ T-cell response and viral replicative capacity to explain it.

Taken together, HLA-B*52:01 is a protective allele not only in HIV B- but C-clade infection. The fitness cost of the mutants at Gag-280, the same P6 position in the targeted epitope might vary due to different structural changes, but the low VRC might be one major factor that benefits the host, resulting in relatively high CD4 counts. Of note, in other autoimmune and infectious diseases, HLA-B*52 (HLA*52:01) has been associated with worse prognosis and outcomes. It is highly susceptible to Takayasu arteritis (Terao *et al*., 2014; Terao *et al*., 2015; Chan *et al*., 2016), cutaneous Leishmaniasis lesions (Ribas-Silva *et al*., 2015), Henoch-Schönlein purpura (Ren *et al*., 2012), Behçet's disease (Soto-Vega *et al*., 2004), Crohn's disease (Kinouchi *et al*., 2003), and Dengue fever (Stephend *et al*., 2002). Evidently, the underlying mechanism for the diverse influences of HLA-B*52:01 in various diseases is yet to be defined.

4.5 Addendum

As previously mentioned, studies of HLA-B*52:01 were mainly done in B-clade infection. Murakoshi and his colleagues have shown that T280X was restricted by HLA-B*52:01 (*n*=358), and all three mutations T280V, T280S and T280A lowered the viral fitness in terms of p24 Gag antigen detected using ELISA (Murakoshi *et al*., 2015). This is different from our findings, where only T280V successfully diminished the fitness in B-clade virus. It might be due to the fact that we used the GXR cell line instead of primary CD4+ cells and measured GFP+ cells percentage for viral fitness. Instead of quantifying the extracellular viral production, our fitness assay was based on the infected cells. In order to verify if T280S, T280A and T280V reduced the

actual amount of p24 Gag protein in our assay, we set up an ELISA following the protocol provided from Leidos Biomedical Research. During the 10-day fitness assay, supernatant was collected on a daily basis and saved for ELISA.

4.5.1 p24 Enzyme-linked immunosorbent assay

An in-house ELISA was run to quantify p24 capsid antigen in the culture supernatant at Day 2, 4, 6, 8 and 10 in fitness assay. All samples' supernatant (in triplicate) were collected before live/dead cell staining by centrifugation at 2,000rpm for 5 minutes, followed by filtration of the supernatant through a 0.22μm filter to ensure it was cell-free. The supernatant was stored at -80°C before use. Capture antibody, $p24^{CA}$ Standard, primary antibody, and the protocol were kindly provided by Ms. Rodman B. Smith from Frederick National Laboratory for Cancer Research, AIDS and Cancer Virus Program. To coat the plates, aliquots of Capture antibody were thawed in a 37°C water bath for 5 minutes and then diluted 1:370 in PBS, mixed at room temperature for 30 minutes. 100μl of Capture antibody solution was added to each well of a flat-bottomed 96-well high binding polystyrene plate (Greiner Bio-One). The plate was sealed with a plate sealer (MP Biomedicals UK) and stored overnight at 4°C.

On Day 2, Blocking solution was prepared as 1% BSA (Insight Biotechnology) in PBS. The solution must be made fresh each time and was allowed to mix for 1 hour before use. The plate was removed from 4°C, the plate sealer was discarded and Capture antibody solution was aspirated. The plate was pat-dried on clean paper towels. 300 μl of Blocking solution was added to each well and incubated at room temperature for 30 minutes. Blocking solution was then aspirated and pat-dried on clean paper towels. The plate was stored upside down on absorbent towel at 4°C overnight.

For Secondary antibody optimal concentration determination, peroxidase-labeled antibody to rabbit IgG (H+L) (Insight Biotechnology) as Secondary antibody was reconstituted on Day 3. 50% Glycerol was added as instructed in KPL product 074-1516 Procedure A, followed by an additional 1:100 dilution in 50% glycerol. Four dilutions (1:50, 1:100, 1:200, 1:500) in 0.45-micron-filtered Secondary antibody diluent containing 2% NMS, 5% NGS, and 0.01% Tween 20 (Sigma) in RPMI were prepared for testing. Plate layout is shown in **Figure 4.7 A**. The plate was washed 5 times with 1x Wash buffer (Perkinelmer LAS) and blotted on a clean paper towel. 100 µl $p24^{CA}$ Standard (Leidos Biomedical Research) at a concentration of 40,170 pg/ml was added to wells in row B, D, F and H from column 1 to 10, while filtered Sample diluent containing 1% BSA and 0.2% Tween 20 in RPMI was added in 4 wells in column 11. Negative controls were defined as column 12 left blank and empty throughout the assay until the substrate incubation step. The plate was sealed and incubated at 37°C with 5% CO_2 for two hours. Primary antibody (Leidos Biomedical Research) was diluted in 200-fold in filtered Primary antibody diluent containing 10% FCS and 2% NMS in RPMI. 100µl of diluted Primary antibody was added in the four rows except column 12 and incubated at 37°C with 5% CO_2 for one hour. The plate was washed 5 times as before. 100 µl Secondary antibody at different dilutions was added accordingly with an additional 1:2 serial dilution each row and incubated at 37°C with 5% CO_2 for one hour. The plate was washed again as before. 100 µl of TMB peroxidase substrate (Insight Biotechnology) was added to every well in row B, D, F, H and incubated at room temperature for 30 minutes. The reaction was stopped by adding 100 µl 1N HCL (Sigma) to all the wells.

The plate was read at 450nm with a reference 650nm as a dual wavelength setting on a SpectraMax 340PC384 Microplate Reader (Molecular Devices LLD, USA). The best dilution was determined as optimal ODs of the top Standard >3.0 and

background <0.1. In our testing, 1:100 dilution (1:10,000 final dilution) was selected. Single use aliquots were prepared and stored at -20°C.

For measuring p24 Gag in the samples, all the filtered supernatant was lysed by adding 1/10 volume of lysing solution containing 10% Triton X-100 (Sigma) in Milli-Q H_2O (Sigma). Samples were diluted at 1:5, 1:10, 1:100 and 1:500 in Sample diluent. On Day 3, plates were washed 5 times with 1x Wash buffer and blotted as previously described. 100 µl of the sample was added in triplicate as indicated in **Figure 4.7 B**. $p24^{CA}$ Standard was diluted at 1:2 in Sample diluent for ten different concentrations and added to the plate accordingly. Each plate contains two wells of 100 µl Sample diluent and two blank wells which were left empty until the TMB step as described above. Plates were incubated at 37°C with 5% CO_2 for two hours. The following steps were the same as previously described by using 1:200 diluted Primary and 1:100 diluted Secondary antibody.

A

	1	2	3	4	5	6	7	8	9	10	11	12
A		1:50										
B	STD	→	→	→	1:2 dilution	→	→	→	→	→	SD	Neg.
C		1:100										
D	STD	→	→	→	1:2 dilution	→	→	→	→	→	SD	Neg.
E		1:250										
F	STD	→	→	→	1:2 dilution	→	→	→	→	→	SD	Neg.
G		1:500										
H	STD	→	→	→	1:2 dilution	→	→	→	→	→	SD	Neg.

B

	1	2	3	4	5	6	7	8	9	10	11	12
A	STD #1	STD #1	STD #2	STD #2	STD #3	STD #3	STD #4	STD #4	STD #5	STD #5	STD #6	STD #6
B	STD #7	STD #7	STD #8	STD #8	STD #9	STD #9	STD #10	STD #10	SD	SD	Blank	Blank
C	Sample #1	Sample #1	Sample #1	Sample #7	Sample #7	Sample #7	Sample #13	Sample #13	Sample #13	Sample #19	Sample #19	Sample #19
D	Sample #2	Sample #2	Sample #2	Sample #8	Sample #8	Sample #8	Sample #14	Sample #14	Sample #14	Sample #20	Sample #20	Sample #20
E	Sample #3	Sample #3	Sample #3	Sample #9	Sample #9	Sample #9	Sample #15	Sample #15	Sample #15	Sample #21	Sample #21	Sample #21
F	Sample #4	Sample #4	Sample #4	Sample #10	Sample #10	Sample #10	Sample #16	Sample #16	Sample #16	Sample #22	Sample #22	Sample #22
G	Sample #5	Sample #5	Sample #5	Sample #11	Sample #11	Sample #11	Sample #17	Sample #17	Sample #17	Sample #23	Sample #23	Sample #23
H	Sample #6	Sample #6	Sample #6	Sample #12	Sample #12	Sample #12	Sample #18	Sample #18	Sample #18	Sample #24	Sample #24	Sample #24

Figure 4.7. p24 ELISA 96-well plate setup. (A) Secondary antibody optimal dilution determination layout. $p24^{CA}$ Standards are shown in blue shading in row B, D, F and H. 2^{nd} antibody at four dilutions (1:50; 1:100, 1:250, 1:500) were added to each row and a further 1:2 serial dilution of each diluted concentration was performed from column 1 to 10. Orange and grey shading represents Standard Diluent (SD) and blank wells as negative controls. **(B)** Sample layout. 10 $p24^{CA}$ Standards in 1:2 dilution, SD and blank wells are shown within row A and B in duplicate. Sample wells are shown in the rest of the wells in triplicate.

4.5.2 Statistical analysis of p24 ELISA

The result of the p24 ELISA data were analysed by using the online tool elisaanalysis.com (http://elisaanalysis.com/app). The ODs of $p24^{CA}$ Standard (40,170 pg/ml) in 1:2 serial dilutions were fit into a 4-parameter logistic regression model to generate a standard curve. In the cases of unavailable data due to the maximum and minimum limit of the model equation, linear regression analysis in Excel was used.

4.5.3 Results of p24 Gag ELISA were comparable to the fitness assay

First we compared the p24 Gag concentrations between B- and C-clade viruses. **Figure 4.8 A** illustrates that except for the first time point, the production of p24 Gag was at a similar level between the wild-types of these two clades, which is different from the results of measuring GFP+ cell percentages where B-clade wild-type NL4-3 was more fit and C-clade wild-type SK-254(M) peaked much later (Day 9; **Figure 4.8 D**).

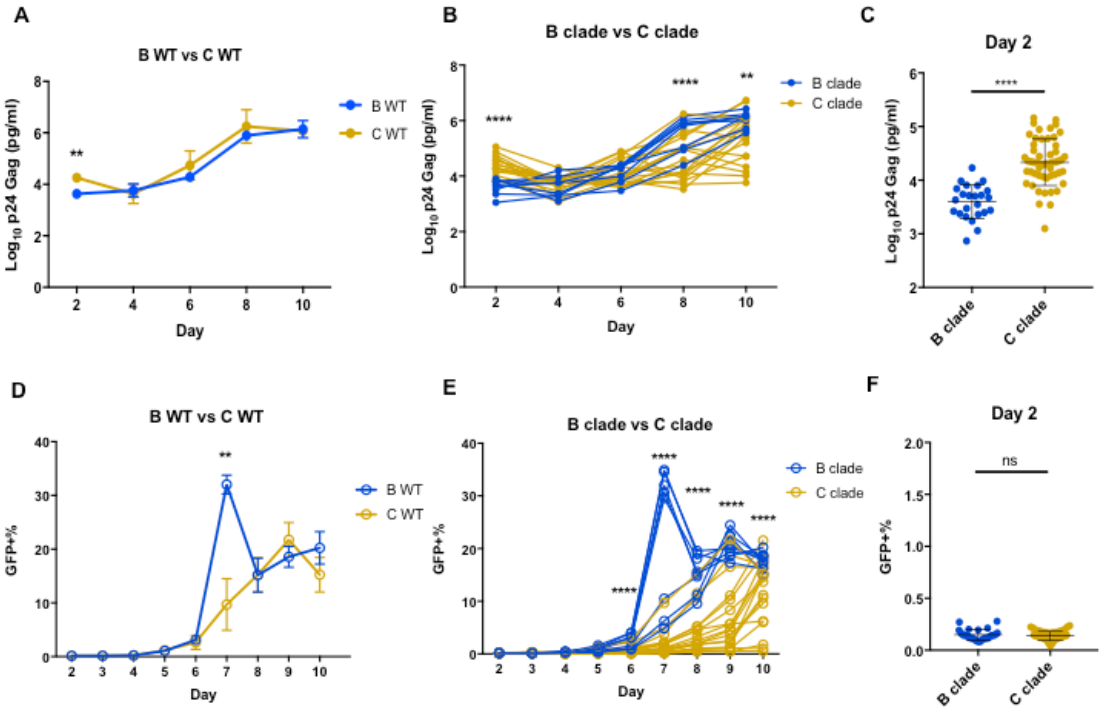
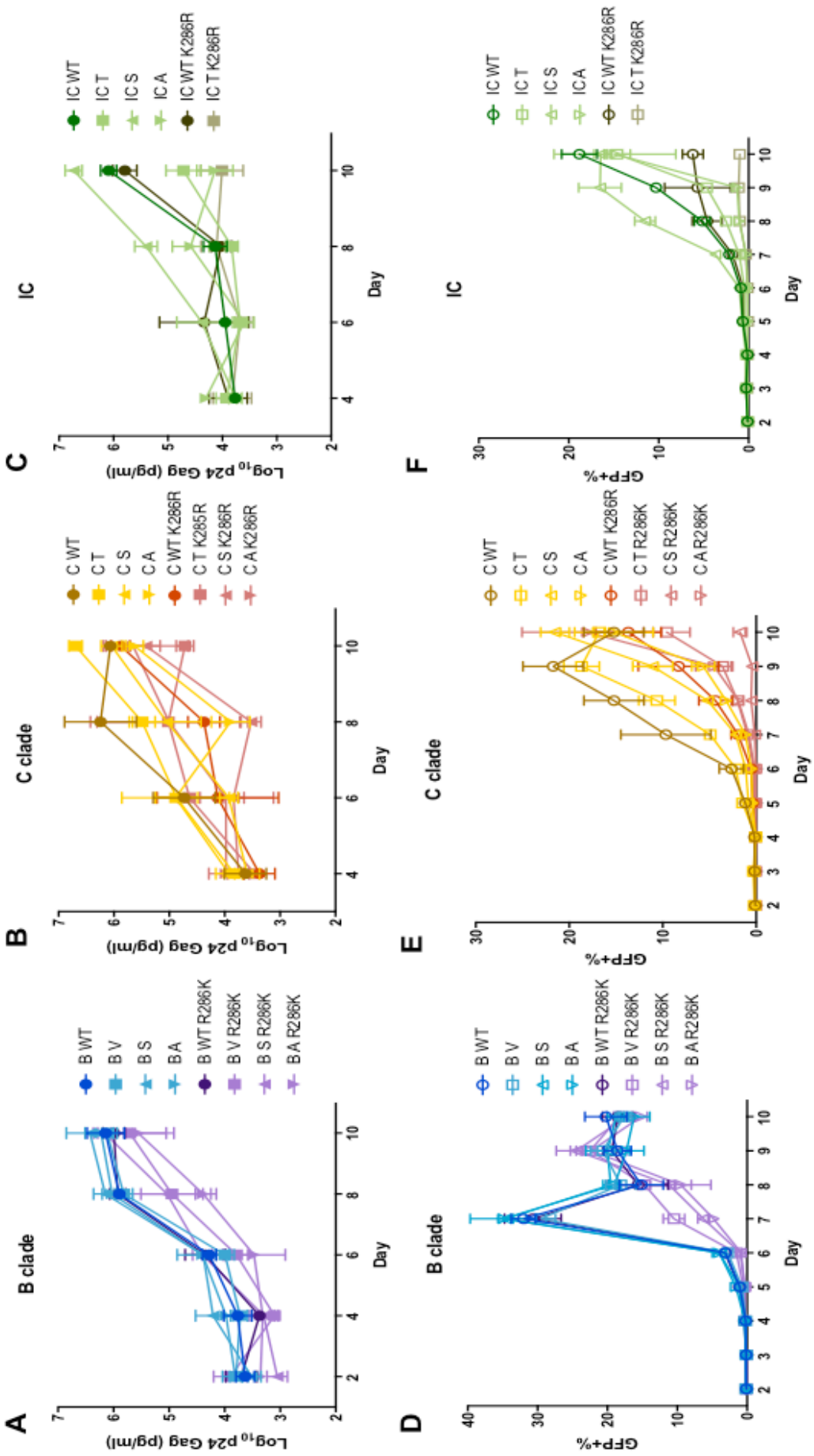


Figure 4.8. Comparison of fitness between B- and C-clade viruses by p24 Gag ELISA. (A-C) and fitness assay (E-G). Log p24 Gag concentration (pg/ml) of **(A)** B-clade wild-type NL4-3 and C-clade wild-type SK-254(M) from Day 2 to Day 10. **(B)** All B-clade and C-clade tested viruses including SK-254(M) and IC from Day 2 to Day 10. **(C)** All B-clade and C-clade tested viruses including SK-254(M) and IC on Day 2. GFP+ cells% of **(D)** B-clade wild-type NL4-3 and C-clade wild-type SK-254(M) from Day 2 to Day 10 **(E)** All B-clade and C-clade tested viruses including SK-254(M) and IC from Day 2 to Day 10. **(F)** All B-clade and C-clade tested viruses including SK-254(M) and IC on Day 2. One-way ANOVA multiple comparison and Dunnett's correction were used for calculating *p* value. (ns, ** $p \leq 0.01$, **** $p \leq 0.0001$)

Figure 4.9. (Next page) Comparison of viral fitness in different backbones by p24 Gag ELISA. (A-C) and fitness assay (E-G) from Day 2 to Day 10. Log p24 Gag concentration (pg/ml) of **(A)** B-clade viruses (NL4-3 backbone). **(B)** All C-clade viruses [SK-254(M) backbone]. **(C)** All IC viruses. GFP+ cells% of **(D)** B-clade viruses (NL4-3 backbone). **(E)** All C-clade viruses (SK-254(M) backbone). **(F)** All IC viruses.


A
B clade
Log10 p24 Gag (pg/ml)
Day
B WT
B V
B S
B A
B WT R286K
B V R286K
B S R286K
B A R286K
B
C clade
Log10 p24 Gag (pg/ml)
Day
C WT
C T
C S
C A
C WT K286R
C T K285R
C S K286R
C A K286R
C
IC
Log10 p24 Gag (pg/ml)
Day
IC WT
IC T
IC S
IC A
IC WT K286R
IC T K286R
D
B clade
GFP+%
Day
B WT
B V
B S
B A
B WT R286K
B V R286K
B S R286K
B A R286K
E
C clade
GFP+%
Day
C WT
C T
C S
C A
C WT K286R
C T R286K
C S R286K
C A R286K
F
IC
GFP+%
Day
IC WT
IC T
IC S
IC A
IC WT K286R
IC T K286R

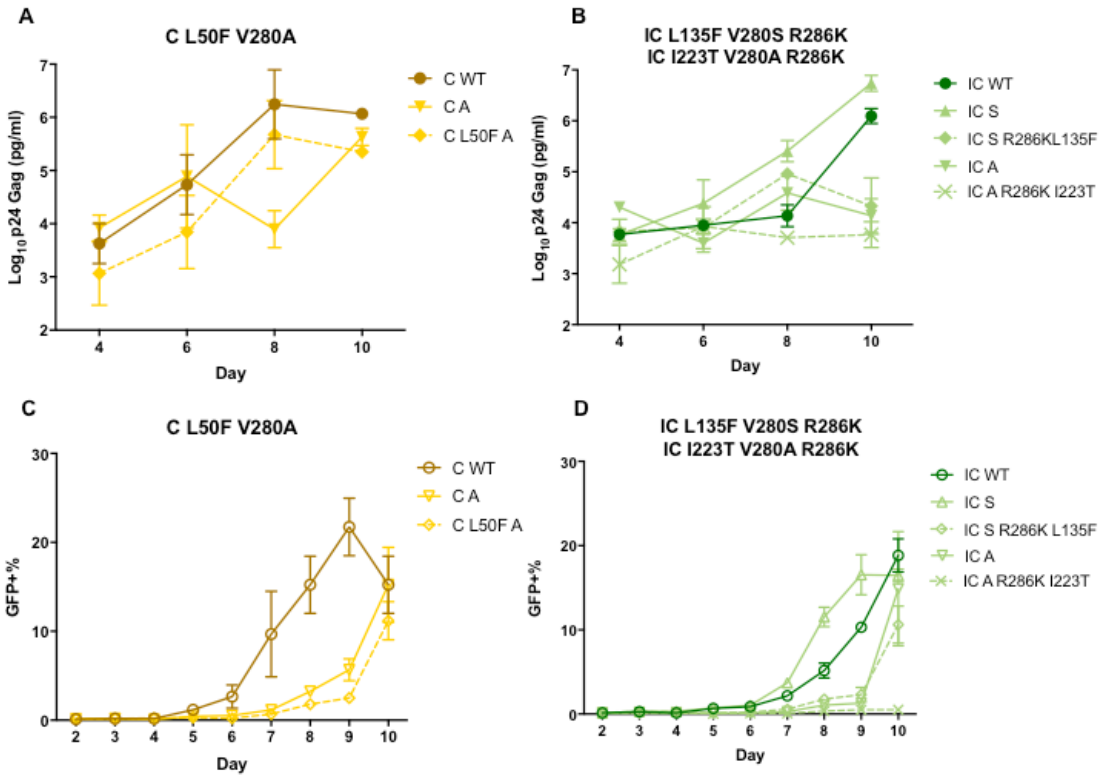


Figure 4.10. Comparison of viral fitness of *in vitro* developed mutations from Day 2 to Day 10. Log p24 Gag concentration (pg/ml) of **(A)** C-clade wild-type, V280A and L50F/V280A. **(B)** IC wild-type, V280S, L135F/V280S/K286R, V280A, I223T/V280A/K286R. GFP+% of **(C)** C-clade wild-type, V280A and L50F/V280A. **(D)** IC wild-type, V280S, L135F/V280S/K286R, V280A and I223T/V280A/K286R.

It should be noted that data collected after the growth peak were not used in the analysis in the GFP fitness assay (e.g. NL4-3 after Day 7 in **Figure 4.8 D**), since the VRC was calculated on the viral exponential phase. On the whole, apart from the first time point (Day 2), B-clade viruses were more competent than C-clade in p24 Gag ELISA (**Figure 4.8 B**). In fitness assay, the VRC of B-clade viruses remained significantly higher after Day 5 (**Figure 4.8 E**). Even when they reached peak growth and declined, the overall GFP+% was still higher than C-clade viruses.

The reason that the p24 Gag concentrations of C-clade viruses were significantly higher on Day 2 might be due to the amount of viruses added at the beginning of the experiment (**Figure 4.8 A** to **C**). One of the key points of conducting the viral fitness

assay was to ensure the same M.O.I applied to every virus at the starting point (**Figure 4.8 F**). In other words, the titration was based on determining the amount of viruses needed to reach the same percentage of GFP+ expression on Day 2, instead of quantifying actual concentrations of the viruses. Hence in our fitness assay, a greater volume of C-clade viruses including both SK-254(M) and IC was required to infect the same percentage of cells when compared with B-clade viruses. However, this does not necessarily mean that the actual titres of C-clade viruses were lower. Presumably, the Day 2 results shown in **Figure 4.8 A** to **C** indicates that the p24 Gag detected in the supernatant at this point might still be from the initial infection from Day 0. On Day 4, the p24 Gag concentrations were from the viruses that finished the life cycle in cells and released progeny into the supernatant. The entry of C-clade viruses into the GXR cells seems to be more difficult and less efficient (Ariën *et al*., 2007; Wilen *et al*., 2012). Therefore, when the p24 Gag ELISA results were compared with the GXR cell fitness assay, data of C-clade viruses on Day 2 were omitted in the following figures and discussion. Of note, even after the viruses infected the largest amount of cells and started to die off as GFP+% dropped (e.g. Day 8), the concentrations of p24 Gag in the culture supernatant was still maintained or continued to increase (**Figure 4.8 A, B, D** and **E**).

Figure 4.9 shows the p24 Gag concentrations by clades with the fitness assay percentages as a comparison in the lower panel. Unlike the previous findings by Murakoshi *et al*., 2015, all three V280X mutant viruses had similar level of p24 Gag production, shown in blue in **Figure 4.9 A**. With the presence of the K286R mutation (shown in purple), the amounts of p24 Gag were much less as the GFP+% peaked later (**Figure 4.9 D**). Similarly in C-clade viruses including both SK-254(M) and IC, the pattern of increase of p24 Gag was comparable to GFP+% (**Figure 4.9 B, C, E** and **F**). V280A and L50F/V280A are comparable, apart from the drop of V280A on Day 8 (**Figure 4.10 A**), the results of the two assays were fairly proportionate and

correspond to each other (**Figure 4.10**). Taken together, the number of extracellular viruses measured by p24 Gag ELISA had a similar trend as the percentage of infected GXR cells. Our findings of the fitness cost of R286K/K286R and the compensatory effect of L135F and I223T were confirmed in both ELISAs and GXR based assays. V280S/K286R and V280A/K286R in IC were not shown in the figures since they remained negative throughout both assays.

4.5.4 Discussion of p24 ELISA results

In our fitness assay, we found T280X in B-clade viruses did not significantly lowered VRC. Results from the p24 antigen ELISA also supported this finding. The main purpose of performing ELISAs as another approach to measure VRC was to confirm the viral fitness evaluated by a fitness assay where viral growth kinetics were calculated. Hence our ELISA results did not agree with Murakoshi's data. There are several potential explanations for these discrepancies. First, the viral titration for the fitness assay was a factor which influenced the amount of extracellular p24 Gag detected at the start. In addition, the use of a cell line in our assay, instead of human primary CD4+ T cells from HLA-B*52:01-positive healthy donors, might be one of the reasons for the different results. The reason that we used GXR cell line instead of primary CD4+ T cells is that we were able to control the other variables such as the phenotype and tropism of the CD4 co-receptor in the experimental settings. One feasible conclusion we can draw is the VRC of T280A measured by fitness assay was significantly lower in B-clade infection, which was not observed in the other T280X mutations.

CHAPTER 5: CLADE-SPECIFIC PATHWAYS OF COMPENSATORY MUTATIONS IN HLA-B*27:05-MEDIATED IMMUNE CONTROL OF HIV

5.1 Introduction

HLA-B*27:05 is one of the most well-characterised of the HLA alleles that are protective against HIV infection. Individuals carrying HLA-B*27:05 tend to maintain better immune control of HIV infection with delayed disease progression (Kaslow *et al*., 1996; Fellay *et al*., 2009; Pereyra *et al.,* 2010). The allele is also enriched in elite controllers (ECs) and long-term non-progressors (LTNPs) (Pereyra *et al.,* 2008).

The characteristics of HLA-B*27:05 binding groove are shown in Table 5.1. As is the case with the other members in B27 superfamily, Arginine (R) is strongly preferred at position 2. At the C-terminus, HLA-B*27:05 is capable of accommodating a wide range of hydrophobic amino acids apart from the largest Tryptophan (W), in addition to the basic residues Arginine, Lysine and Histidine (Sidney *et al.,* 2008; Marsh *et al*., 2000).

Table 5.1. Characteristics of B*27:05-restricted epitope.

	Peptide-Binding Motif				
Position	**P1**	**P2**	**P3**	**P5-7**	**PC**
Binding pocket	A	B	D	C	F
B*27:05	R/K	**R**	I/Y/F/W	I/L/P/V	R/K/Y/F/M/L/I/H

Table is based on The HLA Factsbook (Marsh *et al*., 2000).

*The dominant anchor Arginine(R) at position 2 is shown in bold.

HLA-B*27:05-mediated immunodominant CTL responses targeting the p24 Gag KK10 epitope (263-272, KRWIILGLNK) are associated with the immune control of HIV (Goulder *et al*., 1997; Goulder *et al*., 2001; Feeney *et al*., 2004, Chen *et al*.,

2012a). Escape from this epitope is associated with more rapid progression to AIDS (Goulder *et al*., 1997; Goulder *et al*., 2001), but at the same time, the mutation introduces a significant fitness cost for the virus (Schneidewind *et al*., 2007; Schneidewind *et al*., 2008). The most common escape mutation is R264K at position 2 (Goulder *et al*., 1997; Goulder *et al*., 2001; Ammaranond *et al*., 2005), and an upstream second mutation S173A is selected as a compensatory mutation to fully restore R264K replicative capacity (Schneidewind *et al*., 2007). The compensatory effect can be relatively explained in the secondary p24 Gag structure model where R264K (helix 7) is located at the same surface as S173A (helix 2) (Schneidewind *et al*., 2007). The long time to develop two simultaneous mutations partially explains how HLA-B*27:05-positive individuals initially maintain but eventually lose viral control over time (Goulder and Walker, 2012). Once R264K is accompanied by S173A, the virus is able to evade the CTL surveillance and still retain high viral replicative capacity (VRC). Studies have shown the two mutations seldom reverse back to wild-type sequence even in a transmitted host who does not carry HLA-B*27:05 (Goulder *et al*., 2001; Scheidewind *et al*., 2008).

To date, most research has focused on B-clade infected Caucasians, due to the relatively frequency of HLA-B*27:05 in this population (8-10%, compared to <1% in sub-Saharan African populations) (The Allele Frequency Net Database, www.allelefrequencies.net). R264K and its compensatory mutation S173A were studied and described in this B-clade context. When examining the impact of HLA alleles on immune control of HIV, however, it is important to take the other HIV subtypes into account, not only for considering geographical and epidemical aspects, but also for different VRC and evolutionally rates of the virus associated with different disease outcomes (**Chapter 4**; Kløverpris *et al*., 2014; Brener *et al*., 2015). In addition, different HIV subtype sequences might alter CD8+ T-cell immunodominance hierarchies. An example here is the HLA-B*35:01-restricted Gag

epitope, NPPIPVGDIY (Gag 254-262), which is the immunodominant epitope and associated with some degree of control of viraemia in C clade infection. In the majority of C-clade individuals with HLA-B*35:01, the D260E escape mutant which does not bind HLA-B*35:01 is selected early. In B clade infection, HLA-B*35:01 is associated with more rapid progression and the consensus sequence for the epitope is NPPIPVGEIY, which is non-immunogenic (Matthews *et al.*, 2012). In the case of HLA-B*27:05, the residue at Gag 173 is Threonine in C-clade consensus instead of Serine. We hypothesized that, for this reason, it is possible that the escape and compensatory pathways are different for KK10 between B- and C-clade infection.

We therefore studied a group of HLA-B*27:05-positive, C-clade infected individuals from several large cohorts, detailed in **Section 2.1**. A transmission pair was also included for longitudinal study over 86 months. By identifying HLA-B*27:05 footprints in the sequences, we were able to test the impact of different variants on VRC and their roles in KK10-mediated CTL responses. This offers insights into the potentially different patterns of HLA-B*27:05-restricted escape and particularly compensatory mutations, in particular the latter, which will then give a clearer picture of the evolving relationship between HLA and HIV.

5.2 Chapter-specific Materials and Methods

5.2.1 Study subjects

In total, 26 HLA-B*27:05-positive, chronically C-clade infected, treatment-naïve individuals were selected from different cohorts [North Indian, Sinikithemba, Bloemfontein, ZEHRP, Gaborone and Thames Valley Cohort (TVC)]. Fifteen HLA-B*27:05-negative subjects infected with the autologous virus carrying HLA-B*27:05 footprints were also included in the study. A transmission pair of HLA-B*27:05-negative donor and -positive recipient from the TVC was studied as well.

5.2.2 Site-directed mutagenesis of S165N

The fitness assay performed in this study was based on p6.5 (C-clade) and p83-2 (B-clade) combined with p83-10_{GFP}, this is a different approach from that described in the previous chapters. The chimeric viruses generated in this system carried GFP expression themselves, so the GFP percentages we measured were the replicative viruses in MT4 cell lines. This is in contrast to previous ΔGag-protease system where encoding GFP was in a Tat-dependent manner in GXR cells.

All the mutations were introduced by using the QuikChange Lightning site-directed mutagenesis kit (Agilent technologies). **Table 5.2** lists all the tested mutant viruses, the majority of which were kindly provided by Dr Jacqueline R Brener. The S165N mutation was introduced into different templates including p6.5_C_p24_WT, P6.5_C_p24 R264K/T173A, p83-2, p83-2 R264K and p83-2 R264K/S173A. Custom-designed mutagenesis forward and reverse primers of S165N are shown in **Table 5.3**.

Table 5.2. Tested HLA-B*27:05-associated mutant viruses.

	C-clade p24 Gag backbone	B-clade p24 Gag backbone
wild-type	p6.5_C_p24_WT	p83-2
R264K variants	p6.5_C_p24 R264K p6.5_C_p24 R264K/T173A p6.5_C_p24 R264K/T173S p6.5_C_p24 R264K/S165N p6.5_C_p24 T173S p6.5_C_p24 T173A **p6.5_C_p24 S165N** **p6.5_C_p24 R264K/T173A/S165N** p6.5_C_p24 R264K/V218M p6.5_C_p24 V218M p6.5_C_p24 R264K/V218M/S165N	p83-2 R264K p83-2 R264K/S173A p83-2 R264K/S173T **p83-2 R264K/S165N** **p83-2 S165N** **p83-2 R264K/S173A/S165N**
R264G variants	p6.5_C_p24 R264G p6.5_C_p24 R264G/D260E p6.5_C_p24 R264G/D260N	p83-2 R264G p83-2 R264G/E260D

Newly made S165N introduced viruses are shown in bold. The other viruses were provided by Dr Jacqueline R Brener.

Table 5.3. S165N Custom-designed mutagenesis forward and reverse primers.

HXB2 position in Gag	Clade	Mutant	Primer direction	Primer sequence (5'→3')
165	C	S→N	Forward	TA AAA GTA ATA GAG GAG AAG GCT TTT **AAC** CCG GAG GTA ATA C
			Reverse	G TAT TAC CTC CGG **GTT** AAA AGC CTT CTC CTC TAT TAC TTT TA
165	B	S→N	Forward	TA AAA GTA GTA GAA GAG AAG GCT TTC **AAC** CCA GAA GTA ATA CC
			Reverse	GG TAT TAC TTC TGG **GTT** GAA AGC CTT CTC TTC TAC TAC TTT TA

The work I have contributed to this study: *Of the 22 KK10 variant viruses that were made for this study, I contributed 5. This study was initiated by Dr Brener and was completed by me. The initial findings from Dr Brener's study included the observation that S165N is a compensatory mutation to R264K in C-clade infection. However co-variation analyses undertaken by Dr Jonathan Carlson raised two further questions that I sought to address here. The co-variation analyses showed that S165N is not only a previously undescribed compensatory mutation to R264K in B-clade as well as in C-clade infection, but in fact S165N covaries more strongly with R264K in B-clade than in C-clade infection. To experimentally evaluate the impact of S165N on viral replicative capacity in B clade virus with and without R264K, I therefore generated these variant viruses and undertook the fitness assays shown below. The second question I sought to address here is as follows: The major difference between B- and C- clade viruses with respect to R264K compensation lies in the use of S173A or, more rarely, S173T in B-clade infection. In C-clade, there is no co-variation of the consensus T173 with R264K. In B-clade infection, S165N is observed only in a subset of cases where R264K is present in combination with S173A, never alone in combination with R264K, raising the possibility that, in B-clade viruses, S165N adds to the compensation resulting from S173A. I therefore addressed this question by generating the viruses listed in* ***Table 5.2*** *and in undertaking the fitness assays shown in* ***Figures 5.2-5.4*** *and* ***Tables 5.7****. The work*

described below in pages 146-150 was undertaken prior to my involvement in the study.

5.3 Results

*5.3.1 Distinct pattern of escape and putative compensatory mutation(s) in the HLA-B*27:05-positive recipient*

The transmission pair information and deep sequencing data in the following section were provided by Dr Jacqueline R Brener, Department of Paediatrics, University of Oxford.

Commencing in April 2007, a transmission pair infected with C-clade virus was studied over 86 months. An HLA-B*27:05-negative donor South-African female (HLA-B*14:01/B*58:01) transmitted the virus to an HLA-B*27:05-positive UK Caucasian male recipient (HLA-B*27:05/B*51:01) between November, 2007 (HIV test negative) and September 2008 (diagnosis) (**Figure 5.1 A** and **B**). Phylogenetic analysis and subtyping of the full-length genome confirmed the subtype of virus as C clade infection and the authenticity of the transmission pair (**Figure 5.1 C**). A time-measured phylogenetic analysis showed the estimated time of transmission was approximately in October 2007 (**Figure 5.1 D**).

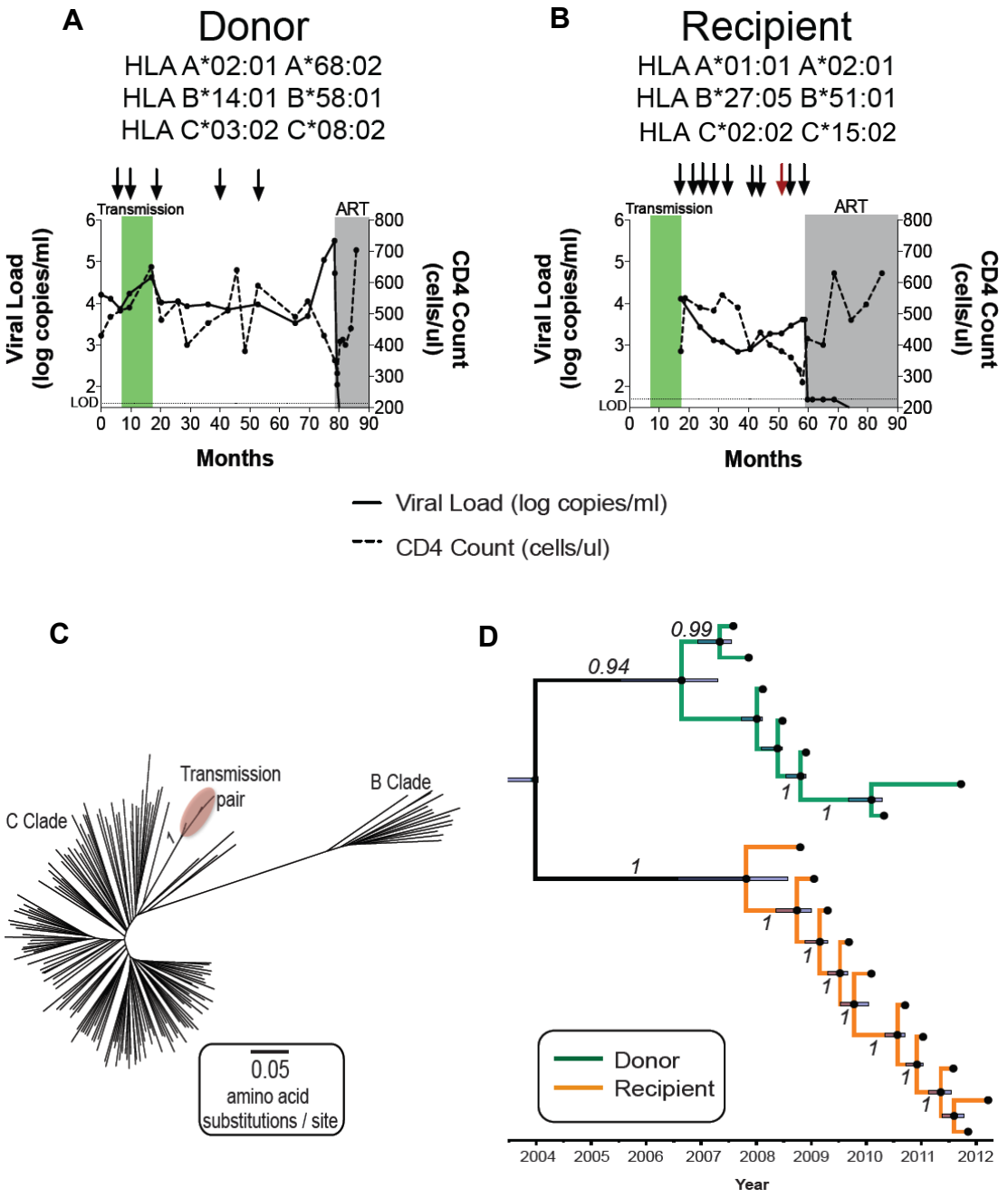


Figure 5.1. Disease progression of C-clade infected transmission pair and the phylogenetic analysis of the autologous virus. Viral load (solid line) and CD4 count (dashed line) of **(A)** HLA-B*27:05-negative female donor and **(B)** HLA-B*27:05-positive male recipient. Timepoint '0' represents the time of diagnosis of the recipient. LOD stands for the limit of detection (40 copies/ml), shown by a horizontal dotted line. Green and grey shaded regions represents the window period and time of the subject receiving antiretroviral therapy (ART). Arrows are the timepoints of sampling for ultra-deep sequencing. Red arrow indicates the first detection of the R264K escape mutation. **(C)** Maximum likelihood tree of 8,739 bp sequence alignment (full-length HIV genome) of the transmission pair (highlighted in red, *n*=7 and 10 from the donor and recipient, respectively); C-clade (*n*=150, South Africa); and B-

clade (*n*=15, US) reference sequences between 2001 and 2011 from the Los Alamos database (LANL HIV Sequence Database, http://www.hiv.lanl.gov/). Bootstrap value for the transmission pair is shown in italics. **(D)** Bayesian time-measured phylogenetic analysis of full-length HIV sequences from multiple timepoints from the donor (green) and recipient (orange). The most recent common ancestor (MRCA) of the recipient was the timing of transmission, in approximately October, 2007 within the 95% height posterior density (HPD) range (blue bars). Posterior probability values are shown in italics.

Longitudinal ultra-deep sequencing of the HLA-B*27:05-positive recipient was focused on the p24 Gag protein, in particular the KK10 epitope (**Table 5.4**). As a result, some distinctive putative compensatory mutations were observed in the C-clade infection. The common escape mutation R264K (Goulder *et al*., 1997; Goulder *et al*., 2001) had gradually overtaken the wild-type from the 44th month, reaching fixation at the 51st month. By the time of the emergence of R264K, the patient had shown clinical deterioration including declining of CD4 count and increasing viral load (**Figure 5.1 B**). Surprisingly, there was no residue change at Gag 173 throughout the follow-up. In B-clade, the S to A mutation serves as the compensatory mutation to R264K (Schneidewind *et al*., 2007). Instead, S165N and V218M were both co-selected with R264K in C-clade infection, almost at the same timepoint with similar percentages (**Table 5.4**). This is highly indicative of the potential compensatory roles S165N and V218M might play in for R264K. Another common escape mutation L268M was observed in earlier timepoints, consistent with previous findings in B-clade (Goulder *et al*., 1997; Kelleher *et al*., 2001). This escape mutation has minimal impact on VRC, but it can still be recognised by some CTL subsets (Lichterfeld *et al*., 2007; Schneidewind *et al*., 2007). Along with other variants such as I267V, they accounted for small percentages at the late timepoints, which demonstrates the different attempts the virus employed and the evolution under CTL pressure.

Table 5.4. Deep sequencing of the HLA-B*27:05-restricted Gag-KK10 epitope and co-varying amino acid positions in a C-clade infected transmission pair.

	C-clade Consensus and Autologous Variants	Frequency (%)										
		Donor	Recipient									
	(month)	19	19	21	24	28	33	41	44	51	54	59
HLA-B*27:05 Gag-KK10 variants	^{263}KRWIILGLNK^{272}	100	72	98	39	43	54	10	32	-	-	-
	-K--------	-	6	-	-	-	-	-	30	95	100	100
	-K---M----	-	-	-	-	-	-	-	-	5	-	-
	-----M----	-	13	2	61	48	25	5	3	-	-	-
	----VM----	-	6	-	-	-	-	67	15	-	-	-
	-----M--H-	-	-	-	-	-	3	14	8	-	-	-
	-----M--T-	-	-	-	-	10	19	2	-	-	-	-
	----V-----	-	3	-	-	-	-	-	-	-	-	-
	--------H-	-	-	-	-	-	-	-	5	-	-	-
	--------T-	-	-	-	-	-	-	2	-	-	-	-
	-----I--T-	-	-	-	-	-	-	-	8	-	-	-
Variants co-selected with R264K	T173X	-	-	-	-	-	-	-	-	-	-	-
	S165N	-	6	-	-	-	-	-	31	100	100	77
	V218M	-	5	-	-	-	-	-	-	81	88	96

Longitudinal Gag-KK10 epitope and R264K co-selected polymorphisms are shown. Percentages of variants in KK10 have been adjusted by excluding the variants <1.5% frequency. Depth of coverage ranges from 174-192,000 reads. The C-clade consensus sequence is shown on the top. R264K escape variants are shown in red.

5.3.2 C-Clade-specific pattern of Gag KK10 compensation

The sequencing data in the following section was provided by Dr Jacqueline R Brener, Department of Paediatrics, University of Oxford.

The Gag KK10-associated variants were investigated in other C-clade infected HLA-B*27:05-positive individuals (*n*=26), along with HLA-B*27-negative subjects (*n*=15) (**Table 5.5**). Presumably, viruses carrying KK10-associated variants were transmitted from HLA-B*27-positive donors to these HLA-B*27-negative individuals. K264R/L268M were the most common selected escape mutations within KK10 epitope in HLA-B*27:05-positive individuals, similar to B-clade infection. Both S165N and T173S only appeared with the existence of R264X in both HLA-B*27:05-positive and -negative groups (**Table 5.5** and **Figure 5.2**). Considering S165N, the co-

selection of S165N with R264K reached significance despite the small sample number (**Figure 5.2 A**; Fisher's exact test, *p*=0.0108). It is likely that S165N acted as a compensatory mutation to R264K. In addition, it is noted that none of the sequences contained T173A in C-clade infection, which is the main compensatory mutation for R264K in B-clade infection.

Table 5.5. Sequences of Gag KK10 epitope and Gag 165/173 of proviral DNA from C-clade infected subjects.

	165	173	263-272	
Consensus C	**S**	**T**	**KRWIILGLNK**	
Consensus B	**-**	**S**	**----------**	**No. of Patients**
B*27:05+ve with KK10 R264X Escape (*n*=15)	-	-	-K---M----	*n*=6*
	-	S	-K---I-X--	*n*=1
	N	-	-K---M----	*n*=1
	N	-	-K---I--H-	*n*=1
	N	S	-K---M--H-	*n*=1
	N	-	-K--------	*n*=1*
	-	I	-K---M----	*n*=1
	-	-	-G--------	*n*=2
	-	S	-G---M----	*n*=1
B*27:05+ve without KK10 R264X Escape (*n*=11)	-	-	----------	*n*=7
	-	-	-----M----	*n*=1
	-	-	--------H-	*n*=1
	-	-	---V------	*n*=1
	-	-	-------F--	*n*=1
B*27-ve with R264X Variants (*n*=15)	-	-	-K--------	*n*=5
	N	-	-K---M----	*n*=3
	N	-	-K--------	*n*=1
	-	-	-K---I----	*n*=1
	N	I	-K--V-----	*n*=1
	-	-	-G--------	*n*=2
	-	-	-G---M----	*n*=1
	-	-	-S--------	*n*=1

Gag 173 and 165 are shown due to its association with R264K in B-clade and potential for compensatory mutation in C-clade (*n*=41). *=one sample sequenced from plasma RNA as opposed to proviral DNA.

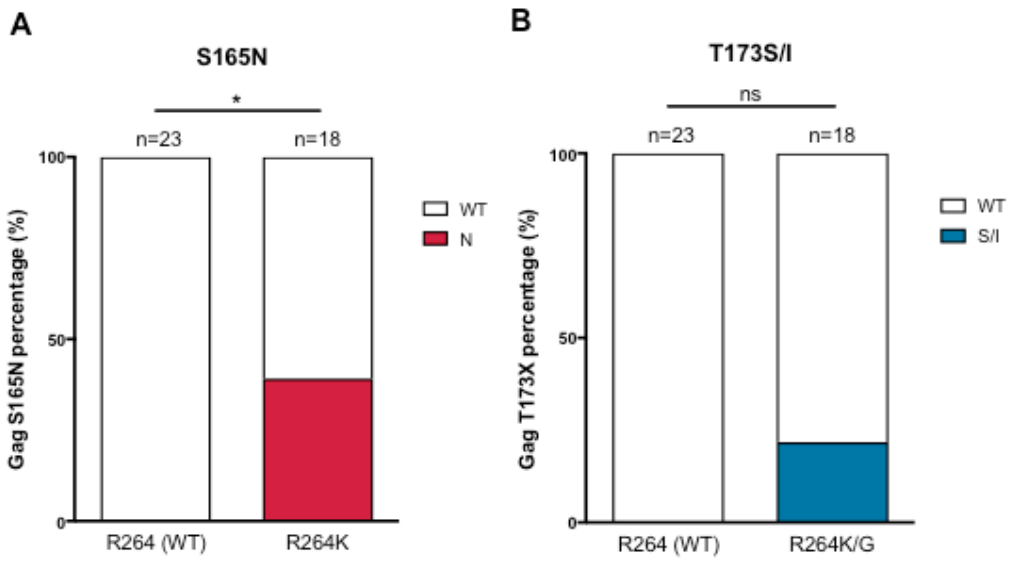


Figure 5.2. Frequencies of putative compensatory mutations for R264K/R264G in C-clade infection. (A) S165N percentage **(B)** T173S/I percentage in both HLA-B*:27:05-positive and -negative subjects with or without the R264K/G mutation, based on Table 5.5. 2x2 Fisher's exact Test was performed. (ns, p≥0.05; * p<0.05)

5.3.3 Undefined role of S165N in relation to Gag 173 in both B- and C-clade infection

The analyses in the following section were performed by Dr Jacqueline R Brener, Department of Paediatrics, University of Oxford; and by Dr Jonathan M Carlson, Microsoft Research, USA.

To understand the different patterns of HLA-B*27:05-associated mutations in different HIV subtypes, statistical analysis was performed to compare differential selection in B- versus C-clade. **Table 5.6** shows the different restrictions in wild-type sequences, sequences with R264K, or sequences with R264G by B- and C-clades. In total, 12 out of 19 co-variation sites with R264X were significantly different between the two clades. As with the R264K compensatory patterns, S165N and T173A were of particular interest (shaded in light blue in **Table 5.6**). Unexpectedly, S165N was associated with R264K in both B- and C-clade infection, instead of C-clade only. In fact, it was even more significant in B-clade (p≤0.0001) than C-clade

($p \leq 0.001$). It is noteworthy that S165N also acts an escape mutation of the Gag KF11 epitope (162-172, KAFSPEVIPMF) restricted by HLA-B*57:03 in C-clade infection (Carlson *et al*., 2012b). Since HLA-B*57:03 expression is relatively high (~7%) in sub-Saharan African populations (The Allele Frequency Net Database, www.allelefrequencies.net), it gives a higher background of S165N in individuals not carrying R264K in C-clade infection. The presence of S165N in African populations therefore obscures the fact that S165N not only is a compensatory mutant for the HLA-B*57:03 KF11 mutant A163G (Crawford JV 2007) but also for the HLA-B*27:05-associated KK10 mutation R264K In fact, the odds ratio (OR) of S165N arising with R264K is substantially higher in B-clade than in C clade infection (27.7 versus 6.5). In other words, the selection of S165N in combination with R264K is more strongly selected in B-clade than in C-clade infection. Moreover, S165N never occurred in the absence of S173A in B-clade infection, which indicates S165N was not an alternative compensatory mutation, but in addition to S173A (q=2e-5). Intriguingly, despite the presence of T173S in C-clade infection, it did not reach significance in the analysis; nor was T173A selected. By contrast, in B-clade infection, our analysis displayed the consistent results of co-selection of S173A with R264K. Collectively, S165N was linked to R264K escape mutation in both clades but it was in addition to S173A in B-clade. T173A was never observed in C-clade infection and the role of Gag 173 mediating the compensatory effect with S165N for R264K remains in question.

Another potential compensatory mutation candidate V218M for R264K shown in the transmission pair (**Table 5.4**) did not yield statistical significance among other individuals in our analysis.

Table 5.6. Co-variation analysis of Gag-R264X in C- and B-clade HIV infection.

	HXB2 position	C-clade Consensus	B-clade Consensus	Associated/ Co-varying residue	B-clade *p* value	C-clade *p* value	Comparison of B- and C-clade		Differential Selection in B- vs C-clade
							p value	*q* value	
R264 (WT)	**163**	**A**	**A**	**A**	**ns**	**≤0.05**	**≤0.0001**	**≤0.0001**	**Y**
	215	L	V	M	≤0.001	ns	ns	ns	N
	256	**I**	**I**	**V**	**≤0.01**	**ns**	**≤0.0001**	**≤0.0001**	**Y**
	256	**I**	**I**	**I**	**≤0.01**	**ns**	**≤0.0001**	**≤0.0001**	**Y**
	268	**L**	**L**	**I**	**≤0.0001**	**≤0.0001**	**≤0.001**	**≤0.001**	**Y**
R264K Escape	9	R	S	R	ns	≤0.05	ns	ns	N
	26	K	K	K	≤0.01	ns	ns	ns	N
	165	**S**	**S**	**N**	**≤0.0001**	**≤0.001**	**≤0.05**	**≤0.05**	**Y**
	165	**S**	**S**	**S**	**≤0.0001**	**≤0.001**	**≤0.05**	**≤0.05**	**Y**
	173	**T**	**S**	**A**	**≤0.0001**	**ns**	**≤0.0001**	**≤0.0001**	**Y**
	173	**T**	**S**	**S**	**≤0.0001**	**ns**	**≤0.01**	**≤0.01**	**Y**
	173	T	S	T	≤0.05	≤0.01	ns	≤0.05	N
	182	**Q**	**Q**	**H**	**ns**	**≤0.0001**	**≤0.01**	**≤0.01**	**Y**
	200	**M**	**M**	**I**	**ns**	**≤0.01**	**≤0.05**	**≤0.05**	**Y**
	268	L	L	L	≤0.0001	≤0.0001	ns	≤0.05	N
	268	L	L	M	≤0.0001	≤0.0001	ns	ns	N
	268	L	L	I	≤0.0001	≤0.0001	ns	ns	N
R264G Escape	**136**	**Q**	**Q**	**R**	**≤0.0001**	**ns**	**≤0.001**	**≤0.001**	**Y**
	136	**Q**	**Q**	**Q**	**≤0.0001**	**ns**	**≤0.001**	**≤0.001**	**Y**

Gag 165 and 173 are shaded in light blue. Variants that reached significance between B- and C-clade are in bold.

5.3.4 Different compensatory effects of S165N and S173A in B- and C-clade infection

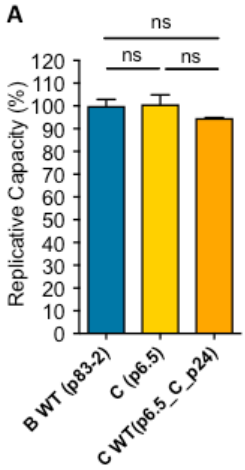
To determine the putative compensatory effect of S165N and S173A/T173S, viral fitness assays were used to evaluate the impact of S165N made in B-clade infection and to offer possible explanations for the absence of T173A in C-clade infection. First of all, the VRCs of both B- and C-clade wild-types viruses were compared. Here we used p6.5_C_24 as C-clade wild-type, which was reverse mutated from p6.5 to make the p24 Gag sequence identical to C-clade consensus. Similar to the SK-254 and SK-254(M) plasmids used in the previous Chapters (Chapter 3 and 4). **Table 5.7** displays the p6.5_C_24 p24 Gag sequence with p83-2 (NL4-3) as C-clade and B-clade wild-type, used in the assay and the following analyses. **Figure 5.3 A** shows there was no difference in VRC between B-clade WT, p6.5 and p6.5_C_24 (C-clade wild-type), which was different to our previous fitness studies where C-clade viruses were less fit (Chapter 3 and 4).

Table 5.7. Alignment of p24 Gag sequence of p83-2 (B-clade wild-type, NL4-3) and p6.5_C_24 (C-clade wild-type, with p24 Gag identical to C-clade consensus).

Plasmid	HXB2 Gag position
	159
p83-2	P I V Q N L Q G Q M V H Q A I S P R T L N A W V K V V E E K A F S P E V I P M F
p6.5_C_24	- I - - - - - - - - - - - - -
	173 203
p83-2	S A L S E G A T P Q D L N T M L N T V G G H Q A A M Q M L K E T I N E E A A E W
p6.5_C_24	T - D - - - - - - - - -
	248 252
p83-2	D R L H P V H A G P I A P G Q M R E P R G S D I A G T T S T L Q E Q I G W M T H
p6.5_C_24	- A - - - S
	260 280 286
p83-2	N P P I P V G E I Y K R W I I L G L N K I V R M Y S P T S I L D I R Q G P K E P
p6.5_C_24	- - - - - - - D - V - - - - - K - - - - - -
	301 310 312 319
p83-2	F R D Y V D R F Y K T L R A E Q A S Q E V K N W M T E T L L V Q N A N P D C K T
p6.5_C_24	- - - - - - - - F - - - - - - - - - T - D - - - - - - - D - - - - - - - - - - - - -
	335 357
p83-2	I L K A L G P G A T L E E M M T A C Q G V G G P G H K A R V L
p6.5_C_24	- - R - S - - - - - -

We then examined the compensatory effect of S173T/A and S165N in B-clade viruses (**Figure 5.3 B**). As reported in previous studies (Goulder *et al*., 1997; Goulder *et al*., 2001; Ammaranond *et al*., 2005), there was a considerable fitness cost of R264K, reducing VRC by more than 50% (p<0.0001). Changing from a B- to C-clade residue at Gag 173 (S173T) was unable to rescue the VRC of R264K, whereas S173A successfully restored VRC to the wild-type level, again consistent with previous findings (Schneidewind *et al*., 2007). S165N alone did not make a great impact on reducing viral fitness (~5%). The compensatory effect of S165N was second to S173A, partially restoring R264K to 87.4%. Interestingly, the presence of S173A and S165N together with R264K did not have an impact on the compensatory effect of S173A alone with R264K, but the S173A raised the VRC of R264K/S165N. The emergence of S165N in addition to S173A clearly did not yield a great advantage for the virus, in terms of VRC. It still remains unclear why S165N emerged when the S173A alone was able to compensate R264K to almost wild-type level.

In C-clade viruses, R264K reduced the VRC significantly, yet not as much as in B-clade viruses (**Figure 5.3 C**). Unlike in B-clade, T173S and T173A did not act as compensatory mutations. In fact, they both further decreased VRC by approximately 10%. T173S and T173A alone lowered the VRC even more than R264K. It is therefore reasonable that T173X variants were rarely seen in C-clade sequences, in particular T173A, which was able to reduce the VRC by 50%. On the other hand, S165N alone had no impact on VRC in C-clade viruses, similar to B-clade with a slight reduction (**Figure 5.3 B** and **5.3 C**). It seems that S165N was a critical compensatory mutation to R264K in C-clade infection. Although it was not as significant as B-clade, S165N was able to increase the VRC of R264K. However, this compensatory effect was lost with the addition of T173A. The combination of T173A, S165N and R264K had lower VRC, compared to R264K alone.



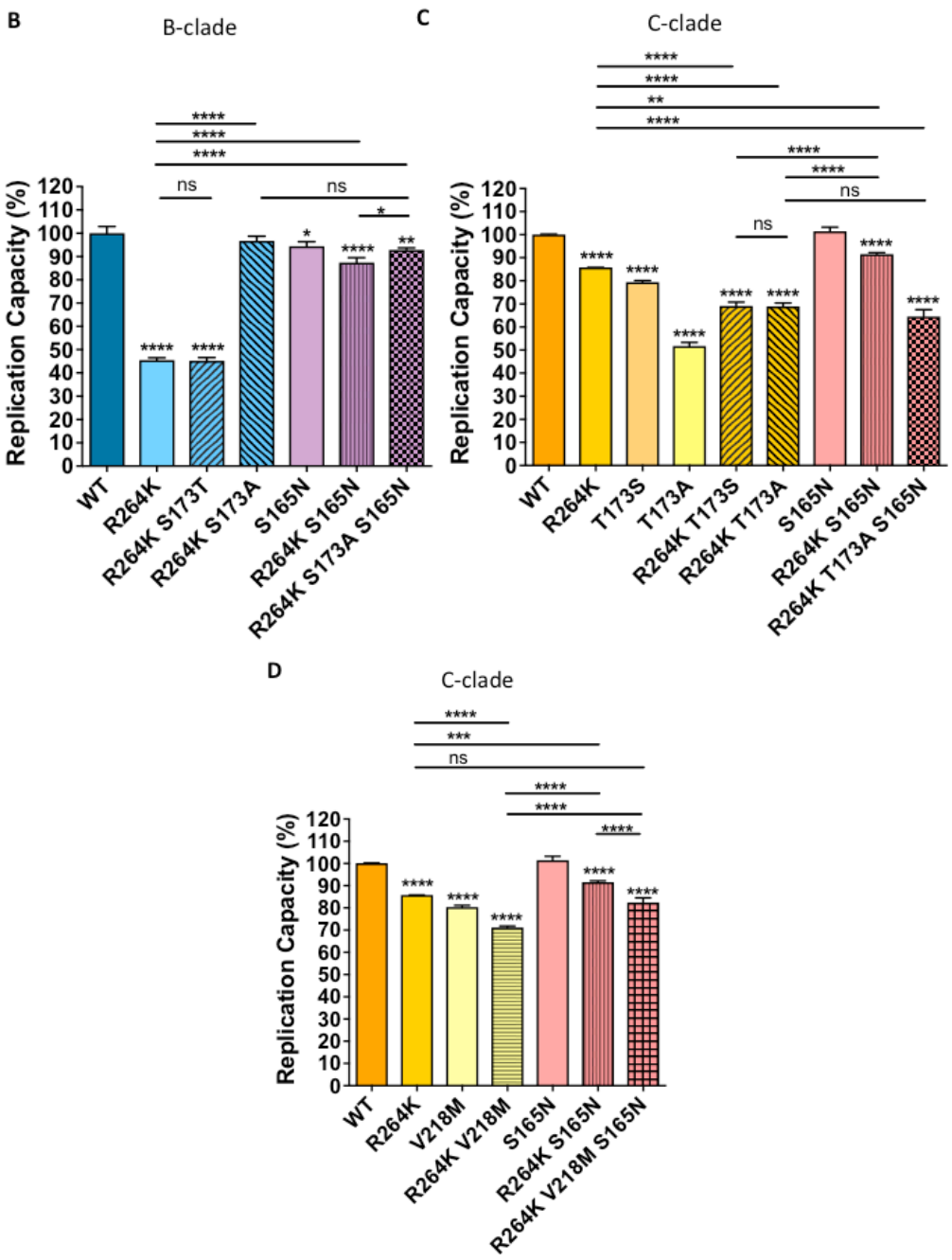


Figure 5.3. Viral replicative capacity (VRC) of R264K-related mutant viruses in B- and C-clade. (A) Relative VRC (%) of C-clade p6.5 and C-clade p6.5_C_p24 as wild-type to B-clade wild-type (NL4-3). **(B)** Relative VRC (%) of R264K/S173X/S165N B-clade viruses to B-clade wild-type (NL4-3). **(C)** Relative VRC (%) of R264K/T173X/S165N C-clade viruses to C-clade wild-type (p6.5_C_p24). **(D)** Relative VRC (%) of R264K/V218M/S165N C-clade viruses to C-clade wild-type (p6.5_C_p24). All bars are shown as mean with SD. Patterned bars represent putative compensatory mutation(s) to R264K. One-way ANOVA multiple comparison and Dunnett's correction were used for calculating *p* value. Asterisks marked on top of each bar are the *p* value when each mutant virus was compared to the wild-type. (ns, p≥0.05; * *p*<0.05, ** *p*<0.01, *** *p*<0.001, **** *p*<0.0001)

Moreover, since V218M was co-selected with S165N and R264K in the C-clade infected HLA-B*27:05-positive recipient in the transmission pair (**Table 5.4**), we examined the VRC of this mutation along with S165N (**Figure 5.3 D**). Similar to T173S, V218M did not serve as a compensatory mutation. V218M alone or in the presence of R264K had lower VRC, compared to the VRC of R264K alone. In the combination of S165N, V218M and R264K, S165N lost its compensatory effect. From the longitudinal study (**Table 5.4**), it is clear that S165N was co-selected almost at the same time with R264K, and V218M followed shortly after. It is very similar to the case in B-clade where the virus took the strategy of accommodating the third mutation with a slight reduction of the VRC (**Figure 5.3 B**), the reasons for which still remain unclear, but it is noted that the three mutation combinations (S165N/S173A/R264K in B-clade and S165N/V218M/R264K in C-clade) did not completely cripple HIV. The VRCs of the two combinations still maintained 92.9% and 82.4% respectively, which suggests S173A and V218M might be functional in other underlying mechanisms to pressure on the virus to develop the third mutation with a cost of marginal decreased VRC.

5.3.5 E260D and D260E/N did not act as compensatory mutations to R264G in B- and C-clade infection, respectively

In some HLA-B*27:05-positive individuals infected with C-clade HIV, the co-selection of R264G and D260N was observed (**Table 5.5**; *n*=3). Due to the different residues in different subtype consensuses sequences (E in B-clade and D in C-clade), this position potentially could act as a compensatory mutation to rescue R264G escape. In our assay, R264G did not have a significant impact on VRC, compared to wild-type in B-clade viruses (**Figure 5.4 A**). Neither did E260D significantly increase or decrease VRC of R264G. This is different from a previous study, which showed E260D was able to partially compensate for R264G selection in B-clade infection

(Schneidewind *et al*., 2008). In C-clade viruses, R264G significantly decreased the VRC more than R264K (**Figure 5.4 B**). Neither D260E nor D260N rescued the VRC of R264G. It is likely that D260N was an irrelevant variant to R264G and KK10-associated fitness cost. Collectively, position 260 did not yield a clade-specific effect, which is consistent with the analysis in Table 5.6 with the absence of Gag 260. Neither E260D nor D260E/N acted as compensatory mutation in either B- or C-clade.

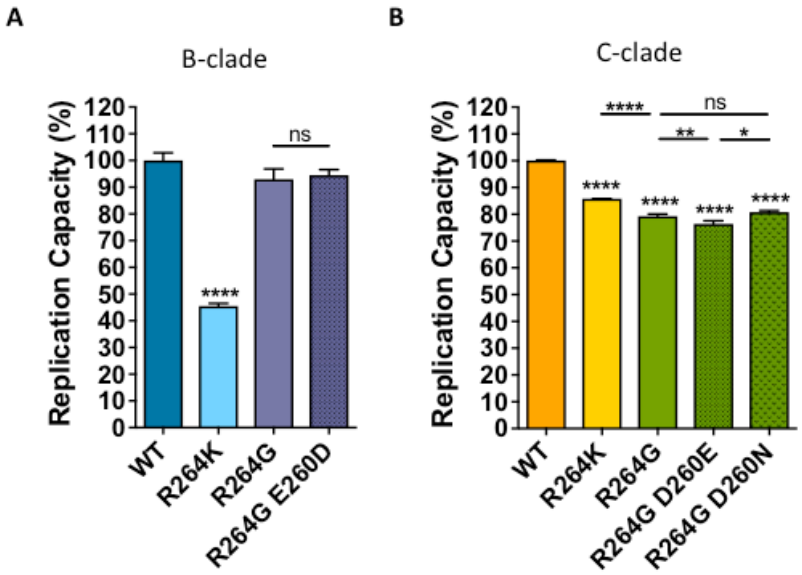


Figure 5.4. Viral replicative capacity (VRC) of R264G-related mutant viruses in B- and C-clade. (A) Relative VRC (%) of R264G/E260D B-clade viruses to B-clade wild-type (NL4-3). **(B)** Relative VRC (%) of R264G/D260X C-clade viruses to C-clade wild-type (p6.5_C_p24). All bars are shown as mean with SD. Patterned bars represent putative compensatory mutation(s) to R264K. One-way ANOVA multiple comparison and Dunnett's correction were used for calculating *p* value. Asterisks marked on top of each bar are the *p* value when each mutant virus was compared to the wild-type. (ns, p≥0.05; * *p*<0.05, ** *p*<0.01, **** *p*<0.0001)

5.4 Discussion

Differences between HIV subtypes are critical not only in HLA-related impact on viral control but also in several protein functional aspects. For instance, both HLA-B*07:02 and HLA-B*35 (HLA-B*35:01) were associated with rapid disease progression in B-clade infection, but not in C-clade (Matthews *et al*., 2012; Kløverpris *et al*., 2014).

With respect to functional clade-specific differences, Tat protein in C-clade is defective for monocyte chemotactic activity, whereas Tat is capable of facilitating monocyte migration in other subtypes (Ranga *et al*., 2004). Another example was the downregulation of HLA-A and B expression. B-clade Nef protein exhibits the greatest reduction of expression of these molecules, while C-clade has the lowest *in vitro* activity (Mann *et al*., 2013). As one of the strongest protective alleles, it is important to address the differences of HLA-B*27:05-mediated immune responses between the two dominant subtypes that are responsible for the main global HIV epidemics (Hemelaar *et al*., 2011).

In this study, the unique co-selection of mutations associated with a HLA-B*27:05-restricted KK10 epitope in C-clade HIV infection was identified. Due to the low frequency of HLA-B*27:05 in sub-Saharan population who are mainly infected with C-clade viruses (The Allele Frequency Net Database, www.allelefrequencies.net; Hemelaar *et al*., 2011), studies to date have not yet evaluated the clade-specific patterns of escape and compensatory mutations with this HLA allele. From our longitudinal study of a transmission pair infected with C-clade viruses, S165N and V218M were identified as putative compensatory mutations as they were co-selected with R264K in the HLA-B*27:05-positive recipient. Among a group of C-clade infected HLA-B*27:05-positive individuals, D260N emerged with R264G in three cases. This mutation was therefore considered as a putative compensatory mutation for R264G, as E260D (a mutation in the same position) was previously reported to partially compensate R264G in B-clade infection (Schneidewind *et al*., 2008). T173X was also included as a potential candidate since S173A has been shown to have a compensatory effect for R264K in B-clade infection (Schneidewind *et al*., 2007). However, frequency analysis of T173X in HLA-B*27:05-positive R264X individuals indicated that changes in this position might not affect KK10 compensation. In fact, T173S did not reach significance in our co-variation analysis, while T173A had not

been observed in C-clade infection, suggesting mutations at this position was not influential in KK10 compensatory mechanism. This was verified in the fitness assay and the greatly diminished VRC partially explained the reason that this position was not preferred in C-clade. T173A was destructive for the virus. Similar patterns were observed for V218M and D260N. Neither reached significance in the co-variation analysis between B- and C-clade. Developing either of these two variants was not beneficial to VRC. They hardly act as compensatory mutations to R264K and R264G respectively in C-clade, since they further decreased the VRC instead of restoring it. This left S165N, among all the putative compensatory mutations we examined, to be the only one that is capable of compensating R264K, bringing its VRC to 91.5% relative to wild-type.

In B-clade viruses, the role S165N played was more complicated. From the co-variation analysis, we found the selection pressure is stronger in B-clade than in C-clade (OR= 27.7 versus 6.5; respectively), regardless of the higher background caused by high expression of HLA-B*57:03 having S165N as an escape mutation in C-clade infection (Carlson *et al*., 2012b). The fitness cost of S165N was not greater in C-clade. In fact, it was more significant in B-clade. Our fitness data showed that the compensatory effect of S165N was less than S173A in B clade infection, consistent with the early selection in combination with R264K. S165N may serve as an additional compensatory mutation, since S165N was only selected in the presence of S173A (q=2e-5). However, since S173A almost completely corrects the fitness cost of R264K, having S165N with S173A and R264K did not have an observable impact on VRC. It remains possible that S165N provides some small incremental additional fitness benefit to the virus on top of the major restoration towards fully efficient replicative capacity resulting from S173A.

The three residues R264K, S173A and S165N are all located in the N-terminal domain (NTD) of the capsid protein (**Figure 5.5**). It is clear that R264K (helix 7) and S173A (helix 2) are proximate and facing the same interface (Schneidewind *et al*., 2007). S165N, on the other hand, is relatively remote from R264K and located between helix 1 and 2. Since helix 1 and 2 are responsible for the assembly of capsid hexamer, it is possible that the co-selection of S165N was to enhance viral assembly, which might be affected by S173A (von Schwedler *et al*., 2003; Mateu, 2009). In addition, S165N has proximity of the intersubunit NTD-CTD interaction where several cellular factors interact and is important for structural stability (Gamble *et al*., 1997; Bhattacharya *et al*., 2014; Price *et al*., 2014). In C-clade infection, a strong compensatory effect might not be required since the reduction in VRC by R264K was relatively small compared to that observed in B-clade. In this case, S165N might compensate R264K by other structural modifications instead of by interacting in the same surface. To further verify if S165N is a compensatory mutation to R264K in both clades, future studies are required to reveal the dynamics of developing S165N and the role it plays in the KK10 compensatory mechanism.

Of note, a similar unusual pattern can be seen in the selection of V218M along with S165N and R264K in C-clade infection. V218M appeared slightly later than S165N and R264K in the longitudinal study and was disadvantageous compared to the S165N compensatory effect. Since V218M is located in the Cyclophilin A (CypA) loop where it is more critical for biological interaction with host cell factors rather than structural stability, the selection of V218M might not be related to KK10. It is also possible that this position was another undefined escape mutation restricted by other HLA alleles (the recipient's HLA class I: HLA-A*01:01/02:01, B*27:05/51:01, C*02:02/15:02).

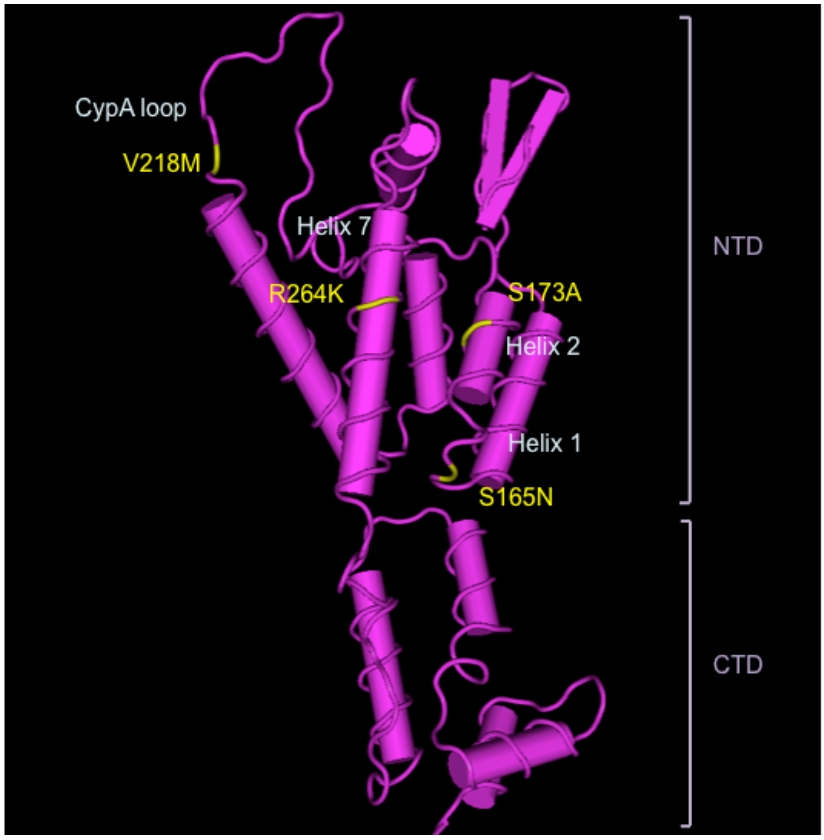


Figure 5.5. HIV-1 NL4-3 p24 Gag secondary structure. Protein's NTD and CTD are indicated. The KK10-related mutations are highlighted in yellow. R264K is located in helix 1; S173A is in helix 2; S165N is located between helix 1 and 2; and V218M is located in CypA loop. Figure is constructed and modified by using Cn3D (https://www.ncbi.nlm.nih.gov/Structure/pdb/2M8N).

The fitness assay used in this chapter is different from the previous two chapters. The site-directed mutagenesis was performed in p6.5 or p83-2 which encodes the 5'end of HIV proteins, whereas p83-10_{GFP} carries the GFP gene along with the 3'end of the viral genome. Details of plasmids can be found in **Section 2.7** and **Appendix**. The biggest advantage of using this system compared to the ΔGag-protease route is that the VRC was measured directly on the virus instead of infected cells. This offers a more accurate evaluation, although the quantity of the viruses should correlate with the number of infected cells. The GFP-positive expression can peak to 70-80% for the wild-type viruses in the MT4 cell line, while the infected cell percentage plateaus at 30-40% in GXR cells. In our examination of B- and C-clade wild-types viruses, there was no difference in VRC between B- and C-clade in the p6.5/p83-2 with p83-

10_{GFP} system. In contrast, our results in the previous chapters using ΔGag-protease system (Chapter 3 and 4) and previously published studies (Ariën *et al.*, 2007; Abraha *et al*., 2009) support the fact that C-clade wild-type virus is less fit than the B-clade NL4-3.

Taken together, a different compensatory pathway of rescuing HLA-B*27:05-restricted-KK10 escape mutation was identified in this study. This demonstrated that even in the 'protective' allele, there are still many different patterns or underlying mechanisms at work. In this case, it is in a clade-specific manner.

CHAPTER 6: NEUTRAL IMPACT OF HLA-B*40:01 ON HIV IMMUNE CONTROL IN AN B-CLADE-INFECTED TAIWANESE COHORT

6.1 Introduction

Since HLA class I molecules were found to be associated with different disease progression in HIV infection, studies have focused on those either associated with increased rates of progression ('disease-susceptible') or those associated with decreased rates of progression ('protective') (Emu *et al*., 2008; Pereyra *et al*., 2008). Protective alleles such as HLA-B*27 (HLA-B*27:05), HLA-B*57 and HLA-B*81 (HLA-B*81:01) are enriched in long-term non-progressors (LTNPs) and elite controllers (ECs) have been particularly well studied (Kiepiela *et al*., 2004; Pereyra *et al*., 2008; Wright *et al*., 2010; Carlson *et al*., 2012a). There is a clear relationship between the strength of selection pressure imposed on the virus by HLA alleles and their degree of protection (Matthews *et al*., 2008). As many studies (Wright *et al*., 2010; Gijsbers *et al*., 2013; Murakoshi *et al*., 2015) and previous chapters of this thesis show (**Section 4.3.2**), the frequency of escape mutations is significantly higher in association with protective HLA alleles (Matthews *et al*., 2008). To date, very little is known about the immune modulation by alleles which do not consistently link to better or worse disease progression. When it comes to vaccine strategy, it is important to understand what accounts for differences in outcome among individuals expressing middle-ranking HLA alleles neither associated with disease susceptibility nor with slow progression, since these account for approximately 50% of most populations.

Herein we selected HLA-B*40:01 as a representative allele in a Taiwanese cohort since its frequency is up to 44% in the population (The Allele Frequency Net

Database, www.allelefrequencies.net). In the natural course of HIV infection, the effect of HLA-B*40:01 on HIV disease progression is not significant against both B-clade and CRF01_AE viruses (Pereyra et al., 2010; Naruto *et al*., 2012; Mori *et al*., 2014). In a Thai cohort, this allele was protective against treatment failure related modification of ART (Tsuchiya *et al*., 2014).

HLA-B*40:01 belongs to the B44 supertype family, with a preference for acidic residues (Aspartic acid or Glutamic acid) at the B pocket. In the F pocket, the B44 family accommodates aromatic, aliphatic and hydrophobic molecules (F, Q, Y, L, I, V, M, W, A) (Sidney *et al.,* 2008). **Table 6.1** shows the characteristics of HLA-B*40:01-restricted epitopes (Marsh *et al*., 2000). HLA-B*40:01 prefers Glutamic acid (E) at position 2 and Leucine (L) at the carboxyl-terminus (Marsh *et al*., 2000). Of note, with respect to the discussion above regarding effective CD8+ T-cell responses imposing strong selection pressure on the virus, HLA-B*40:01 drives the selection of multiple escape mutants. In three well-characterised HLA-B*40:01-restricted epitopes, E to D mutations at position 2 are escape mutations in Gag and Pol epitopes, whilst in the Nef epitope the footprint involves a substitution at the carboxy-terminal residue Leucine to Methionine. There are six epitopes restricted by HLA-B*40:01 in total (Altfeld *et al*., 2000; Carlson *et al*., 2012a). However, there is no restricted mutation or variant within two of these: the Pol RT IL8 (202-209, IEELRQHL) and Nef LS9 (37-45, LEKHGAITS) epitopes. In particular, the Nef epitope LS9 is not well-defined and its existence is somewhat dubious in view of the Serine at the carboxy terminal position, and the extreme conservation of the motif across the other 5 epitopes, with Glutamic acid at P2 and Leucine at Carboxyl-terminus (Altfeld *et al*., 2000). The Nef KL9 (92-100, KEKGGLEGL) was also excluded due to the lack of evidence of selection pressure driven by this response. In this study, we focused on the three epitopes in Gag and Pol known to impose selection pressure on the virus (**Figure 6.1**) since these epitopes are potentially associated with effective CD8+ T cell

responses (Kiepiela *et al*., 2007; Satcha *et al*., 2007; Matthews *et al*, 2008; Chen *et al*., 2012a). Though not always, the escape mutations occurring in these conserved regions tend to have a greater fitness cost (Dahirel *et al*., 2011; Miura *et al*., 2009a).

Table 6.1. Characteristics of B*40:01-restricted epitope.

	Peptide-Binding Motif				
Position	**P1**	**P2**	**P3**	**P5-7**	**PC**
Binding pocket	A	B	D	C	F
B*40:01	-	E	-	I/V	L

Table is based on The HLA Factsbook (Marsh *et al*., 2000).

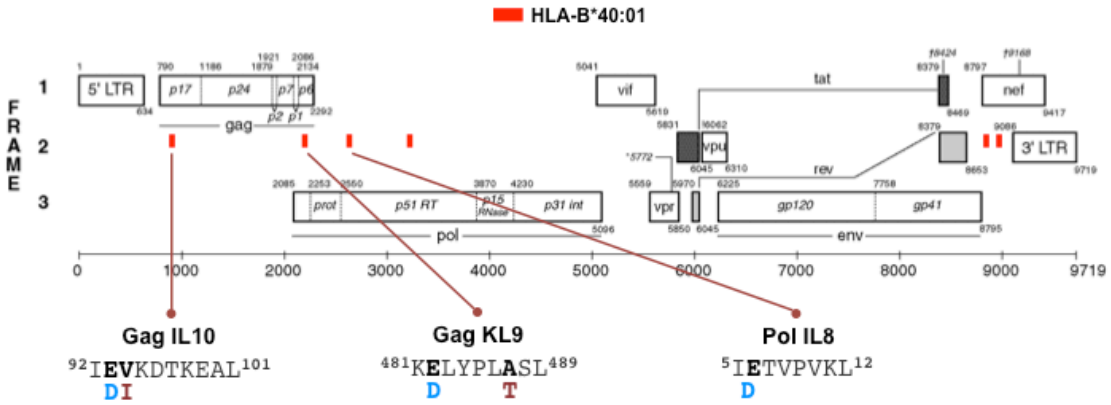


Figure 6.1. Schematic views of HLA-B*40:01-restricted epitopes in HIV genome. Red boxes indicate the HLA-B*40:01 epitopes. Three epitopes accentuated in this study are listed with the restricted mutations (blue) as well as the variants (dark red). Figure is based on Carlson *et al.,* 2012a; modified from LANL HIV Sequence Database, https://www.hiv.lanl.gov/content/sequence/HIV/MAP/landmark.html.

The aim of this study was, first, to compare the frequency of HLA-B*40:01-restricted mutations between HLA-B*40:01 positive progressors and controllers, along with a HLA-B*40:01 negative group. Very few studies have examined the relationship between HLA and HIV in Taiwanese populations, and given the high prevalence of this allele, this should provide high power to detect protective or disease-susceptible effects in this population. Second, we sought to identify features of subjects expressing HLA-B*40:01 associated with control of viraemia, and features in subject

expressing HLA-B*40:01 associated with lack of control of viraemia. We hypothesized that HLA-B*40:01-positive subjects who were controllers of HIV infection would be more likely not to have viral mutations within key epitopes associated with control of viraemia, whereas HLA-B*40:01-positive progressors would be more likely to carry these mutations. In this way we hypothesized we might identify key epitopes capable of mediating control of HIV among HLA-B*40:01-positive individuals.

6.2 Chapter-specific Materials and Methods

6.2.1 Study subjects

In total, 156 chronically HIV-infected, treatment-naïve individuals were recruited in Taipei Veterans General Hospital, Taipei, Taiwan. Characteristics of the participants are listed in **Table 6.2**. This is a unique cohort for the following reasons: first, it is a monoethnic (100%) and a male-dominant cohort (99.4%). Secondly, the majority of the patients were men who have sex with men (MSM), and were shown to be infected with B-clade virus. All subtypes were further confirmed with NCBI (website) and REGA (website) HIV genotyping tools. Lastly, the frequency of HLA-B*40:01 was high at 32.9%. Although it was slightly lower than the frequency in large cohorts, the number was sufficient to conduct the statistical analysis (The Allele Frequency Net Database, www.allelefrequencies.net).

In addition to the Taiwanese cohort, PBMCs of 13 B-clade chronic infected Caucasian individuals from the Thames Valley Cohort [mean CD4 count: 432 cells/mm3, median viral load: 50,914 copies/ml (IQR 13,314-105,352 copies/ml)] were used in the *ex vivo* ELISPOT assay for functional studies.

Table 6.2. Characteristics of studied participants from Taipei, Taiwan (*n*=156).

Characteristic	
Recruitment year	**21/10/2013 – 01/02/2016**
Ethnicity	
Chinese Han	**156 (100)**
Sex (*n*=156)	
Male	**155 (99.4)**
Female	**1 (0.6)**
Risk Group (*n*=156)	
MSM	**147 (94.2)**
IDU	**9 (5.8)**
CD4+ T cell count, cells/mm^3 (*n*=155)[a]	**355 [239-466]**
Viral load, copies/ml (*n*=155)[a]	**27,600 [8,945-28,400]**
HLA-B typing (*n*=140)	
HLA-B*40:01+ve	**46 (32.9)**
HLA-B*27:04+ve	**4 (2.9)**
HIV subtype (*n*=142)[b]	
B	**121 (85.2)**
CRF07_BC	**10 (7.0)**
CRF01_AE	**10 (7.0)**
CRF01_AE + CRF07_BC	**1 (0.7)**

*Data are no. (%) of patients, otherwise indicated. IQR is shown in brackets. [a]CD4+ T-cell and viral load measurements were performed at enrolment. [b]HIV subtype was determined by p17 and p24 Gag sequences.

6.2.2 Defining viral controllers and progressors

Previous studies have defined a viraemic controller as an HIV-seropositive individual whose viral load is <2,000 copies/ml, whereas a progressor's viral load is >10,000 copies/ml (Deeks and Walker, 2007; Pereyra et al Science 2010, Goulder and Walker, 2012). Here, due to the limited number of samples, we set the cut-off value of the viral load to be 10,000 copies/ml for a viraemic controller and ≥10,000 copies/ml for a progressor. These same values have been applied in other studies (Poropatich and Sullivan, 2010; de Medeiros *et al*., 2016).

6.2.3 Statistical analysis

All the data were analysed via the D'Agostino and Pearson normality test. A one-way ANOVA (parametric) or Kruskall-Wallis test (non-parametric) was applied according to normality of data distribution in group analyses. For individual tests, a Mann-Whitney U-test was performed, in the case when the sample size was small. When the results of both group and individual tests were not significant, correlational tests (e.g. Bonferroni correction) were not applied.

6.3 Results

6.3.1 Dominance of B-clade and the identification of a unique recombinant form of CRF01_AE/CRF07_BC

In the Taiwanese HIV epidemic, each HIV subtype is associated with a specific high-risk group, namely, B-clade in MSM, CRF07_BC in IDUs and CRF01_AE in heterosexuals (CDC, R.O.C. Taiwan, 2016). Our cohort reflects this, as B-clade is the dominant circulating strain in MSM. These current analyses described below were restricted to the B-clade infected subjects (**Table 6.2** and **Figure 6.2**). However, there were patients in the MSM group infected with CRF07_BC. A unique recombinant form (URF) of virus between CRF01_AE and CRF07_BC has been identified (**Figure 6.2**). This is a reinforcement of the 'template switching' during the reverse transcription of HIV (Jung *et al*., 2002). The CRF01_AE strain itself clustered into two different groups in the phylogenetic tree (**Figure 6.2**). One is the cluster of sequences closely related to the consensus highlighted in dark purple, likely transmitted from Thailand (CDC, R.O.C. Taiwan, 2016). The other cluster shown in light purple is shifting from Consensus AE to B-clade sequences, supporting the idea of mixing and adaption to the main circulating strain.

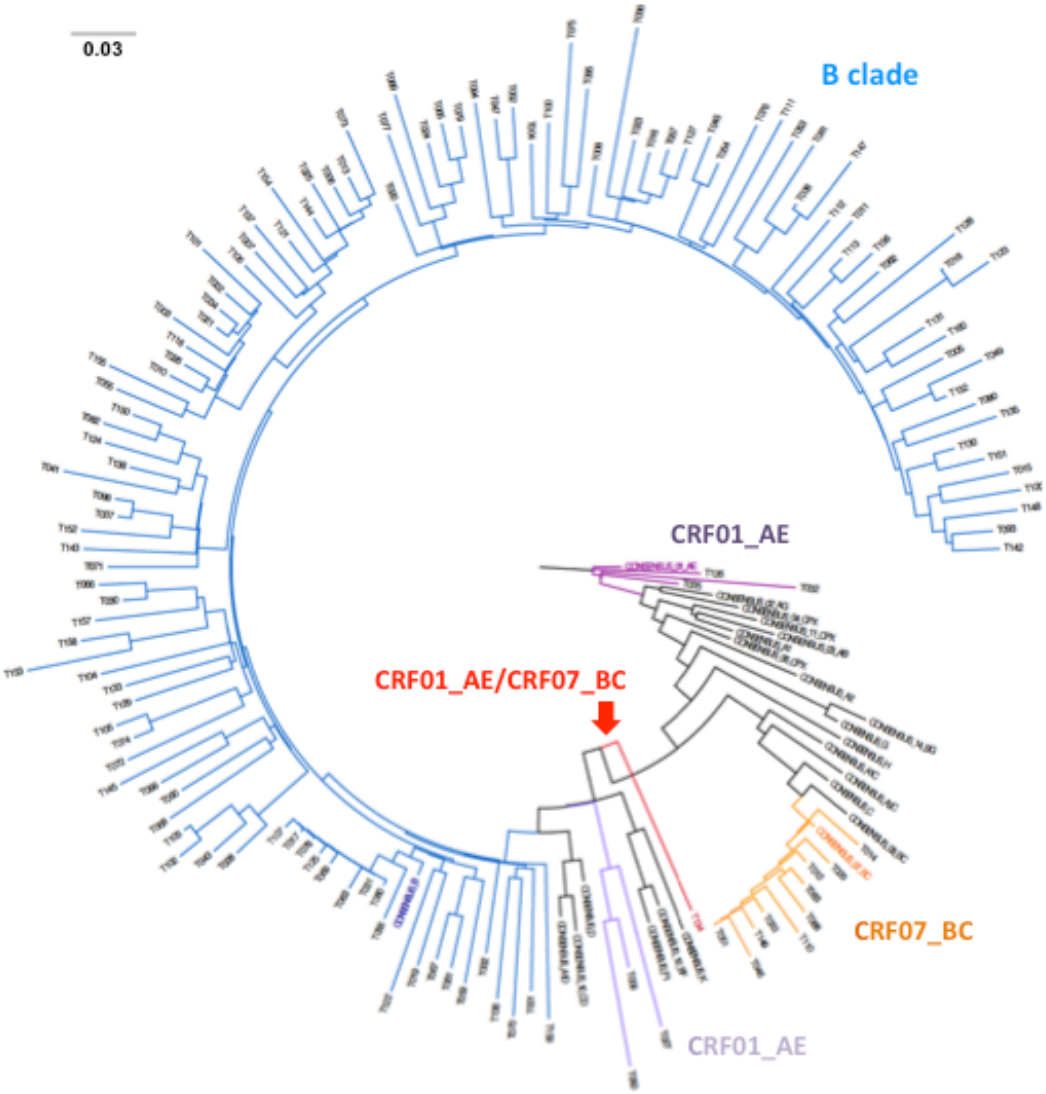


Figure 6.2. Maximum-likelihood phylogenetic tree of p17+p24 Gag sequences of all individuals in Taiwanese cohort. B-clade, CRF07_BC, CRF01_AE and CRF01_AE/CRF07_BC are shown in blue, orange, purple and red, respectively. Nucleotide length represented in the tree is indicated at top-left corner.

*6.3.2 Wide range of both CD4 count and viral load among HLA-B*40:01-positive individuals*

Similar to previous findings, HLA-B*40:01 in our study was not associated with either better or worse disease progression in B-clade infection (Pereyra *et al.*, 2010; Naruto *et al.*, 2012). The effect of HLA-B*40:01 was minimal, compared with some protective alleles such as HLA-B*52:01 (**Chapter 4**; Murakoshi *et al.*, 2015; Sakai *et al.*, 2015).

In addition, among all the HLA-B alleles, a wide range of both CD4 count and viral load of HLA-B*40:01-positive individuals was observed (**Figure 6.3 A** and **B**). In this cohort, the HLA-B*40:01-positive group included an individual with the highest viral load (>10^7 $\log_{10}$ copies/ml) and the only one viral controller with undetectable viral load in this cohort.

The mean CD4 count in this cross-sectional ART-naive cohort was 355 cells/mm^3, which indicated that almost half of the patients remained treatment-naïve even thought their CD4 count was under 350 cells/mm^3, the recommended CD4 criterion for ART initiation in 2013 as per WHO guidelines (WHO, 2013). Since the viral load is relatively stable after establishment of the viral set-point, and the CD4 count drops at the rate of 50-70 cells/mm^3 per year (Keller *et al*., 2009), HLA effects may be harder to discern from CD4 count than from viral load in a heterogeneous cross-sectional cohort whose duration of infection is mixed and entirely unknown. This is further supported by statistical analysis. The impact of HLA differences on viral load was closer to significance than on CD4 count (p=0.08 in Kruskall-Wallis test compared to p=0.50 in one-way ANOVA, **Figure 6.3 A** and **B**). There was a weak correlation between CD4 count and viral load (r^2=0.10) (**Figure 6.3 C**), lower than in most published studies (typically r^2=0.30) (Lyle *et al*., 2000). **Figure 6.3 D** displays the grouping of HLA-B*40:01-positive individuals used in these studies. In total, there were 16 and 28 individuals in the low and high viral load groups, respectively.

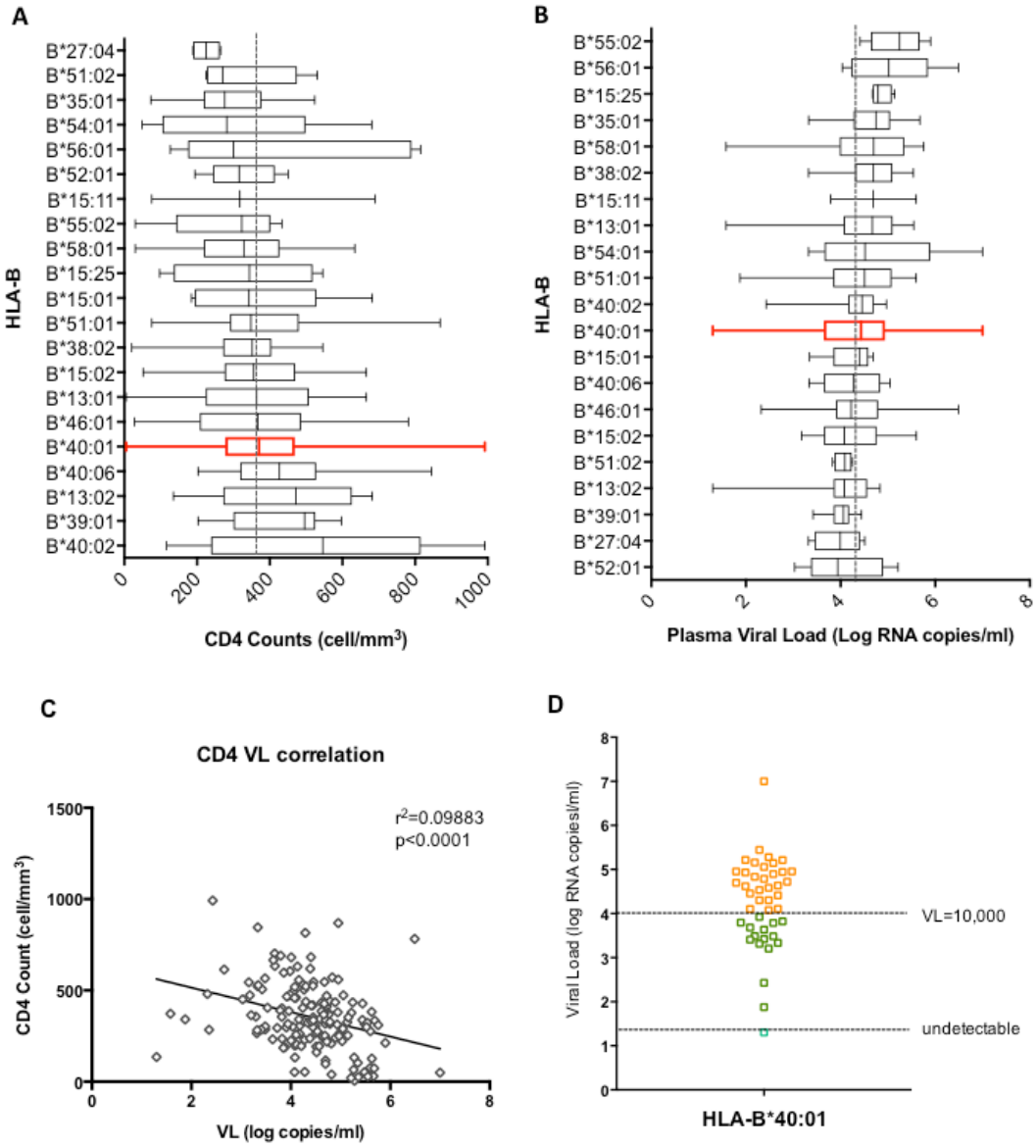


Figure 6.3. Association of HLA-B alleles with CD4 count and viral load (only n>3 alleles were included, *n*=125); and the determination of high (progressors) and low viral load group (controllers). Relationship of **(A)** CD4 count and **(B)** viral load with HLA-B alleles. One-way ANOVA and Kruskall-Wallis test were applied to each allele and the mean or median were calculated. HLA-B*40:01 is highlighted in red, whereas HLA-B*27:04 is shown in pink. Bonferreni correction test was applied after individual Student's t-test or Mann-Whitney U-test. *m*=21, corrected *p* value threshold is 0.0023. **(C)** Correlation of CD4 and viral load (*n*=125). Pearson's correlation *t*-test was applied. **(D)** Grouping for HLA-B*40:01-positive individuals according to their viral loads (*n*=44). High and low viral load groups are shown in orange and green, respectively.

*6.3.3 No evidence of escape selection within HLA-B*40:01 p17 Gag IL10*

The HLA-B*40:01-restricted p17 Gag epitope IL10 (92-101, IEVKDTKEAL) has an HLA-B*40:01 footprint E93D (Carlson *et al*., 2012a). However, in the current study cohort there was no difference in the frequency of E93X between the HLA-B*40:01-positive and -negative groups (**Figure 6.4 A**), neither was there a frequency difference in this mutation between HLA-B*40:01-positive low and high viral load groups (**Figure 6.4 B**). This may reflect the high prevalence of HLA-B*40:01 in Taiwan and the accumulation of the E93D mutation, but more likely there are other factors influencing the nature of the residue at this position since there is no evidence at all of selection pressure at this position in this cohort. Furthermore, the P2 escape along with the other variants (G, Q and K) had no impact either on viral load or CD4 counts in all subjects and in the HLA-B*40:01-positive group (**Figure 6.4 C**), though there was a weak trend of higher viral loads and lower CD4 counts. Similar results could be seen in V94I, the variant at position 3 in the same IL10 epitope (**Figure 6.5**) with no evidence of escape selection in HLA-B*40:01-positive subjects in this cohort.

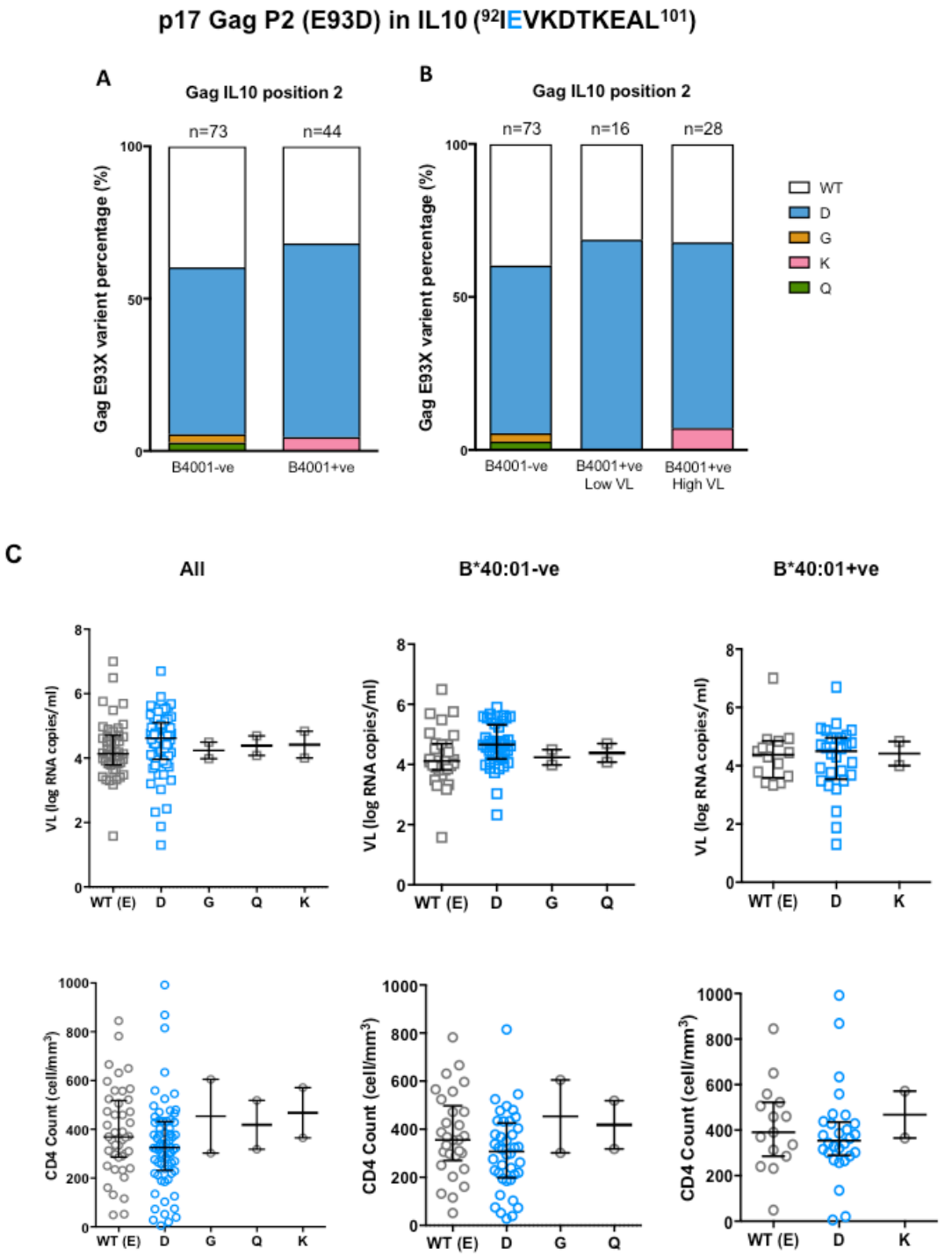


Figure 6.4. Frequency and viral load of HLA-B*40:01-restricted Gag E93D (P2) mutation in IL10 epitope. (A) Percentages of E93X variants in HLA-B*40:01-positive and -negative groups. **(B)** Percentages of E93X variants in HLA-B*40:01-positive high and low viral load group as well as HLA-B*40:01-negative groups. **(C)** Viral loads (upper panels) and CD4 count (lower panels) of E93X variants in HLA-B*40:01-positive, -negative and all sample groups. Fisher's exact test for frequency difference; and Kruskall-Wallis test and Mann-Whitney U-test for CD4 count and viral load were applied. Multiple comparison test was performed.

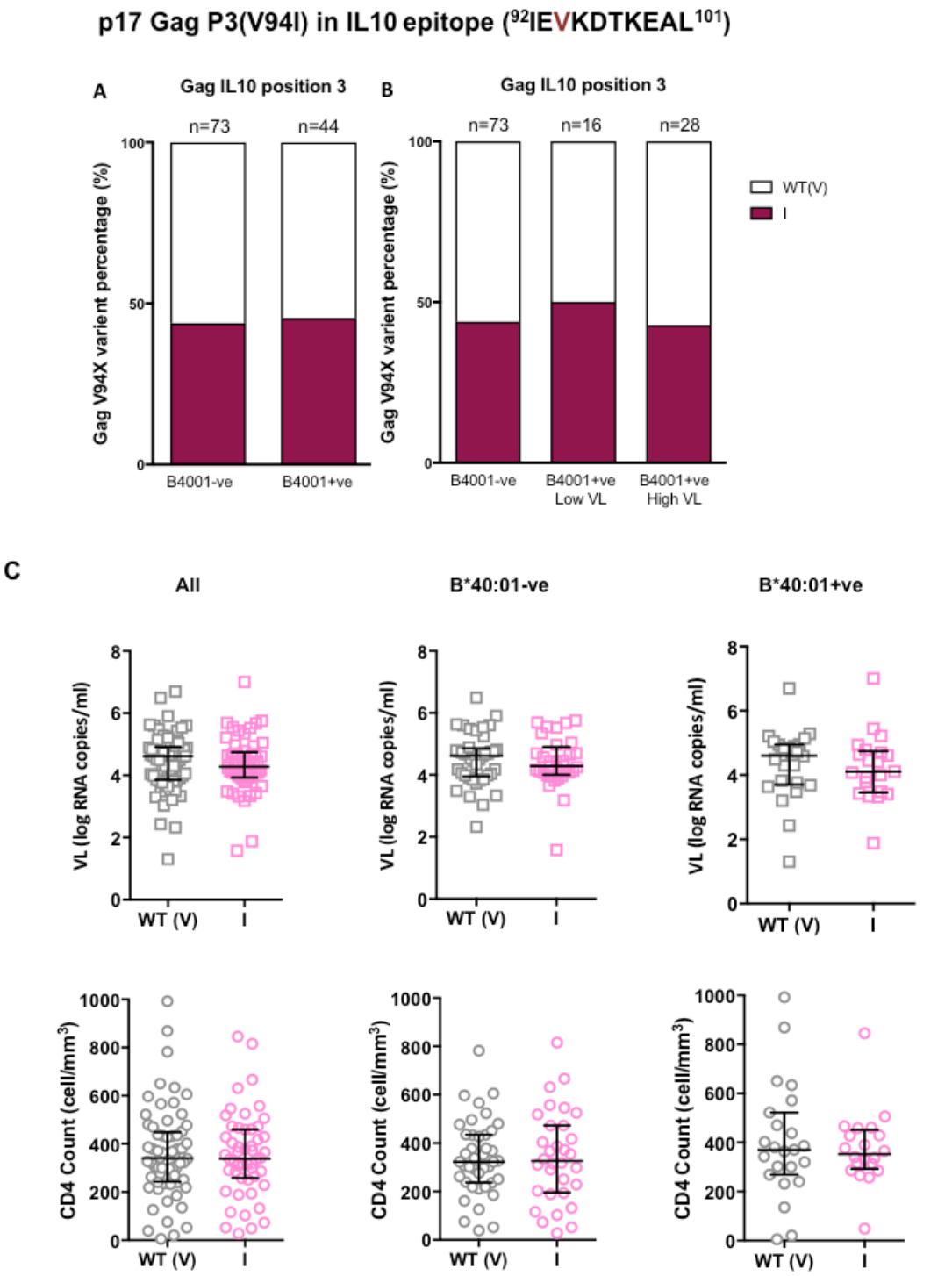


Figure 6.5. Frequency and viral load of HLA-B*40:01-restricted Gag V94I (P3) variant in IL10 epitope. (A) Percentages of V94X variants in HLA-B*40:01-positive and -negative groups. **(B)** Percentages of V94X variants in HLA-B*40:01-positive high and low viral load group as well as HLA-B*40:01-negative groups. **(C)** Viral loads (upper panels) and CD4 count (lower panels) of V94X variants in HLA-B*40:01-positive, -negative and all sample groups. Fisher's exact test for frequency difference; and Mann-Whitney U-test for CD4 count and viral load were applied.

*6.3.4 Selection of HLA-B*40:01 escape mutants in the p6 Gag KL9 epitope*

The frequencies of E482X and A487X were significantly higher in the HLA-B*40:01-positive group compared to HLA-B*40:01-negatives in the p6 Gag KL9 epitope (481-489, KELYPLTSL) (**Figure 6.6 A** and **6.7 A**). Particularly, at position 7, both low and high viral load groups had significantly higher percentages of variants compared to the negative group (**Figure 6.7 B**). However, there was no difference between the two HLA-B*40:01-positive groups (**Figure 6.6 B** and **6.7 B**). The common E482D mutation and A487T variant was not associated with differences in the viral load or CD4 count (**Figure 6.6 C** and **6.7 C**), suggesting that these mutants did not yield a significant viral fitness cost. At position 2, E482G tended to have a higher viral load and lower CD4 count in the HLA-B*40:01-positive group, which (notwithstanding the data shown in relation to HLA-B*52:02 mutants in Chapter 4) is to be expected (Carlson *et al*., 2016) – the escape mutations that are selected by the virus to improve viral fitness (**Figure 6.6 C**).

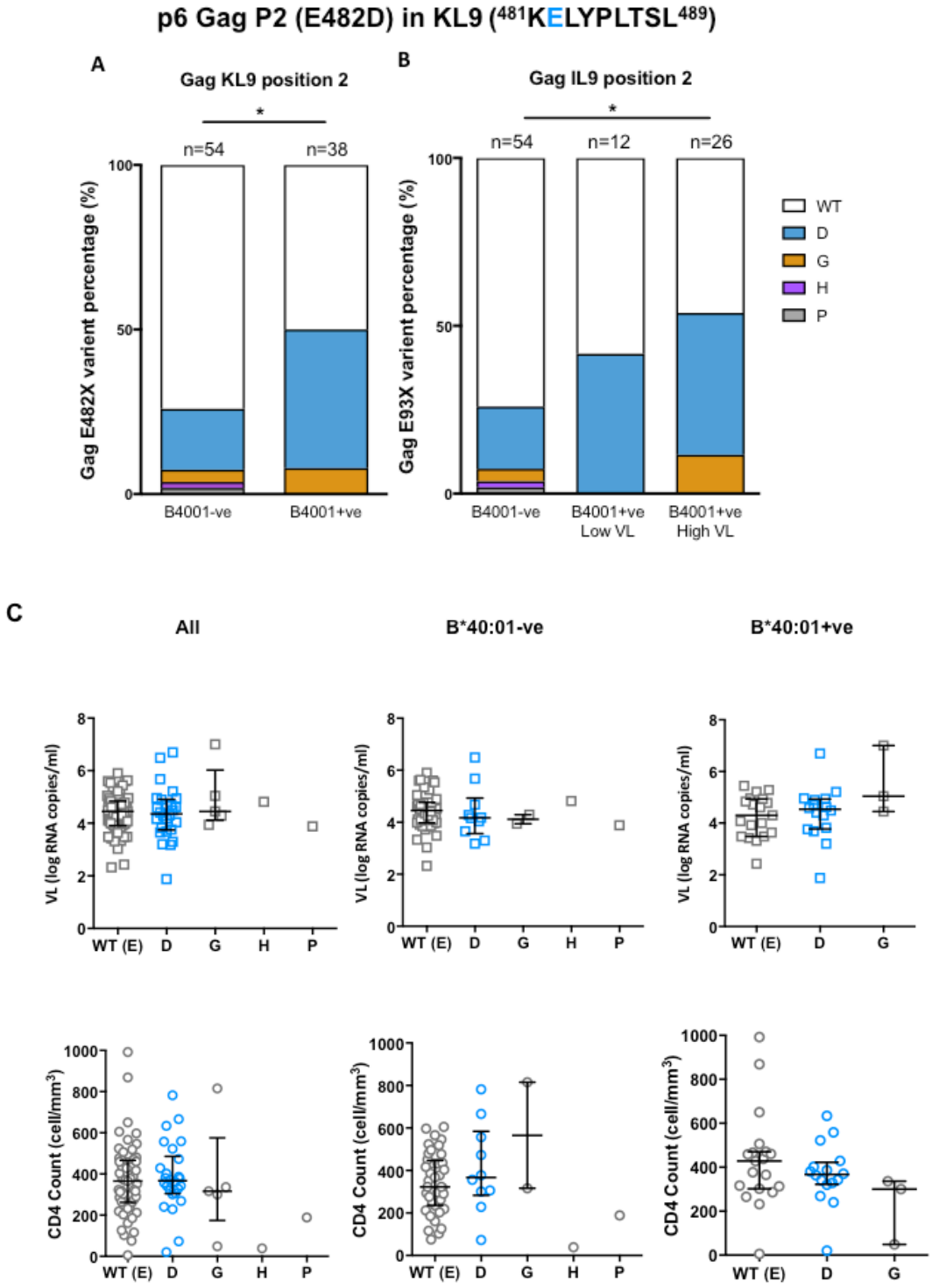


Figure 6.6. Frequency and viral load of HLA-B*40:01-restricted Gag E482D (P2) mutation in KL9 epitope. (A) Percentages of E482X mutations in HLA-B*40:01-positive and -negative groups. **(B)** Percentages of E482X mutations in HLA-B*40:01-positive high and low viral load group as well as HLA-B*40:01-negative groups. **(C)** Viral loads (upper panels) and CD4 counts (lower panels) of E482X mutations in HLA-B*40:01-positive, -negative and all sample groups. Fisher's exact test for frequency difference; and Kruskall-Wallis test and Mann-Whitney U-test for CD4 count and viral load were applied. (* p<0.05)

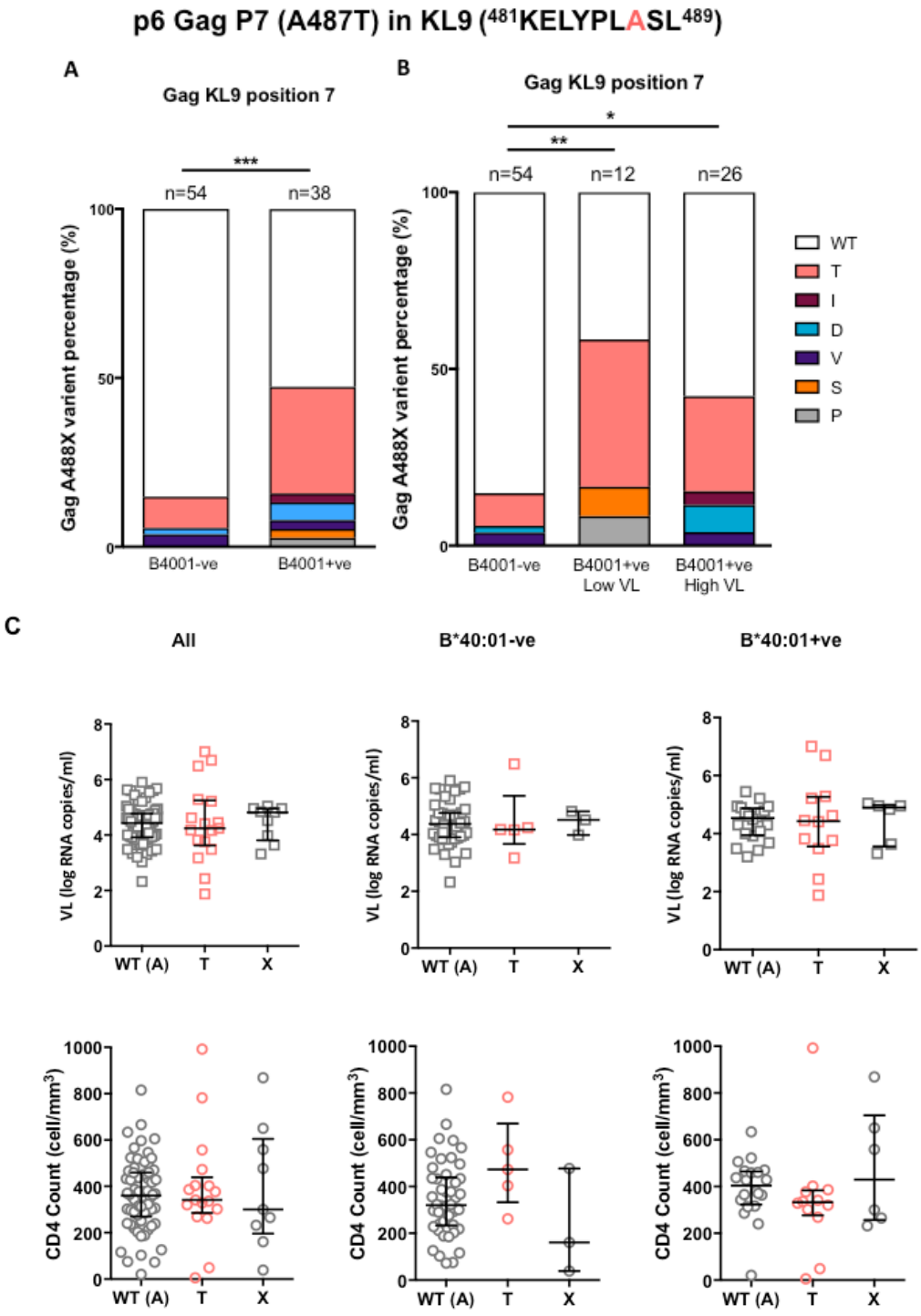


Figure 6.7. Frequency and viral load of HLA-B*40:01-restricted Gag A487T (P3) variant in KL9 epitope. (A) Percentages of A487X variants in HLA-B*40:01-positive and -negative groups. **(B)** Percentages of A487X variants in HLA-B*40:01-positive high and low viral load group as well as HLA-B*40:01-negative groups. **(C)** Viral loads (upper panels) and CD4 counts (lower panels) of A487X variants in HLA-B*40:01-positive, -negative and all sample groups. Fisher's exact test for frequency difference; and Kruskall-Wallis test and Mann-Whitney U-test for CD4 count and viral load were applied. (* p<0.05, ** p<0.01, *** p<0.001)

*6.3.5 No evidence of escape selection within HLA-B*40:01 RT IL8 epitope*

There was a trend of higher percentages of Pol E6X variants in both HLA-B*40:01-positive low and high viral load groups, compared to the HLA-B*40:01-negative group (**Figure 6.8 A** and **B**). However, the difference was not statistically significant and the mutant percentages were similar between the low and high viral load groups (**Figure 6.8 B**). This indicates that the E6D and E6G of the Pol IL8 epitope (5-12, IETVPVKL) did not contribute to the different viral loads in the two groups. In **Figure 6.8 C**, none of the variants had a significant impact on viral load or CD4 count.

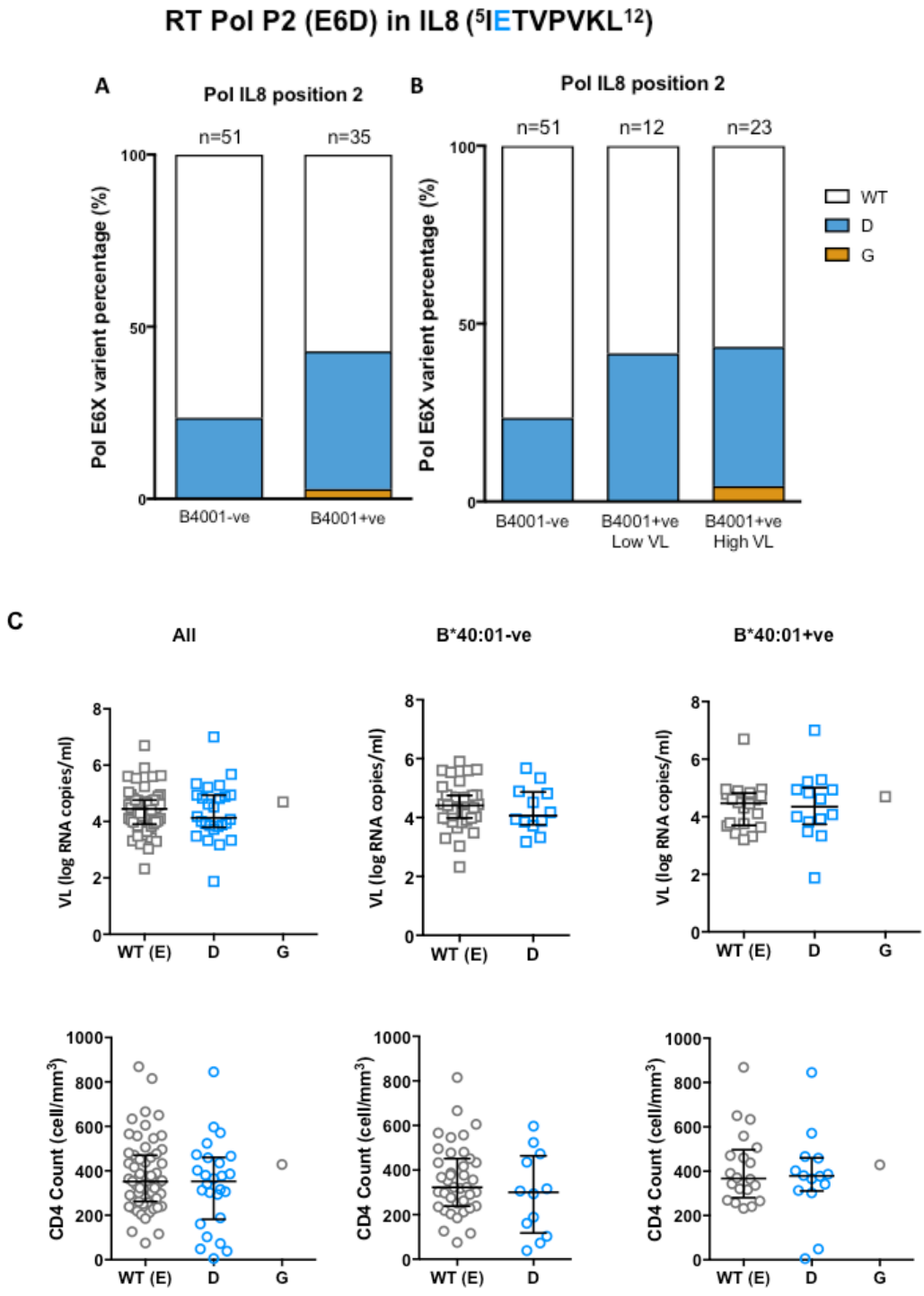


Figure 6.8. Frequency and viral load of HLA-B*40:01-restricted Pol E6D (P2) mutation in IL8 epitope. (A) Percentages of E6X mutations in HLA-B*40:01-positive and -negative groups. **(B)** Percentages of E6X mutations in HLA-B*40:01-positive high and low viral load groups as well as HLA-B*40:01-negative groups. **(C)** Viral loads (upper panels) and CD4 counts (lower panels) of E6X mutations in HLA-B*40:01-positive, -negative and all sample groups. Fisher's exact test for frequency difference; and Kruskall-Wallis test and Mann-Whitney U-test for CD4 count and viral load were applied.

*6.3.6 Higher numbers of targeted epitopes in HLA-B*40:01-positive viral controllers*

As mentioned above, examining the frequency of HLA-B*40:01-restricted mutations cannot offer an insight for an effective CTL response. Here we included HLA-B*40:01-positive B-clade infected 13 individuals from TVC and 1 individual from Taiwanese cohort in the ELISPOT assay testing recognition of 410 OLPs spanning the B-clade proteome to evaluate the magnitude and breadth of CTL responses (**Figure 6.9 A**). In total there were four controllers whose viral loads were less than 10,000 copies/ml (we included this one Taiwanese individual to increase the number in this controller group) and 10 progressors with viral load higher than 10,000 copies/ml. Overall, although there was no difference between the magnitudes of elicited responses, more targeted epitopes throughout the HIV genome were observed in the low viral load group (**Figure 6.9 B** and **C;** p=0.0005). For each epitope, response magnitudes were similar (**Figure 6.9 D**). For each protein, the low viral load group tended to have a greater magnitude of responses, particularly in immunogenic regions including Gag, Pol and Nef. This difference did not yield statistical significance consistent with previous studies (Addo *et al*., 2000) (**Figure 6.9 E**). Regarding the number of targeted peptides, low viral load groups showed a significantly higher number of Gag responses (**Figure 6.9 F**). None of the four individuals in low viral load group targeted Vif. Taken all together, HLA-B*40:01-positive individuals who were able to maintain viral control targeted more epitopes; but not in the HLA-B*40:01-restricted epitopes we have examined above. Together these data are consistent with the hypothesis that the HLA-B*40:01-restricted CD8+ T-cell responses make little contribution to immune control of HIV, and that although differences in viral load are strongly related to the breadth of the CD8+ T-cell response in HLA-B*40:01-positive individuals as previously shown in all HIV-infected subjects (Kiepiela *et al*., 2007), HLA-B*40:01 itself makes a neutral contribution.

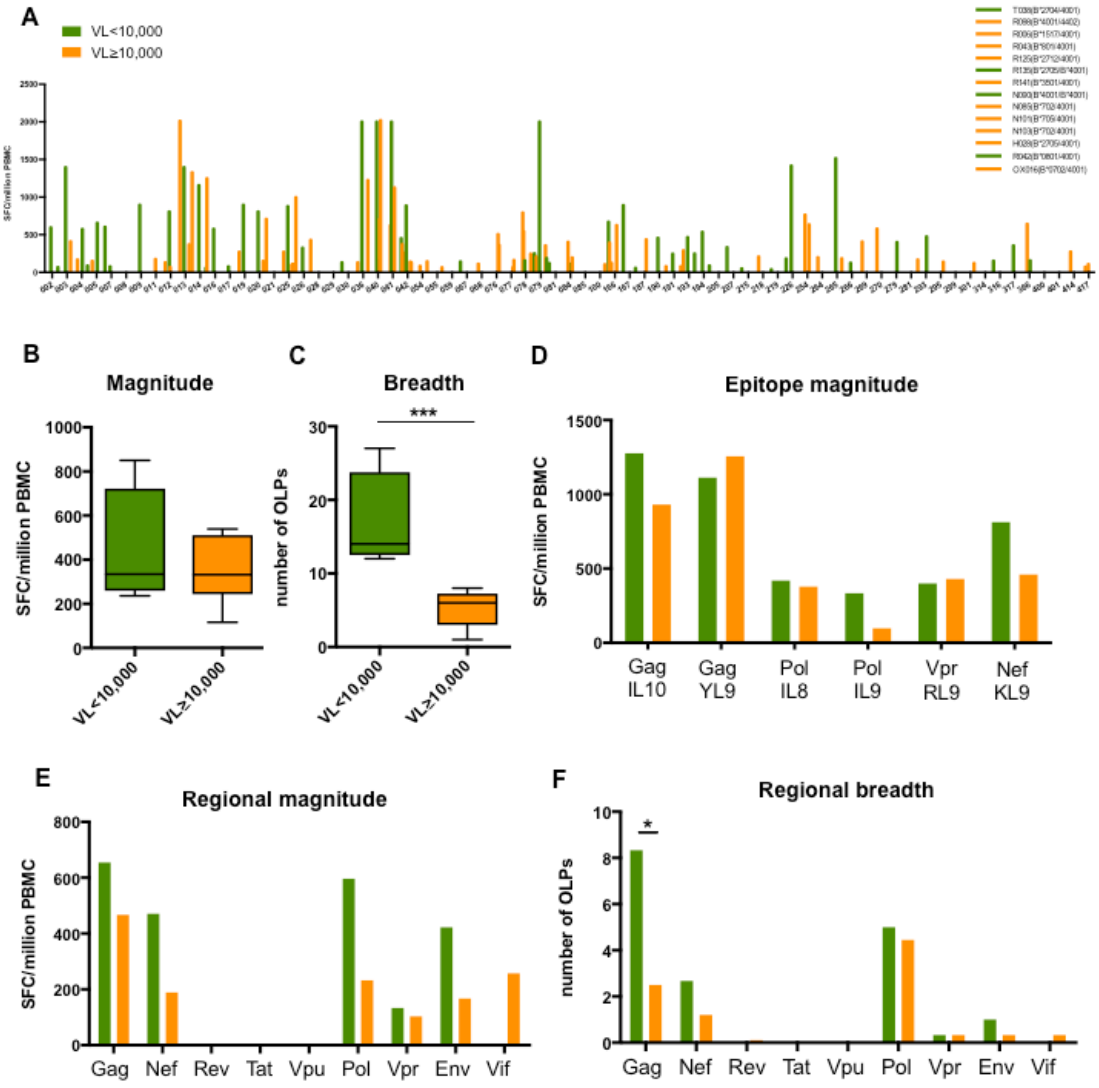


Figure 6.9. CTL IFN-γ ELISPOT results of B-clade infected HLA-B*40:01-positive individuals (*n*=13, TVC; *n*=1, Taiwanese cohort; *n*=4, low viral load group; *n*=10, high viral load group; viral load <10,000 and ≥10,000 are shown in green and orange). **(A)** Overall responses. **(B)** Magnitude of responses (SFC/10^6 PBMC) between low and high viral load groups. **(C)** Breadth of responses (number of OLPs) between low and high viral load groups. **(D)** Magnitude of responses (SFC/10^6 PBMC) for each epitope between low and high viral load group. **(E)** Regional magnitude of responses (SFC/10^6 PBMC) between low and high viral load groups. **(F)** Regional breadth of responses (number of OLPs) between low and high viral load groups. Mann-Whitney U-test was applied. (* p<0.05, *** p<0.001)

6.4 Discussion

HLA-B*40:01 is an example of an HLA allele with apparently neutral effects on HIV progression. However, it is clear from our study of a Taiwanese cohort, in which HLA-B*40:01 is highly prevalent, that HLA-B*40:01-positive individuals may have a diverse range of both viral loads and CD4 counts. To date, most of the HLA class I studies have been focused on the protective alleles (Kiepiela *et al*., 2004; Wright *et al*., 2010; Carlson *et al*., 2012a). It was hypothesized that analysis of viraemic controllers and non-controllers both expressing HLA-B*40:01 might provide insights into the mechanisms by which HLA-B*40:01 might modulate immune control of HIV infection, despite overall having a neutral effect. Here we traced the footprint of HLA-B*40:01 in the Taiwanese cohort and divided the HLA-B*40:01-positive subjects into low and high viral load groups based on a cut-off value of 10,000 copies/ml. To compare with global B-clade infected sequences, **Figure 6.10** shows the sequence data from Los Alamos database for the following discussion (LANL HIV Sequence Database, http://www.hiv.lanl.gov/). By definition, the frequency of an HLA footprint is significantly higher in the individuals carrying the respective HLA. However, HLA-B*40:01 escape mutants were only observed in one of the 3 epitopes analysed. Although HLA-B*40:01 is highly prevalent in the Taiwan population, the accumulation of escape mutants similar to the lost effect of HLA-B*51:01 in Japanese cohorts (Kawashima *et al*., 2009; Kawashima *et al*., 2010) does not explain this lack of any selection effect. The HLA-B*40:01-associated variants that have been studied here are observed in a minority of the population in most cases, and in the case of the p17 Gag epitope the E92D variant is more common in the HLA-B*40:01-negative population. In contrast, for HLA-B*51-TI8 mutation, the I135X variant was significantly more common among HLA-B*51-positive subjects even when the frequency was as high at 66% in the HLA-B*51-negative population. Similar findings apply to the D260E HLA-B*35:01-associated mutant which continued to accumulate

in populations even when the frequency was as high as 85-90% in the HLA-B*35:01-negative population.

An alternative and more likely explanation for lack of escape effects than escape accumulation is the opposing impact of other prevalent HLA alleles. For example, the p17 Gag escape mutant E93D is driven by HLA-A*11:01 (Carlson *et al*., 2012a) and thus the frequency of the E93D mutant among HLA-B*40:01-positive and negative groups needs to be undertaken taking this into account. Since the Taiwanese study cohort were not HLA-A typed this could not be done in this cohort. Approximately 55% of Taiwanese have HLA-A*11:01 (The Allele Frequency Net Database, www.allelefrequencies.net). The HLA-B*40:01-associated escape mutants were observed however in the Gag KL9 to a greater degree. However, there was no difference in the frequency of the E to D mutation at position 2 between the low and high viral load groups, suggesting that this epitope and its recognition is not crucial to immune control of viraemia.

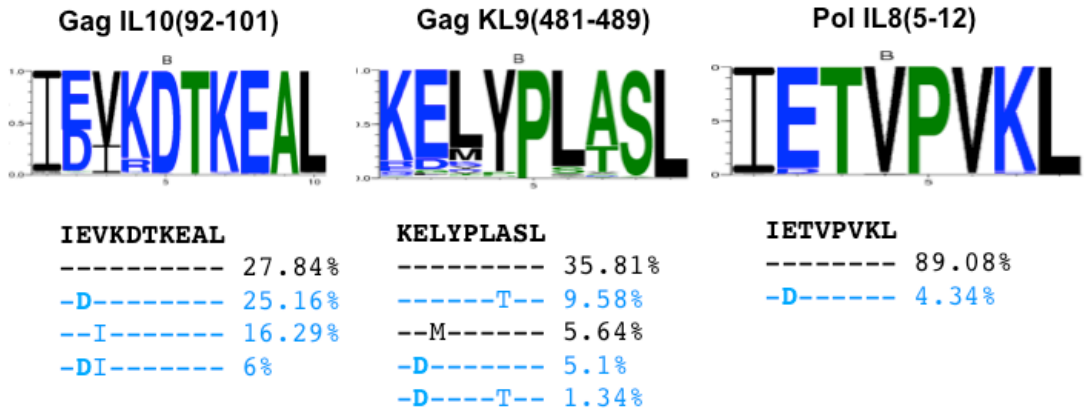


Figure 6.10. Variant percentages of HLA-B*40:01- and HLA-C*03:04-restricted epitopes in LANL HIV Sequence Database, http://www.hiv.lanl.gov/**.** The size of each letter in the epitope is proportional to the percentage to the level of expression within the population. All sequences are B-clade (*n*=1341). Figure is adapted and based on HIV sequence quick align, LANL HIV Sequence Database, https://www.hiv.lanl.gov/cgi-bin/ANALYZEALIGN/analyze_align.cgi.

Another HLA allele that might affect the frequency of the restricted mutation is HLA-B*40:02, a HLA-B*40 subtype which is common in East Asian population, in particular in Taiwan which can account for up to 23% (The Allele Frequency Net Database, www.allelefrequencies.net). However, in our cohort the expression of HLA-B*40:02 is only 4.4% and HLA-B*40:02 does not share the same epitopes we examined in Gag and Pol as HLA-B*40:01 (Carlson *et al*., 2012a). Therefore the HLA-B*40:01-restricted mutations are independent from the effect of HLA-B*40:02.

The data from analysis of the HIV-specific CD8+ T-cell response using a panel of overlapping peptides spanning the B clade proteome were notable for the presence of a strong effect on viral load mediated by breadth of the response, and especially the Gag response; in contrast, the lack of any evident HLA-B*40:01-restricted CD8+ T-cell response. These findings are fully consistent with previous studies that have highlighted the importance of breadth of the CD8+ T-cell response and in particular the Gag response (Kiepiela *et al*., 2007). The magnitude of the CD8+ T-cell response, however, has consistently not been associated with immune control (Addo *et al*., 2003). The conclusions that one might draw from these findings are that, although HLA-B*40:01 does present a number of HIV-specific CD8+ T-cell epitopes, and in some of them escape mutants are selected, the effectiveness of these responses does not appear to be such that they influence immune control in HLA-B*40:01-positve subjects. Presumably it is better to have the modest Gag- and Pol-specific CD8+ T-cell responses restricted by HLA-B*40:01 than none at all, as one might seen in subjects expressing HLA-B*58:02, a disease-associated molecule only presenting Env epitope (Ngumbela *et al*., 2008). In all likelihood, immune control or lack of it in HLA-B*40:01 depend on whatever the other HLA molecules are that accompany it, together with all the other immunological and non-immunological factors that contribute to outcome in HIV infection.

CHAPTER 7: CONCLUSION

HLA class I molecules have been shown to be the main weapon involved in combatting HIV infection. The immune responses mediated by each HLA are distinct, reflecting the extensive polymorphism observed. Much can be learned from the association between patterns of immune response and the differential HIV disease outcomes and that are associated with particular HLA molecules. Several studies have shown that the HLA-mediated CTL responses play a central role in suppressing HIV infection and maintaining immune control both in early and later time points during the course of infection (Kiepiela *et al*., 2004; Ndhlovu *et al*., 2015; Pereyra *et al*., 2010). However, HIV-specific CD8+ T cell-mediated immunity is incapable of eliminating the infection completely. Only a rare group of individuals (<1% of a given population), in most cases expressing protective HLA-B alleles, are able to suppress the viral load to undetectable level over ten years or more. To date, several studies have been focusing on unveiling the underlying mechanisms in these elite controllers, in the hope of achieving a so-called 'functional cure' or developing a 'therapeutic vaccine.'

One of the critical factors that impair the CTL immune control is the escape mutation(s) from the targeted epitope. Escape mutations act as a double-edged sword for the virus. On the one hand, these mutations help the virus evade the host's immune surveillance. On the other hand, this may occur at the cost of reducing the replicative capacity. The latter can be redeemed by further compensatory mutations. The attenuated viruses might offer a strategy for vaccine design.

In this thesis, I have examined four distinct HLA-B molecules from different supertype families, comprising from the protective HLA-B*81:01, HLA-B*27:05, and HLA-B*52:01 to the neutral HLA-B*40:01 in B- and C-clade infected adult populations.

These two HIV subtypes account for the majority of the HIV infections that have occurred in the global epidemic. Focusing on HLA footprints made by the immune response on the virus, I have investigated the impact of these HLA molecules on viral replicative capacity and on disease outcome. The relevance of HIV subtype is also highlighted here. Together, these studies have yielded a clearer picture of the escape and compensatory pathways available to the virus, and potential mechanisms by which the virus might be cornered by the CD8+ T-cell response.

The tool for evaluating VRC in this thesis is the fitness assay. Recently, several methods have been developed to introduce interested variant site(s) or region(s) of interest into laboratory strains. In the order of ascendant size of the introduced sequence, there have been approaches of introducing a single or multiple point mutation(s); inserting autologous viral fragment of sequence between enzyme restriction sites; and cloning autologous viral sequence between deleted sites into a standardised lab strain backbone (Miura *et al*., 2009a ; Prado *et al*., 2009). The full-length molecular clone of the autologous virus is available as well (Claiborne *et al*., 2015; Deymier *et al*., 2015). Certainly, the larger the fragment of autologous viral sequence introduced, the more representative are the results of the *in vivo* scenario. However, the disadvantage of this approach is not being able to rule out the irrelevant variants when examining the impact made by one or a few mutations. To best identify the pattern of escape and compensatory mutations, the method of introducing single point mutation by site-directed mutagenesis was called upon. Two systems were used in this thesis: ΔGag-protease pNL4-3 and p83-2/p6.5 and p83-10_{GFP}. The former approach was applied in Chapter 3 and Chapter 4. Most of the VRC studies in these two chapters focused on the point mutations, except one family member's Gag-protease sequence were inserted in to the ΔGag-protease plasmid for better understanding of the actual VRC in the peadiatric slow progressor. Because the reporter GFP signal is encoded in a Tat-dependent pathway in GXR

cell, the VRC measured by this method relies on the percentage of the infected cells instead of the actual amount of virus. Hence, in Chapter 4, another quantifying assay (p24 ELISA) was performed and the result correlated nicely with the fitness assay. In Chapter 5, since the focus of the study was the dynamics and pattern of particular mutations, the p83-2/p6.5 and p83-10_{GFP} system was used and the results were comparable to the ΔGag-protease system.

The focus of this thesis is Gag protein since it is normally the target of dominant CTL responses in these protective HLA-B alleles. Multiple studies have shown that Gag-specific CTL response is associated with slower disease progression and better immune control (Zuniga *et al*., 2006; Kiepiela *et al*., 2007; Sacha *et al*., 2007; Chen *et al*., 2012; Leitman *et al*., 2016). The other advantage of targeting Gag is the potential high fitness cost the virus has to afford when developing an escape mutation, largely because this is the most conserved part of the viral genome encoding the essential structural capsid protein (Miura *et al*., 2009a; Dahirel *et al*., 2011). One approach used in recent T cell vaccine trials was to skew the CD8+ T cell immunodominance pattern toward a broad Gag-specific response (Gray *et al*., 2011; Goulder *et al*., 2016; Leitman *et al*., 2016). In this thesis, I have identified several critical escape mutations that were capable of crippling the virus completely in HLA-B*81:01- and HLA-B*52:01-restricted epitopes. The fitness data were also comparable to the secondary structure model of p24 Gag capsid protein. These findings are important for future vaccine design particularly with a view to reducing the fitness of HIV to the level of highly attenuated viruses that are easy to suppress.

Contrasting with HLA-mediated control of HIV in adults, the HLA-independent nature of paediatic disease non-progression was addressed. Previous studies have shown the association between immune activation and disease progression (Leitman et al., 2016; Muenchhoff et al., 2016). However, it is still possible that the immune control of

HIV in children is still HLA-dependent but it is mediated through a different pathway or through different 'protective HLA allele.' For example, NK activity mediated through HLA recognition plays a role in control of HIV infection in adults (Martin *et al*, 2002; Martin *et al*, 2007), and it is possible that this may be the more important contributor in early paediatric infection when Th1-driven CTL responses are relatively weak. It is also noteworthy that in paediatric infection, the standard to determine whether a patient is a progressor or not varies for the time being according to the changes of WHO treatment guideline (<1 years old, CD4%<25%, CD4 count<350 cells/mm^3 in 2010 versus <5 years old, CD4 count < 500 cells/mm^3; WHO. 2010 and 2013). Although the widely recognised strict standard (Adland *et al*., 2015; Leitman *et al*., 2016) for paediatric slow progressors is >5 years old and the absolute CD4 count is over 750 cells/mm^3, the patients' CD4 counts ranging from 500-750 cells/mm^3 after WHO 2013 guideline, in our perspective, still have the value in controlling HIV infection regardless of different classifications.

In adults, compared to the rare elite controllers and long-term non-progressors, the majority of the HIV-infected population does not have the protective HLA alleles. We attempted to determine via study of HLA-B*40:01, an allele with neutral impact on disease progression in adults, whether there are features of the HLA-B*40:01-restricted response that could be utilized to achieve more successful control of HIV than in the typical HLA-B*40:01-posotive subject. In fact what I observed was that diverse outcomes in subjects expressing HLA-B*40:01 depends not solely on HLA-B*40:01-restricted immune responses. Indeed, even among the individuals with protective HLA alleles, the majority of individuals progress to disease, indicating that many factors in addition to host HLA, play an important role (Goulder and Walker, 2012). Indeed, the more findings the global scientific community gains, the more it is acknowledged that we must consider the multifactorial aspects of immune defence in HIV infection. Current research has shown that the operating mechanisms in elite

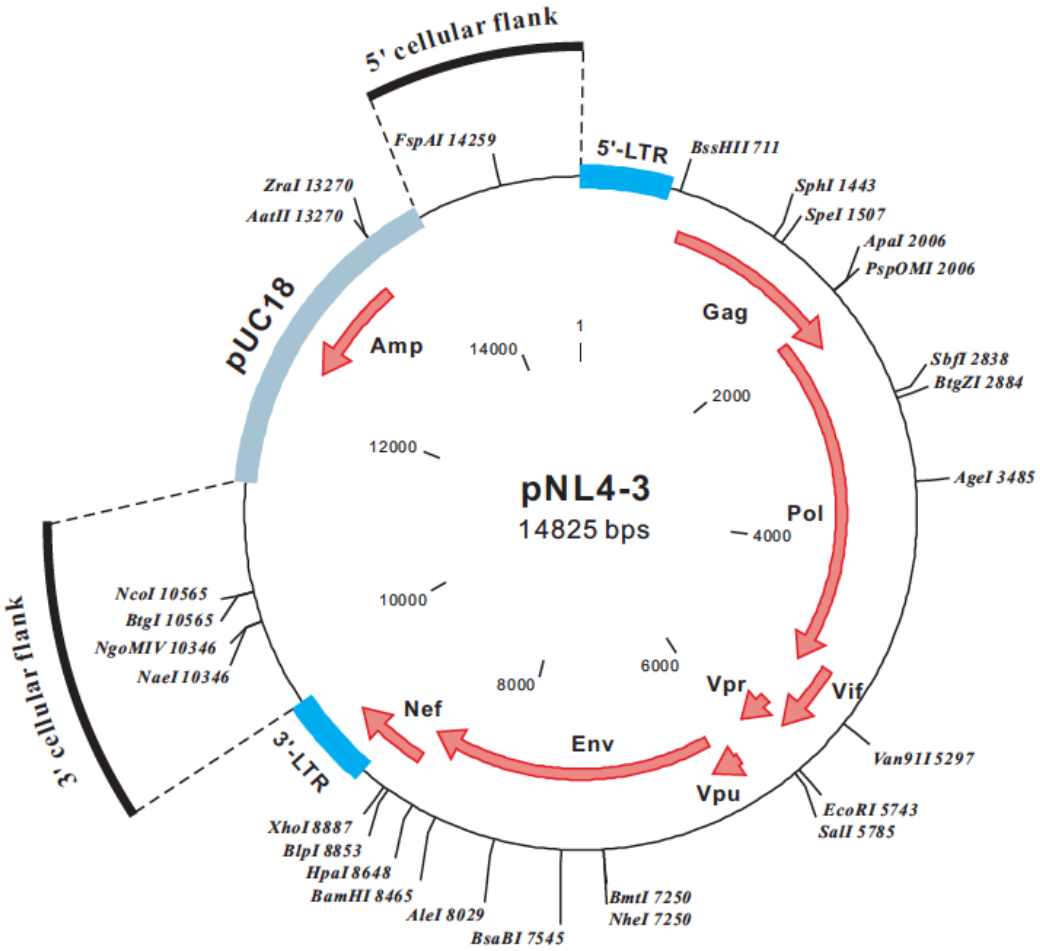
controllers are multiple and heterogeneous, instead of being restricted by a single factor (Walker and Yu, 2013). Hence the elicitation of cellular and humoral responses (e.g. ADCC or bNAbs) is critical and the immune response needs to be cooperative. This thesis has aimed to contribute to the understanding of HLA-mediated immunity and of the impact of altered VRC on an efficient cellular response. T cell vaccines have been shown to operate in an HLA-dependent fashion to improve outcome. For instance, skewing CD8+ T cell responses in a more host-beneficial direction and reordering the CTL hierarchy from Env towards Gag in subjects expressing HLA-B*58:02 was one result of the MRK Ad5 vaccine that improved viral load and CD4 counts in vaccinees expressing the disease-susceptible HLA-B*58:02 who subsequently became infected (Leitman *et al*., 2016).

Thirty years of research has produced a wealth of knowledge and some critical insights into combatting HIV; this thesis has supported the increased understanding for some aspects of this battle. Given the great potential of CTL immunity, developing a therapeutic vaccine that is capable of protecting most of the population will be hugely beneficial in further resisting the HIV/AIDS pandemic. Furthermore, the work of unveiling Gag-specific immune responses and HLA-driven viral adaptation is a vital step closer to the cure of HIV.

APPENDIX

Viral Plasmids (NIH reagents)

NL4-3. GenBank: The plasmid is consisted of the whole HIV-1 B-clade genome replication and an ampicillin resistant pUC18 as the laboratory strain (Adachi *et al.*, 1986).



Appendix Figure 1. Plasmid Map of pNL4-3. Figure is extracted from NIH AIDS Program (https://aidsreagent.org/index.cfm)

HIV-1_{NL4-3} 5´ Clone (p83-2). The p83-2 contains the 5'end of pNL4-3, encoding 5'LTR, Gag, Pol, Vif, and Vpr (Martinez-picado *et al.*, 2006; Pardo *et al.*, 2009).

HIV-1_{NL4-3} 3´ Clone (p83-10). The p83-10_{GFP} contains the 3'end of pNL4-3 and a GFP reporter gene, encoding Tat, Rev, Vpu, Env, 3'LTR, and GFP (Martinez-picado *et al*., 2006; Pardo *et al*., 2009).

p6.5 and p6.5_C_p24. The two plasmids are p-83-2 with C-clade p24 Gag sequences replacement from SapI to ApaI restriction sites (positions 1098 to 2021 in HXB2 sequence). The p24 Gag sequence of p6.5 is patient-derived, and p6.5_C_p24 is constructed by using site-directed mutagenesis at six positions to reverse-mutate the variant to C-clade consensus (Martinez-picado *et al*., 2006; Pardo *et al*., 2009).

HLA-B*27:04 in Taiwanese study

The Taiwanese cohort started as a study of HLA-B*27:04, but in the event the numbers precluded any analysis (*n*=4).

HLA-B*27:04 is a HLA-B*27 subtype prevalent in the Chinese population, compared to HLA-B*27:05 and HLA-B*27:02 dominant in Caucasians (Sheehan, 2010). Studies have shown that different mechanisms were moderated in HLA-B*27 subtypes due to the slight difference of the amino acids in the peptide-binding groove (**Appendix Table 1**). HLA-B*27:04 has similar features to HLA-B*27:02 which possesses non-negative charged amino acids at the critical anchoring site 77. It also accommodates fewer peptides at this position (F, Y and L). The CTL hierarchy of HLA-B*27:04 remains unclear as does its ability to protect against HIV infection. In Ankylosing Spondylitis, HLA-B*27:09 differing only one amino acid from HLA-B*27:05 led to different disease prognosis (Nurzia *et al.*, 2010).

In our cohort, although HLA-B*27:04 tended to have lower CD4 counts and lower viral loads compared to other alleles, due to the small number of individuals expressing this allele and without the longitudinal studies, we were unable to define whether HLA-B*27:04 was protective or not. The enrolment of patients ceased in February 2016, yet ongoing local studies had shown slower disease progression and high risk of acquiring Ankylosing Spondylitis among the HLA-B*27:04-positive chronic-HIV-infected individuals (personal communication). Future studies examining the role HLA-B*27:04 plays in combatting HIV infection are required. It may either assist us in verifying the HLA-B*27:05/B*27:02-mediated mechanisms, or provide evidence that our current hypothesis needs to be revisited.

Appendix Table 1. Characteristics of B*27 subtypes.

B*27 Subtypes	HLA residues					PC Motif	P2	KK10 -Gag	KY9- Pol	VW9- Nef	RL9 (EBV)
	77	80	81	152	211						
B*27:05	D	T	L	V	A	RKHYFMLI	R	✔	✔		✔
B*27:02	N	I	A	-	-	FYILW	R		✔	✔	✔
B*27:04	S	-	-	E	G	FYL	R	N/A	N/A	N/A	✔

Table is based on The HLA Factsbook (Marsh *et al*., 2000).

p6Δ7 in CRF07_BC-infected Taiwanese individuals

Similar to HLA-B*27:04, the small numbers of available CRF07_BC-infected sequences precluded any analysis (*n*=5).

Previous studies showed that a 7-mer amino acid deletion in p6 Gag (478-484, PIDKELY) was observed in most of the Taiwanese patients infected with CRF07_BC (Lin *et al*., 2013). This 7-mer deletion leads to less p24/Pr55 processing and therefore less Gag and protease was produced. This further results in tethering of immature virions on cell membrane, which reduces the viral replicative capacity (Lin *et el*., 2013). However, some studies have reported no influence or an increase in viral load when acquiring p6Δ7 (Song *et al*., 2007; Zhefeng *et al*., 2011).

In our study, all the available sequences from CRF07_BC-infected subjects had p6Δ7 (**Appendix Figure 2**).

	K	Q	D	P	I	D	K	E	L	Y	P	L	T	S	L
T012	R	H	D	-	-	-	-	-	-	-	P	S	F	S	L
T033	K	Q	D	-	-	-	-	-	-	-	P	L	A	S	L
T083	K	Q	E	-	-	-	-	-	-	-	P	L	T	S	L
T088	K	Q	D	-	-	-	-	-	-	-	P	L	I	S	L
T149	K	Q	D	-	-	-	-	-	-	-	P	L	T	S	L

Appendix Figure 2. CRF07_BC p6Δ7 in Taiwanese cohort. p6 Gag sequences of CRF07_BC-infected individuals (*n*=5). The 7-mer deletion is shown in black on the top consensus sequence.

REFERENCE

Accola M, Hoglund S, and Gottlinger H. 1998. A putative alpha-helical structure which overlaps the capsid-p2 boundary in the human immunodeficiency virus type 1 Gag precursor is crucial for viral particle assembly. Journal of Virology 72:2072-2078.

Adachi A, Gendelman H, Koening S, Folks T, Willey R, Rabson A, Martin M. 1986. 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.

Addo M, Yu X, Rathod A, Cohen D, Eldridge R, Strick D, Johnston M, Corcoran C, Wurcel A, Fitzpatrick C, Feeney M, Rodriguez W, Basgoz N, Draenert R, Stone D, Brander C, Goulder P, Rosenberg E, Altfeld M, Walker B. 2003. Comprehensive Epitope Analysis of Human Immunodeficiency Virus Type 1 (HIV-1)-Specific T-Cell Responses Directed against the Entire Expressed HIV-1 Genome Demonstrate Broadly Directed Responses, but No Correlation to Viral Load. Journal of Virology 77:2081-2092.

Adland E, Laker L, Mori L, Hurst J, Jooste P, Goulder P. 2014 Mechanisms of non-pathogenicity in HIV: lessons from paediatric infection. In: 20th International AIDS conference. International AIDS Society, Melbourne, Abstract.

Adland E, Paioni P, Thobakgale C, Laker L, Mori L, Muenchhoff M, Csala A, Clapson M, Flynn J, Novelli V, Hurst J, Naidoo V, Shapiro R, Huang K, Frater J, Prendergast A, Prado J, Ndung'u T, Walker B, Carrington M, Jooste P, Goulder P. 2015. Discordant Impact of HLA on Viral Replicative Capacity and Disease Progression in Pediatric and Adult HIV Infection. PLoS Pathog 11:e1004954.

Abraha A, Nankya I, Gibson R, Demers K, Tebit D, Johnston E, Katzenstein D, Siddiqui A, Herrera C, Fischetti L, Shattock R, Arts E. 2009. CCR5- and CXCR4-Tropic Subtype C Human Immunodeficiency Virus Type 1 Isolates Have a Lower Level of Pathogenic Fitness than Other Dominant Group M Subtypes: Implications for the Epidemic. Journal of Virology 83:5592-5605.

Allen T, Altfeld M, Yu X, O'Sullivan K, Lichterfeld M, Le Gall S, John M, Mothe B, Lee P, Kalife E, Cohen D, Freedberg K, Strick D, Johnston M, Sette A, Rosenberg E, Mallal S, Goulder P, Brander C, Walker B. 2004. Selection, Transmission, and Reversion of an Antigen-Processing Cytotoxic T-Lymphocyte Escape Mutation in Human Immunodeficiency Virus Type 1 Infection. Journal of Virology 78:7069-7078.

Allen T, Yu X, Kalife E, Reyor L, Lichterfeld M, John M, Cheng M, Allgaier R, Mui S, Frahm N, Alter G, Brown N, Johnston M, Rosenberg E, Mallal S, Brander C, Walker B, Altfeld M. 2005. De Novo Generation of Escape Variant-Specific CD8+ T-Cell Responses following Cytotoxic T-Lymphocyte Escape in Chronic Human Immunodeficiency Virus Type 1 Infection. Journal of Virology 79:12952-12960.

Allen T and Altfeld M. 2008. Crippling HIV one mutation at a time. The Journal of Experimental Medicine 205:1003-1007.

Almeida J, Price D, Papagno L, Arkoub Z, Sauce D, Bornstein E, Asher T, Samri A, Schnuriger A, Theodorou I, Costagliola D, Rouzioux C, Agut H, Marcelin A, Douek D, Autran B, Appay V. 2007. Superior control of HIV-1 replication by CD8+T cells is reflected by their avidity, polyfunctionality, and clonal turnover. The Journal of Experimental Medicine 204:2473-2485.

Almeida J, Sauce D, Price D, Papagno L, Shin S, Moris A, Larsen M, Pancino G, Douek D, Autran B, Saez-Cirion A, Appay V. 2009. Antigen sensitivity is a major determinant of CD8+ T-cell polyfunctionality and HIV-suppressive activity. Blood 113:6351-6360.

Altfeld M, Trocha A, Eldridge R, Rosenberg E, Phillips M, Addo M, Sekaly R, Kalams S, Burchett S, McIntosh K, Walker B, Goulder P. 2000. Identification of Dominant Optimal HLA-B60- and HLA-B61-Restricted Cytotoxic T-Lymphocyte (CTL) Epitopes: Rapid Characterization of CTL Responses by Enzyme-Linked Immunospot Assay. Journal of Virology 74:8541-8549.

Altfeld M and Goulder P. 2007. 'Unleashed' natural killers hinder HIV. Nature Genetics 39:708-710.

Ammaranond P, Zaunders J, Satchell C, Van Bockel D, Cooper D, Kelleher A. 2005. A New Variant Cytotoxic T Lymphocyte Escape Mutation in HLA-B27-Positive Individuals Infected with HIV Type 1. AIDS Research and Human Retroviruses 21:395-397.

Antoni G, Guergnon J, Meaudre C, Samri A, Boufassa F, Goujard C, Lambotte O, Autran B, Rouzioux C, Costagliola D, Meyer L, Theodorou I. 2013. MHC-driven HIV-1 control on the long run is not systematically determined at early times post-HIV-1 infection. AIDS 27:1707-1716.

Apps R, Qi Y, Carlson J, Chen H, Gao X, Thomas R, Yuki Y, Del Prete G, Goulder P, Brummw Z, Brumme C, John M, Mallal S, Nelson G, Bosch R, Heckerman D, Stein J, Soderberg K, Moody M, Denny T, Zheng X, Fang J, Moffett A, Lifson J, Goedert J, Buchbinder S, Kirk G, Fellay J, McLaren P, Deeks S, Pereyra F, Walker B, Michael N, Weintrob A, Wolinsky S, Liao W, Carrington M. 2013. Influence of HLA-C expression level on HIV control. Science 340(6128):87-91.

Baggaley R, Boily M, White R, Alary M. 2006. Risk of HIV-1 transmission for parenteral exposure and blood transfusion: a systematic review and meta-analysis. AIDS 20:805-812.

Bailey J, Brennan T, O'Connell K, Siliciano R, Blankson J. 2009. Evidence of CD8+ T-Cell-Mediated Selective Pressure on Human Immunodeficiency Virus Type 1 nef in HLA-B*57+ Elite Suppressors. Journal of Virology 83:88-97.

Barre-Sinoussi F, Chermann J, Rey F, Nugeyre M, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vezinet-Brun F, Rouzioux C, Rozenbaum W, Montagnier L. 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science 220:868-871.

Barreto-de-Souza V, Arakelyan A, Margolis L, Vanpouille C. 2014. HIV-1 Vaginal Transmission: Cell-Free or Cell-Associated Virus?. American Journal of Reproductive Immunology 71:589-599.

Battivelli E, Migraine J, Lecossier D, Yeni P, Clavel F, Hance A. 2011. Gag Cytotoxic T Lymphocyte Escape Mutations Can Increase Sensitivity of HIV-1 to Human TRIM5 , Linking Intrinsic and Acquired Immunity. Journal of Virology 85:11846-11854.

Baum L. 2010. Role of Humoral Immunity in Host Defense Against HIV. Current HIV/AIDS Reports 7:11-18.

Berthet-Colominas C. 1999. Head-to-tail dimers and interdomain flexibility revealed by the crystal structure of HIV-1 capsid protein (p24) complexed with a monoclonal antibody Fab. The EMBO Journal 18:1124-1136.

Betts M, Nason M, West S, De Rosa S, Migueles S, Abraham J, Lederman M, Benito J, Goepfert P, Connors M, Roederer M, Koup R. 2006. HIV nonprogressors preferentially maintain highly functional HIV-specific CD8+ T cells. Blood 107:4781-4789.

Bevan M. 2004. Helping the CD8+ T-cell Response. Nature Reviews Immunology 4:595-602.

Bhattacharya A, Alam S, Fricke T, Zadrozny K, Sedzicki J, Taylor A, Demeler B, Pornillos O, Ganser-Pornillos B, Diaz-Griffero F, Ivanov D, Yeager M. 2014. Structural basis of HIV-1 capsid recognition by PF74 and CPSF6. Proceedings of the National Academy of Sciences 111:18625-18630.

Bjorkman P, Saper M, Samraoui B, Bennett W, Strominger J, Wiley D. 1987. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. Nature 329:512-518.

Bjorkman P, Saper M, Samraoui B, Bennett W, Strominger J, Wiley D. 1987. Structure of the human class I histocompatibility antigen, HLA-A2. Nature 329:506-512.

Blais M, Dong T, Rowland-Jones S. 2011. HLA-C as a mediator of natural killer and T-cell activation: spectator or key player?. Immunology 133:1-7.

Blanche S, Newell M, Mayaux M, Dunn D, Teglas J, Rouzioux C, Peckham C. 1997. Morbidity and Mortality in European Children Vertically Infected by HIV-1. Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology 14:442-450.

Boily M, Baggaley R, Wang L, Masse B, White R, Hayes R, Alary M. 2009. Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies. The Lancet Infectious Diseases 9:118-129.

Borrow P, Lewicki H, Hahn B, Shaw G, Oldstone M. 1994. 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):6103-6110.

Borthwick N, Ahmed T, Ondondo B, Hayes P, Rose A, Ebrahimsa U, Hayton E, Black A, Bridgeman A, Rosario M, Hill A, Berrie E, Moyle S, Frahm N, Cox J, Colloca S, Nicosia A, Gilmour J, McMichael A, Dorrell L, Hanke T. 2014. Vaccine-elicited Human T Cells Recognizing Conserved Protein Regions Inhibit HIV-1. Molecular Therapy 22:464-475.

Borthwick N, Lin Z, Akahoshi T, Llano A, Silva-Arrieta S, Ahmed T, Dorrell L, Brander C, Murakoshi H, Takiguchi M, Hanke T. 2017. Novel, in-natural-infection subdominant HIV-1 CD8+ T-cell epitopes revealed in human recipients of conserved-region T-cell vaccines. PLOS ONE 12:e0176418.

Boyington J and Sun P. 2002. A structural perspective on MHC class I recognition by killer cell immunoglobulin-like receptors. Molecular Immunology 38:1007-1021.

Brander C. and Walker B. 1996. The HLA class I restricted CTL response in HIV infection: systematic identification of optimal epitopes. In B. Korber, C. Brander, B. Walker, R. Koup, J. Moore, B. Haynes, and G. Meyers (ed.), HIV molecular immunology database. Los Alamos National Lab- oratory: Theoretical Biology and Biophysics, Los Alamos, N.Mex. http://hiv-web.lanl.gov/immunology/index.html. Accessed on 12 December 2016.

Brander C, Yang O, Jones N, Lee Y, Goulder P, Johnson R, Trocha A, Colbert D, Hay C, Buchbinder S, Bergmann C, Zweerink H, Wolinsky S, Blattner W, Kalams S, Walker B. 1999. Efficient Processing of the Immunodominant, HLA-A*0201- Restricted Human Immunodeficiency Virus Type 1 Cytotoxic T-Lymphocyte Epitope despite Multiple Variations in the Epitope Flanking Sequences. Journal of Virology 73:10191–10198.

Brener J, Gall A, Batorsky R, Riddell L, Buus S, Leitman E, Kellam P, Allen T, Goulder P, Matthews P. 2015. Disease progression despite protective HLA expression in an HIV-infected transmission pair. Retrovirology 12.

Briggs J, Simon M, Gross I, Kräusslich H, Fuller S, Vogt V, Johnson M. 2004. The stoichiometry of Gag protein in HIV-1. Nature Structural & Molecular Biology 11:672-675.

Brocca-Cofano E, McKinnon K, Demberg T, Venzon D, Hidajat R, Xiao P, Daltabuit-Test M, Patterson L, Robert-Guroff M. 2011. Vaccine-elicited SIV and HIV envelope-specific IgA and IgG memory B cells in rhesus macaque peripheral blood correlate with functional antibody responses and reduced viremia. Vaccine 29:3310-3319.

Brockman M, Schneidewind A, Lahaie M, Schmidt A, Miura T, DeSouza I, Ryvkin F, Derdeyn C, Allen S, Hunter E, Mulenga J, Goepfert P, Walker B, Allen T. 2007. Escape and Compensation from Early HLA-B57-Mediated Cytotoxic T-Lymphocyte Pressure on Human Immunodeficiency Virus Type 1 Gag Alter Capsid Interactions with Cyclophilin A. Journal of Virology 81:12608-12618.

Brockman M, Brumme Z, Brumme C, Miura T, Sela J, Rosato P, Kadie C, Carlson J, Markle T, Streeck H, Kelleher A, Markowitz M, Jessen H, Rosenberg E, Altfeld M, Harrigan P, Heckerman D, Walker B, Allen T. 2010. Early Selection in Gag by Protective HLA Alleles Contributes to

Reduced HIV-1 Replication Capacity That May Be Largely Compensated for in Chronic Infection. Journal of Virology 84:11937-11949.

Bronke C, Almeida C, McKinnon E, Roberts S, Keane N, Chopra A, Carlson J, Heckerman D, Mallal S, John M. 2013. HIV escape mutations occur preferentially at HLA-binding sites of CD8 T-cell epitopes. AIDS 27:899-905.

Brooks J, Murray R, Thomas W, Kurilla M, Rockinson A. 1993. Different HLA-B27 subtypes present the same immunodominant Epstein-Barr virus peptide. Journal of Experimental Medicine 178:879-887.

Brumme Z, Brumme C, Carlson J, Streeck H, John M, Eichbaum Q, Block B, Baker B, Kadie C, Markowitz M, Jessen H, Kelleher A, Rosenberg E, Kaldor J, Yuki Y, Carrington M, Allen T, Mallal S, Altfeld M, Heckerman D, Walker B. 2008. Marked Epitope- and Allele-Specific Differences in Rates of Mutation in Human Immunodeficiency Type 1 (HIV-1) Gag, Pol, and Nef Cytotoxic T-Lymphocyte Epitopes in Acute/Early HIV-1 Infection. Journal of Virology 82:9216-9227.

Buchbinder S, Mehrotra D, Duerr A, Fitzgerald D, Mogg R, Li D, Gilbert P, Lama J, Marmor M, del Rio C, McElrath M, Casimiro D, Gottesdiener K, Chodakewitz J, Corey L, Robertson M. 2008. Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. The Lancet 372:1881-1893.

Byeon I, Meng X, Jung J, Zhao G, Yang R, Ahn J, Shi J, Concel J, Aiken C, Zhang P, Gronenborn A. 2009. Structural Convergence between Cryo-EM and NMR Reveals Intersubunit Interactions Critical for HIV-1 Capsid Function. Cell 139:780-790.

Campbell E and Hope T. 2015. HIV-1 capsid: the multifaceted key player in HIV-1 infection. Nature Reviews Microbiology 13:471-483.

Cano P, Klitz W, Mack SJ, Maiers M, Marsh SG, Noreen H, Reed EF, Senitzer D, Setterholm M, Smith A, Fernández-Viña M. 2007. Human Immunology 68(5):392-417.

Carlson J, Brumme C, Martin E, Listgarten J, Brockman M, Le A, Chui C, Cotton L, Knapp D, Riddler S, Haubrich R, Nelson G, Pfeifer N, DeZiel C, Heckerman D, Apps R, Carrington M, Mallal S, Harrigan P, John M, Brumme Z. 2012a. Correlates of Protective Cellular Immunity Revealed by Analysis of Population-Level Immune Escape Pathways in HIV-1. Journal of Virology 86:13202-13216.

Carlson J, Listgarten J, Pfeifer N, Tan V, Kadie C, Walker B, Ndung'u T, Shapiro R, Frater J, Brumme Z, Goulder P, Heckerman D. 2012b. Widespread Impact of HLA Restriction on Immune Control and Escape Pathways of HIV-1. Journal of Virology 86:5230-5243.

Carlson J, Schaefer M, Monaco D, Batorsky R, Claiborne D, Prince J, Deymier M, Ende Z, Klatt N, DeZiel C, Lin T, Peng J, Seese A, Shapiro R, Frater J, Ndung'u T, Tang J, Goepfert P, Gilmour J, Price M, Kilembe W, Heckerman D, Goulder P, Allen T, Allen S, Hunter E. 2014. Selection bias

at the heterosexual HIV-1 transmission bottleneck. Science 345:1254031-1254031.

Carlson J, Le A, Shahid A, Brumme Z. 2015. HIV-1 adaptation to HLA: a window into virus–host immune interactions. Trends in Microbiology 23:212-224.

Carlson J, Du V, Pfeifer N, Bansal A, Tan V, Power K, Brumme C, Kreimer A, DeZiel C, Fusi N, Schaefer M, Brockman M, Gilmour J, Price M, Kilembe W, Haubrich R, John M, Mallal S, Shapiro R, Frater J, Harrigan P, Ndung'u T, Allen S, Heckerman D, Sidney J, Allen T, Goulder P, Brumme Z, Hunter E, Goepfert P. 2016. Impact of pre-adapted HIV transmission. Nature Medicine 22:606-613.

Carr J, Salminen M, Johansson B, Sönnerborg A, Ayehunie S, Gotte D, Leinikki P, Burke D, Mccutchan F. 1996. Full-Length Sequence of an Ethiopian Human Immunodeficiency Virus Type 1 (HIV-1) Isolate of Genetic Subtype C. AIDS Research and Human Retroviruses 12:1329-1339.

Carr J. 2006. Viral diversity as a challenge to HIV-1 vaccine development. Current Opinion in HIV and AIDS 1:294-300.

Carrington M, Nelson G, Martin M, Kissner T, Vlahov D, Goedert J, Kaslow R, Buchbinder S, Hoots K, O'Brien S. 1999. HLA and HIV-1: Heterozygote Advantage and B*35-Cw*04 Disadvantage. Science 283:1748-1752.

Carrington M, Martin M, van Bergen J. 2008. KIR-HLA intercourse in HIV disease. Trends in Microbiology 16:620-627.

Cenderello G and Maria A. 2015. Discordant responses to cART in HIV-1 patients in the era of high potency antiretroviral drugs: clinical evaluation, classification, management prospects. Expert Review of Anti-infective Therapy 14:29-40.

CDC, R.O.C. Taiwan. 2016. Statistics of HIV/AIDS. In http://www.cdc.gov.tw/english/info.aspx?treeid=00ED75D6C887BB27&nowtreeid=2FA0168FF47E02BE&tid=C15B9E0B7383CA7E

Centers for Disease Control (CDC). 1981. Pneumocystis pneumonia—Los Angeles. Morbidity and Mortality Weekly Report. 30(21):250–252.

Centers for Disease Control (CDC). 1982. Update on Kaposi's sarcoma and opportunistic infections in previously healthy persons—United States". Morbidity and Mortality Weekly Report 31(22):294, 300–201.

Centers for Disease Control (CDC). 1982. Pneumocystis carinii pneumonia among persons with hemophilia A. Morbidity and Mortality Weekly Report 31(27): 365–267.

Centers for Disease Control (CDC). 1982. Opportunistic infections and Kaposi's sarcoma among Haitians in the United States. Morbidity and Mortality Weekly Report 31(26): 353–354, 360–361.

Chan D and Kim P. 1998. HIV Entry and Its Inhibition. Cell 93:681-684.

Chevalier M, Julg B, Pyo A, Flanders M, Ranasinghe S, Soghoian D, Kwon D,

Rychert J, Lian J, Muller M, Cutler S, McAndrew E, Jessen H, Pereyra F, Rosenberg E, Altfeld M, Walker B, Streeck H. 2010. HIV-1-Specific Interleukin-21+ CD4+ T Cell Responses Contribute to Durable Viral Control through the Modulation of HIV-Specific CD8+ T Cell Function. Journal of Virology 85:733-741.

Chen Z, Craiu A, Shen L, Kuroda M, Iroku U, Watkins D, Voss G, Letvin N. 2000. Simian Immunodeficiency Virus Evades a Dominant Epitope-Specific Cytotoxic T Lymphocyte Response Through a Mutation Resulting in the Accelerated Dissociation of Viral Peptide and MHC Class I. The Journal of Immunology 164:6474-6479.

Chen H, Ndhlovu Z, Liu D, Porter L, Fang J, Darko S, Brockman M, Miura T, Brumme Z, Schneidewind A, Piechocka-Trocha A, Cesa K, Sela J, Cung T, Toth I, Pereyra F, Yu X, Douek D, Kaufmann D, Allen T, Walker B. 2012a. TCR clonotypes modulate the protective effect of HLA class I molecules in HIV-1 infection. Nature Immunology 13:691-700.

Chen Y, Lee C, Chen M, Chuang S, Liu H, Wong W, Lin Y, Tsai H, Wang J, Chen Y. 2012b. Molecular epidemiology of HIV-1 infection in Taiwan from 2005 to 2008: Further spread of CRF07-BC and emergence of CRF07-BC/subtype B dual infection. Journal of Acquired Immune Deficiency Syndromes 59(5):438-446.

Chen S, Luan H, Li L, Zeng X, Wang T, Li Y, Yuan H. 2016. Relationship of HLA-B*51 and HLA-B*52 alleles and TNF-α-308A/G polymorphism with susceptibility to Takayasu arteritis: a meta-analysis. Clinical Rheumatology 36:173-181.

Chopera D, Mlotshwa M, Woodman Z, Mlisana K, de Assis Rosa D, Martin D, Abdool Karim S, Gray C, Williamson C. 2011. Virological and Immunological Factors Associated with HIV-1 Differential Disease Progression in HLA-B*58:01-Positive Individuals. Journal of Virology 85:7070-7080.

Chikata T, Carlson J, Tamura Y, Borghan M, Naruto T, Hashimoto M, Murakoshi H, Le A, Mallal S, John M, Gatanaga H, Oka S, Brumme Z, Takiguchi M. 2014. Host-Specific Adaptation of HIV-1 Subtype B in the Japanese Population. Journal of Virology 88:4764-4775.

Choe H, Farzan M, Sun Y, Sullivan N, Rollins B, Ponath P, Wu L, Mackay C, LaRosa G, Newman W, Gerard N, Gerard C, Sodroski J. 1996. The β-Chemokine Receptors CCR3 and CCR5 Facilitate Infection by Primary HIV-1 Isolates. Cell 85:1135-1148.

Claiborne D, Prince J, Scully E, Macharia G, Micci L, Lawson B, Kopycinski J, Deymier M, Vanderford T, Nganou-Makamdop K, Ende Z, Brooks K, Tang J, Yu T, Lakhi S, Kilembe W, Silvestri G, Douek D, Goepfert P, Price M, Allen S, Paiardini M, Altfeld M, Gilmour J, Hunter E. 2015. Replicative fitness of transmitted HIV-1 drives acute immune activation, proviral load in memory CD4+T cells, and disease progression. Proceedings of the National Academy of Sciences.112(12):E1480-E1489.

Clavel F, Guetard D, Brun-Vezinet F, Chamaret S, Rey M, Santos-Ferreira M,

Laurent A, Dauguet C, Katlama C, Rouzioux C, et al. 1986. Isolation of a new human retrovirus from West African patients with AIDS. Science 233:343-346.

Coffin J, Haase A, Levy JA, Montagnier L, Oroszlan S, Teich N, Temin H, Toyoshima K, Varmus H, Vogt P. 1986. What to call the AIDS virus?. Nature. 321(6065):10.

Cohen G, Gandhi R, Davis D, Mandelboim O, Chen B, Strominger J, Baltimore D. 1999. The Selective Downregulation of Class I Major Histocompatibility Complex Proteins by HIV-1 Protects HIV-Infected Cells from NK Cells. Immunity 10:661-671.

Cohen M, Gay C, Busch M, Hecht F. 2010. The Detection of Acute HIV Infection. Journal of Infectious Diseases 202: S270-S277.

Connor R, Sheridan K, Ceradini D, Choe S, Landau N. 1997. Change in Coreceptor Use Correlates with Disease Progression in HIV-1–Infected Individuals. The Journal of Experimental Medicine 185:621-628.

Cotton M, Marais B, Andersson M, Eley B, Rabie H, Slogrove A, Dramowski A, Schaaf H, Mehtar S. 2012. Commentary: minimizing the risk of non-vertical, non-sexual HIV infection in children – beyond mother to child transmission. Journal of the International AIDS Society 15:173-177.

Cotton M, Violari A, Otwombe K, Panchia R, Dobbels E, Rabie H, Josipovic D, Liberty A, Lazarus E, Innes S, van Rensburg A, Pelser W, Truter H, Madhi S, Handelsman E, Jean-Philippe P, McIntyre J, Gibb D, Babiker A. 2013. Early time-limited antiretroviral therapy versus deferred therapy in South African infants infected with HIV: results from the children with HIV early antiretroviral (CHER) randomised trial. The Lancet 382:1555-1563.

Craven R, Leure-dePree A, Weldon R, Wills J. 1995. Genetic analysis of the major homology region of the Rous sarcoma virus Gag protein. Journal of Virology 69(7):4213-4227.

Crawford H, Prado J, Leslie A, Hue S, Honeyborne I, Reddy S, van der Stok M, Mncube Z, Brander C, Rousseau C, Mullins J, Kaslow R, Goepfert P, Allen S, Hunter E, Mulenga J, Kiepiela P, Walker B, Goulder P. 2007. Compensatory Mutation Partially Restores Fitness and Delays Reversion of Escape Mutation within the Immunodominant HLA-B*5703-Restricted Gag Epitope in Chronic Human Immunodeficiency Virus Type 1 Infection. Journal of Virology 81:8346-8351.

Crawford H, Matthews P, Schaefer M, Carlson J, Leslie A, Kilembe W, Allen S, Ndung'u T, Heckerman D, Hunter E, Goulder P. 2010. The Hypervariable HIV-1 Capsid Protein Residues Comprise HLA-Driven CD8+ T-Cell Escape Mutations and Covarying HLA-Independent Polymorphisms. Journal of Virology 85:1384-1390.

Cullen B. 1998. Retroviruses as Model Systems for the Study of Nuclear RNA Export Pathways. Virology 249:203-210.

Currier J, Visawapoka U, Tovanabutra S, Mason C, Birx D, McCutchan F, Cox J.

2006. CTL epitope distribution patterns in the Gag and Nef proteins of HIV-1 from subtype A infected subjects in Kenya: Use of multiple peptide sets increases the detectable breadth of the CTL response. BMC Immunology 7:8

Dahirel V, Shekhar K, Pereyra F, Miura T, Artyomov M, Talsania S, Allen T, Altfeld M, Carrington M, Irvine D, Walker B, Chakraborty A. 2011. Coordinate linkage of HIV evolution reveals regions of immunological vulnerability. Proceedings of the National Academy of Sciences 108:11530-11535.

Davis Z, Cogswell A, Scott H, Mertsching A, Boucau J, Wambua D, Le Gall S, Planelles V, Campbell K, Barker E. 2016. A Conserved HIV-1-Derived Peptide Presented by HLA-E Renders Infected T-cells Highly Susceptible to Attack by NKG2A/CD94-Bearing Natural Killer Cells. PLOS Pathogens 12:e1005421.

de Medeiros R, Valverde-Villegas J, Junqueira D, Gräf T, Lindenau J, de Mello M, Vianna P, Almeida S, Chies J. 2016. Rapid and Slow Progressors Show Increased IL-6 and IL-10 Levels in the Pre-AIDS Stage of HIV Infection. PLOS ONE 11:e0156163.

Deeks S, Schweighardt B, Wrin T, Galovich J, Hoh R, Sinclair E, Hunt P, McCune J, Martin J, Petropoulos C, Hecht F. 2006. Neutralizing Antibody Responses against Autologous and Heterologous Viruses in Acute versus Chronic Human Immunodeficiency Virus (HIV) Infection: Evidence for a Constraint on the Ability of HIV To Completely Evade Neutralizing Antibody Responses. Journal of Virology 80:6155-6164.

Deeks S and Walker B. 2007. Human Immunodeficiency Virus Controllers: Mechanisms of Durable Virus Control in the Absence of Antiretroviral Therapy. Immunity 27:406-416.

Deeks S. 2011. HIV infection, inflammation, immunosenescence, and aging. Annual Review of Medicine. 62:141–155.

Deeks S, The International AIDS Society Scientific Working Group on HIV Cure. 2012. Towards an HIV cure: a global scientific strategy. *Nature reviews Immunology* 12(8):607-614.

Deymier M, Ende Z, Fenton-May A, Dilernia D, Kilembe W, Allen S, Borrow P, Hunter E. 2015. Heterosexual Transmission of Subtype C HIV-1 Selects Consensus-Like Variants without Increased Replicative Capacity or Interferon-α Resistance. PLOS Pathogens 11:e1005154.

Dick T, Bangia N, Peaper D, Cresswell P. 2002. Disulfide Bond Isomerization and the Assembly of MHC Class I-Peptide Complexes. Immunity 16:87-98.

Dinoso J, Kim S, Wiegand A, Palmer S, Gange S, Cranmer L, O'Shea A, Callender M, Spivak A, Brennan T, Kearney M, Proschan M, Mican J, Rehm C, Coffin J, Mellors J, Siliciano R, Maldarelli F. 2009. Treatment intensification does not reduce residual HIV-1 viremia in patients on highly active antiretroviral therapy. Proceedings of the National Academy of Sciences 106:9403-9408.

Doherty P and Zinkernagel R. 1975. A Biological Role For The Major

Histocompatibility Antigens. The Lancet 305:1406-1409.

Donahue R, Johnson M, Zon L, Clark S, Groopman J. 1987. Suppression of in vitro haematopoiesis following human immunodeficiency virus infection. Nature 326:200-203.

Draenert R, Le Gall S, Pfafferott K, Leslie A, Chetty P, Brander C, Holmes E, Chang S, Feeney M, Addo M, Ruiz L, Ramduth D, Jeena P, Altfeld M, Thomas S, Tang Y, Verrill C, Dixon C, Prado J, Kiepiela P, Martinez-Picado J, Walker B, Goulder P. 2004. Immune Selection for Altered Antigen Processing Leads to Cytotoxic T Lymphocyte Escape in Chronic HIV-1 Infection. The Journal of Experimental Medicine 199:905-915.

Duda A, Lee-Turner L, Fox J, Robinson N, Dustan S, Kaye S, Fryer H, Carrington M, McClure M, Mclean A, Fidler S, Weber J, Phillips R, Frater A. 2008. HLA-Associated Clinical Progression Correlates with Epitope Reversion Rates in Early Human Immunodeficiency Virus Infection. Journal of Virology 83:1228-1239.

Duh E, Maury W, Folks T, Fauci A, Rabson A. 1989. Tumor necrosis factor alpha activates human immunodeficiency virus type 1 through induction of nuclear factor binding to the NF-kappa B sites in the long terminal repeat. Proceedings of the National Academy of Sciences 86:5974-5978.

Dunham R, Pagliardini P, Gordon S, Sumpter B, Engram J, Moanna A, Paiardini M, Mandl J, Lawson B, Garg S, McClure H, Xu Y, Ibegbu C, Easley K, Katz N, Pandrea I, Apetrei C, Sodora D, Staprans S, Feinberg M, Silvestri G. 2006. The AIDS resistance of naturally SIV-infected sooty mangabeys is independent of cellular immunity to the virus. Blood 108:209-217.

Dunn D. HIV Paediatric Prognostic Markers Collaborative Study Group. 2003. Short-term risk of disease progression in HIV-1-infected children receiving no antiretroviral therapy or zidovudine monotherapy: a meta-analysis. The Lancet 362:1605-1611.

Elahi S, Dinges W, Lejarcegui N, Laing K, Collier A, Koelle D, McElrath M, Horton H. 2011. Protective HIV-specific CD8+ T cells evade Treg cell suppression. Nature Medicine 17:989-995.

Ellery P, Tippett E, Chiu Y, Paukovics G, Cameron P, Solomon A, Lewin S, Gorry P, Jaworowski A, Greene W, Sonza S, Crowe S. 2007. The CD16+ Monocyte Subset Is More Permissive to Infection and Preferentially Harbors HIV-1 In Vivo. The Journal of Immunology 178:6581-6589.

Emsley P and Cowtan K. 2004. Coot: model-building tools for molecular graphics. Acta Crystallographica Section D Biological Crystallography 60:2126-2132.

Emu B, Sinclair E, Hatano H, Ferre A, Shacklett B, Martin J, McCune J, Deeks S. 2008. HLA Class I-Restricted T-Cell Responses May Contribute to the Control of Human Immunodeficiency Virus Infection, but Such Responses Are Not Always Necessary for Long-Term Virus Control. Journal of Virology 82:5398-5407.

Epstein F, Epstein F, Greene W. 1991. The Molecular Biology of Human

Immunodeficiency Virus Type 1 Infection. New England Journal of Medicine 324:308-317.

Espíndola M, Soares L, Galvão-Lima L, Zambuzi F, Cacemiro M, Brauer V, Frantz F. 2016. HIV Infection: Focus on the Innate Immune Cells. Immunologic Research 64(5):1118-1132.

Fabio G, Scorza R, Lazzarin A, Marchini M, Zarantonello M, D'arminio A, Marchisio P, Plebani A, Luzzati R, Costigliola P. 1992. HLA-associated susceptibility to HIV-1 infection. Clinical & Experimental Immunology 87:20-23.

Falk K, Rötzschke O, Stevanović S, Jung G, Rammensee H. 1991. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. Nature 351:290-296.

Fantuzzi L, Purificato C, Donato K, Belardelli F, Gessani S. 2004. Human Immunodeficiency Virus Type 1 gp120 Induces Abnormal Maturation and Functional Alterations of Dendritic Cells: a Novel Mechanism for AIDS Pathogenesis. Journal of Virology 78:9763-9772.

Farahani M, Mulinder H, Farahani A, Marlink R. 2016. Prevalence and distribution of non-AIDS causes of death among HIV-infected individuals receiving antiretroviral therapy: a systematic review and meta-analysis. International Journal of STD & AIDS 095646241663242.

Faria N, Rambaut A, Suchard M, Baele G, Bedford T, Ward M, Tatem A, Sousa J, Arinaminpathy N, Pepin J, Posada D, Peeters M, Pybus O, Lemey P. 2014. The early spread and epidemic ignition of HIV-1 in human populations. Science 346:56-61.

Feeney M, Tang Y, Roosevelt K, Leslie A, McIntosh K, Karthas N, Walker B, Goulder P. 2004. Immune Escape Precedes Breakthrough Human Immunodeficiency Virus Type 1 Viremia and Broadening of the Cytotoxic T-Lymphocyte Response in an HLA-B27-Positive Long-Term-Nonprogressing Child. Journal of Virology 78:8927-8930.

Feeney M, Tang Y, Pfafferott K, Roosevelt K, Draenert R, Trocha A, Yu X, Verrill C, Allen T, Moore C, Mallal S, Burchett S, McIntosh K, Pelton S, St. John M, Hazra R, Klenerman P, Altfeld M, Walker B, Goulder P. 2005. HIV-1 Viral Escape in Infancy Followed by Emergence of a Variant-Specific CTL Response. The Journal of Immunology 174:7524-7530.

Fellay J, Shianna K, Ge D, Colombo S, Ledergerber B, Weale M, Zhang K, Gumbs C, Castagna A, Cossarizza A, Cozzi-Lepri A, De Luca A, Easterbrook P, Francioli P, Mallal S, Martinez-Picado J, Miro J, Obel N, Smith J, Wyniger J, Descombes P, Antonarakis S, Letvin N, McMichael A, Haynes B, Telenti A, Goldstein D. 2007. A Whole-Genome Association Study of Major Determinants for Host Control of HIV-1. Science 317:944-947.

Feinberg M and Greene W. 1992. Molecular insights into human immunodeficiency virus type 1 pathogenesis. Current Biology 2:532.

Feng Y, Broder C, Kennedy P, Berger E. 1996. HIV-1 Entry Cofactor: Functional

cDNA Cloning of a Seven-Transmembrane, G Protein-Coupled Receptor. Science 272:872-877.

Ferguson A, Mann J, Omarjee S, Ndung'u T, Walker B, Chakraborty A. 2013.

Translating HIV Sequences into Quantitative Fitness Landscapes Predicts Viral Vulnerabilities for Rational Immunogen Design. Immunity 38:606-617.

Fernandez C, Stratov I, De Rose R, Walsh K, Dale C, Smith M, Agy M, Hu S, Krebs K, Watkins D, O'Connor D, Davenport M, Kent S. 2005. Rapid Viral Escape at an Immunodominant Simian-Human Immunodeficiency Virus Cytotoxic T-Lymphocyte Epitope Exacts a Dramatic Fitness Cost. Journal of Virology 79:5721-5731.

Flores-Villanueva P, Yunis E, Delgado J, Vittinghoff E, Buchbinder S, Leung J, Uglialoro A, Clavijo O, Rosenberg E, Kalams S, Braun J, Boswell S, Walker B, Goldfeld A. 2001. Control of HIV-1 viremia and protection from AIDS are associated with HLA-Bw4 homozygosity. Proceedings of the National Academy of Sciences 98:5140-5145.

Fischer W, Perkins S, Theiler J, Bhattacharya T, Yusim K, Funkhouser R, Kuiken C, Haynes B, Letvin N, Walker B, Hahn B, Korber B. 2007. Polyvalent vaccines for optimal coverage of potential T-cell epitopes in global HIV-1 variants. Nature Medicine 13:100-106.

Fischer W, Liao H, Haynes B, Letvin N, Korber B. 2008. Coping with Viral Diversity in HIV Vaccine Design: A Response to Nickle et al. PLoS Computational Biology 4:e15.

Foley B. 2000. An overview of the molecular phylogeny of lentiviruses. HIV sequence compendium. 2000:35-43.

Frahm N, Yusim K, Hraber P, Adams S, Marincola F, Korber B, Brander C. 2004. Extensive HLA class I binding promiscuity among EBV- and HIV- derived cytotoxic T lymphocyte epitopes. Human Immunology 65:S34.

Frahm N, Yusim K, Suscovich T, Adams S, Sidney J, Hraber P, Hewitt H, Linde C, Kavanagh D, Woodberry T, Henry L, Faircloth K, Listgarten J, Kadie C, Jojic N, Sango K, Brown N, Pae E, Zaman M, Bihl F, Khatri A, John M, Mallal S, Marincola F, Walker B, Sette A, Heckerman D, Korber B, Brander C. 2007. Extensive HLA class I allele promiscuity among viral CTL epitopes. European Journal of Immunology 37:2419-2433.

Fraser C, Lythgoe K, Leventhal G, Shirreff G, Hollingsworth T, Alizon S, Bonhoeffer S. 2014. Virulence and Pathogenesis of HIV-1 Infection: An Evolutionary Perspective. Science 343:1243727.

Frange P, Faye A, Avettand-Fenoël V, Bellaton E, Descamps D, Angin M, David A, Caillat-Zucman S, Peytavin G, Dollfus C, Le Chenadec J, Warszawski J, Rouzioux C, Sáez-Cirión A. 2016. HIV-1 virological remission lasting more than 12 years after interruption of early antiretroviral therapy in a perinatally infected teenager enrolled in the French ANRS EPF-CO10 paediatric cohort: a case report. The Lancet HIV 3:e49-e54.

Friedrich T, Dodds E, Yant L, Vojnov L, Rudersdorf R, Cullen C, Evans D,

Desrosiers R, Mothé B, Sidney J, Sette A, Kunstman K, Wolinsky S, Piatak M, Lifson J, Hughes A, Wilson N, O'Connor D, Watkins D. 2004. Reversion of CTL escape–variant immunodeficiency viruses in vivo. Nature Medicine 10:275-281.

Furutsuki T, Hosoya N, Kawana-Tachikawa A, Tomizawa M, Odawara T, Goto M, Kitamura Y, Nakamura T, Kelleher A, Cooper D, Iwamoto A. 2004. Frequent Transmission of Cytotoxic-T-Lymphocyte Escape Mutants of Human Immunodeficiency Virus Type 1 in the Highly HLA-A24-Positive Japanese Population. Journal of Virology 78:8437-8445.

Gaardbo J, Hartling H, Gerstoft J, Nielsen S. 2012. Thirty Years with HIV Infection—Nonprogression Is Still Puzzling: Lessons to Be Learned from Controllers and Long-Term Nonprogressors. AIDS Research and Treatment 2012:1-14.

Gail M, Tan W, Pee D, Goedert J. 1997. Survival After AIDS Diagnosis in a Cohort of Hemophilia Patients. Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology 15:363-369.

Gall A, Ferns B, Morris C, Watson S, Cotten M, Robinson M, Berry N, Pillay D, Kellam P. 2012. Universal Amplification, Next-Generation Sequencing, and Assembly of HIV-1 Genomes. Journal of Clinical Microbiology 50:3838-3844.

Gallo R, Salahuddin S, Popovic M, Shearer G, Kaplan M, Haynes B, Palker T, Redfield R, Oleske J, Safai B. 1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science 224:500-503.

Gamble T, Vajdos F, Yoo S, Worthylake D, Houseweart M, Sundquist W, Hill C. 1996. Crystal Structure of Human Cyclophilin A Bound to the Amino-Terminal Domain of HIV-1 Capsid. Cell 87:1285-1294.

Gamble T, Shanghee Y, Vajdos F, von Schwedler U, Worthylake D, Wang H, McCutcheon J, Sundquist W, Hill C. 1997. Structure of the Carboxyl-Terminal Dimerization Domain of the HIV-1 Capsid Protein. Science 278:849-853.

Ganser B, Li S, Klishko V, Finch J, Sundquist W. 1999. Assembly and Analysis of Conical Models for the HIV-1 Core. Science 283:80-83.

Ganser-Pornillos B, Yeager M, Sundquist W. 2008. The structural biology of HIV assembly. Current Opinion in Structural Biology 18:203-217.

Gao F, Robertson D, Morrison S, Hui H, Craig S, Decker J, Fultz P, Girard M, Shaw G, Hahn B, Sharp P. 1996. The heterosexual human immunodeficiency virus type 1 epidemic in Thailand is caused by an intersubtype (A/E) recombinant of African origin. Journal of Virology 70(10):7013-7029.

Gao F, Bailes E, Robertson D, Chen Y, Rodenburg C, Michael A, Cummins L, Arthur L, Peeters M, Shaw G, Sharp P, Hahn B. 1998. Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. Nature 397:436-441.

Gao X, Nelson G, Karacki P, Martin M, Phair J, Kaslow R, Goedert J,

Buchbinder S, Hoots K, Vlahov D, O'Brien S, Carrington M. 2001. Effect of a Single Amino Acid Change in MHC Class I Molecules on the Rate of Progression to AIDS. New England Journal of Medicine 344:1668-1675.

Garbi N, Tan P, Diehl A, Chambers B, Ljunggren H, Momburg F, Hämmerling G. 2000. Impaired immune responses and altered peptide repertoire in tapasin-deficient mice. Nature Immunology 1(3):234-238.

Gartner S, Markovits P, Markovitz D, Kaplan M, Gallo R, Popovic M. 1986. The role of mononuclear phagocytes in HTLV-III/LAV infection. Science 233:215-219.

Gasper-Smith N, Crossman D, Whitesides J, Mensali N, Ottinger J, Plonk S, Moody M, Ferrari G, Weinhold K, Miller S, Reich C, Qin L, Self S, Shaw G, Denny T, Jones L, Pisetsky D, Haynes B. 2008. Induction of Plasma (TRAIL), TNFR-2, Fas Ligand, and Plasma Microparticles after Human Immunodeficiency Virus Type 1 (HIV-1) Transmission: Implications for HIV-1 Vaccine Design. Journal of Virology 82:7700-7710.

Geels M, Cornelissen M, Schuitemaker H, Anderson K, Kwa D, Maas J, Dekker J, Baan E, Zorgdrager F, van den Burg R, van Beelen M, Lukashov V, Fu T, Paxton W, van der Hoek L, Dubey S, Shiver J, Goudsmit J. 2003. Identification of Sequential Viral Escape Mutants Associated with Altered T-Cell Responses in a Human Immunodeficiency Virus Type 1-Infected Individual. Journal of Virology 77:12430-12440.

Geijtenbeek T, Kwon D, Torensma R, van Vliet S, van Duijnhoven G, Middel J, Cornelissen I, Nottet H, KewalRamani V, Littman D, Figdor C, van Kooyk Y. 2000. DC-SIGN, a Dendritic Cell–Specific HIV-1-Binding Protein that Enhances trans-Infection of T Cells. Cell 100:587-597.

Geretti A. 2006. Antiretroviral resistance in clinical practice, 1st ed. Mediscript Ltd., London, England.

Gervaix A, West D, Leoni L, Richman D, Wong-Staal F, Corbeil J. 1997. A new reporter cell line to monitor HIV infection and drug susceptibility in vitro. Proceedings of the National Academy of Sciences 94:4653-4658.

Goedhals D, Rossouw I, Hallbauer U, Mamabolo M, de Oliveira T. 2012. The tainted milk of human kindness. The Lancet 380:702.

Goepfert P, Lumm W, Farmer P, Matthews P, Prendergast A, Carlson J, Derdeyn C, Tang J, Kaslow R, Bansal A, Yusim K, Heckerman D, Mulenga J, Allen S, Goulder P, Hunter E. 2008. Transmission of HIV-1 Gag immune escape mutations is associated with reduced viral load in linked recipients. The Journal of Experimental Medicine 205:1009-1017.

González-Ortega E, Ballana E, Badia R, Clotet B, Esté J. 2011. Compensatory mutations rescue the virus replicative capacity of VIRIP-resistant HIV-1. Antiviral Research 92:479-483.

Goonetilleke N, Liu M, Salazar-Gonzalez J, Ferrari G, Giorgi E, Ganusov V, Keele B, Learn G, Turnbull E, Salazar M, Weinhold K, Moore S, Letvin N, Haynes B, Cohen M, Hraber P, Bhattacharya T, Borrow P, Perelson A, Hahn B, Shaw G, Korber B, McMichael A. 2009. The first T cell response to

transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection. The Journal of Experimental Medicine 206:1253-1272.

Gottlieb M, Schroff R, Schanker H, Weisman J, Fan P, Wolf R, Saxon A. 1981. Pneumocystis carinii Pneumonia and Mucosal Candidiasis in Previously Healthy Homosexual Men. New England Journal of Medicine 305:1425-1431.

Gosselin A, Monteiro P, Chomont N, Diaz-Griffero F, Said E, Fonseca S, Wacleche V, El-Far M, Boulassel M, Routy J, Sekaly R, Ancuta P. 2009. Peripheral Blood CCR4+CCR6+ and CXCR3+CCR6+ CD4+ T Cells Are Highly Permissive to HIV-1 Infection. The Journal of Immunology 184:1604-1616.

Goulder P, Phillips R, Colbert R, McAdam S, Ogg G, Nowak M, Glangrande P, Luzzi G, Morgana B, Edwards A, McMichael A, Rowland-Jones S. 1997. Late escape from an immunodominant cytotoxic T lymphocyte response associated with progression to AIDS. Immunology Letters 56:25.

Goulder P, Brander C, Tang Y, Tremblay C, Colbert R, Addo M, Rosenberg E, Nguyen T, Allen R, Trocha A, Altfeld M, He S, Bunce M, Funkhouser R, Pelton S, Burchett S, McIntosh K, Korber B, Walker B. 2001. Evolution and transmission of stable CTL escape mutations in HIV infection. Nature. 412:334–338.

Goulder P and Walker B. 2012. HIV and HLA Class I: An Evolving Relationship. Immunity 37:426-440.

Goulder P, Lewin S, Leitman E. 2016. Paediatric HIV infection: the potential for cure. Nat Rev Immunol.16:259–271.

Granier C, Battivelli E, Lecuroux C, Venet A, Lambotte O, Schmitt-Boulanger M, Delaugerre C, Molina J, Chakrabarti L, Clavel F, Hance A. 2013. Pressure from TRIM5 Contributes to Control of HIV-1 Replication by Individuals Expressing Protective HLA-B Alleles. Journal of Virology 87:10368-10380.

Granoff A and Webster R. 1999. Encyclopedia of Virology, Three-Volume Set, 2nd ed. 763-774.

Gray G, Allen M, Moodie Z, Churchyard G, Bekker L, Nchabeleng M, Mlisana K, Metch B, de Bruyn G, Latka M, Roux S, Mathebula M, Naicker N, Ducar C, Carter D, Puren A, Eaton N, McElrath M, Robertson M, Corey L, Kublin J. 2011. Safety and efficacy of the HVTN 503/Phambili Study of a clade-B-based HIV-1 vaccine in South Africa: a double-blind, randomised, placebo-controlled test-of-concept phase 2b study. The Lancet Infectious Diseases 11:507-515.

Greene W and Peterlin B. 2002. Charting HIV's remarkable voyage through the cell: Basic science as a passport to future therapy. Nature Medicine 8:673-680.

Gross I, Hohenberg H, Huckhagel C, Kräusslich G. 1998. N-Terminal extension of human immunodeficiency virus capsid protein converts the in vitro assembly phenotype from tubular to spherical particles. Journal of Virology 72:4798-4810.

Gijsbers E, Feenstra K, van Nuenen A, Navis M, Heringa J, Schuitemaker H,

Kootstra N. 2013. HIV-1 Replication Fitness of HLA-B*57/58:01 CTL Escape Variants Is Restored by the Accumulation of Compensatory Mutations in Gag. PLoS ONE 8:e81235.

Guyader M, Emerman M, Sonigo P, Clavel F, Montagnier L, Alizon M. 1987. Genome organization and transactivation of the human immunodeficiency virus type 2. Nature 326:662-669.

Hadida F, Haas G, Zimmermann N, Hosmalin A, Spohn R, Samri A, Jung G, Debre P, Autran B. 1995. CTLs from lymphoid organs recognize an optimal HLA-A2-restricted and HLA-B52-restricted nonapeptide and several epitopes in the C-terminal region of HIV-1 Nef. The Journal of Immunology 154:4174-4186.

Han C, Kawana-Tachikawa A, Shimizu A, Zhu D, Nakamura H, Adachi E, Kikuchi T, Koga M, Koibuchi T, Gao G, Sato Y, Yamagata A, Martin E, Fukai S, Brumme Z, Iwamoto A. 2014. Switching and emergence of CTL epitopes in HIV-1 infection. Retrovirology 11:38.

Hang I, Lin C, Grant O, Fleurkens S, Villiger T, Soos M, Morbidelli M, Woods R, Gauss R, Aebi M. 2015. Analysis of site-specificN-glycan remodeling in the endoplasmic reticulum and the Golgi. Glycobiology 25:1335-1349.

Harada S, Koyanagi Y, Yamamoto N. 1985. Infection of HTLV-III/LAV in HTLV-I-carrying cells MT-2 and MT-4 and application in a plaque assay. Science 229:563-566.

Harjanto S, Ng L, Tong J. 2014. Clustering HLA Class I Superfamilies Using Structural Interaction Patterns. PLoS ONE 9:e86655.

Hansen S, Wu H, Burwitz B, Hughes C, Hammond K, Ventura A, Reed J, Gilbride R, Ainslie E, Morrow D, Ford J, Selseth A, Pathak R, Malouli D, Legasse A, Axthelm M, Nelson J, Gillespie G, Walters L, Brackenridge S, Sharpe H, López C, Früh K, Korber B, McMichael A, Gnanakaran S, Sacha J, Picker L. 2016 Broadly targeted CD8+ T cell responses restricted by major histocompatibility complex E. Science 351(6274): 714-720.

Hatziioannou T, Cowan S, von Schwedler U, Sundquist W, Bieniasz P. 2004. Species-Specific Tropism Determinants in the Human Immunodeficiency Virus Type 1 Capsid. Journal of Virology 78:6005-6012.

Hayday A. 2009. γδ T Cells and the Lymphoid Stress-Surveillance Response. Immunity 31:184-196.

Hemelaar J, Gouws E, Ghys P, Osmanov S, WHO-UNAIDS Network for HIV Isolation and Characterisation. 2011. Global trends in molecular epidemiology of HIV-1 during 2000–2007. AIDS 25:679-689.

Hessell A, Hangartner L, Hunter M, Havenith C, Beurskens F, Bakker J, Lanigan C, Landucci G, Forthal D, Parren P, Marx P, Burton D. 2007. Fc receptor but not complement binding is important in antibody protection against HIV. Nature 449:101-104.

HLA Nomenclature. In hla.alleles.org. Accessed on 19 July 2017.

Hölzemer A, Thobakgale C, Jimenez Cruz C, Garcia-Beltran W, Carlson J, van Teijlingen N, Mann J, Jaggernath M, Kang S, Körner C, Chung A, Schafer J, Evans D, Alter G, Walker B, Goulder P, Carrington M, Hartmann P, Pertel T, Zhou R, Ndung'u T, Altfeld M. 2015. Selection of an HLA-C*03:04-Restricted HIV-1 p24 Gag Sequence Variant Is Associated with Viral Escape from KIR2DL3+ Natural Killer Cells: Data from an Observational Cohort in South Africa. PLOS Medicine 12:e1001900.

Honeyborne I, Codoner F, Leslie A, Tudor-Williams G, Luzzi G, Ndung'u T, Walker B, Goulder P, Prado J. 2010. HLA-Cw*03-Restricted CD8+ T-Cell Responses Targeting the HIV-1 Gag Major Homology Region Drive Virus Immune Escape and Fitness Constraints Compensated for by Intracodon Variation. Journal of Virology 84:11279-11288.

Horowitz A, Djaoud Z, Nemat-Gorgani N, Blokhuis J, Hilton H, Béziat V, Malmberg K, Norman P, Guethlein L, Parham P. 2016. Class I HLA haplotypes form two schools that educate NK cells in different ways. Science Immunology 1:eaag1672-eaag1672.

Horton R, Wilming L, Rand V, Lovering R, Bruford E, Khodiyar V, Lush M, Povey S, Talbot C, Wright M, Wain H, Trowsdale J, Ziegler A, Beck S. 2004. Gene map of the extended human MHC. Nature Reviews Genetics 5:889-899.

Hinkley T, Martins J, Chappey C, Haddad M, Stawiski E, Whitcomb J, Petropoulos C, Bonhoeffer S. 2011. A systems analysis of mutational effects in HIV-1 protease and reverse transcriptase. Nature Genetics 43:487-489.

Huang K, Goedhals D, Fryer H, van Vuuren C, Katzourakis A, De Oliveira T, Brown H, Cassol S, Seebregts C, McLean A, Klenerman P, Phillips R, Frater J. 2009. Prevalence of HIV type-1 drug-associated mutations in pre-therapy patients in the Free State, South Africa. Antiviral Therapy 14:975-984.

Hubert J, Burgard M, Dussaix E, Tamalet C, Deveau C, Le Chenadec J, Chaix M, Marchadier E, Vildé J, Delfraissy J, Meyer L, Rouzioux C. 2000. Natural history of serum HIV-1 RNA levels in 330 patients with a known date of infection. AIDS 14:123-131.

Hunt G, Hopkins M, Lidgard S. 2015. Simple versus complex models of trait evolution and stasis as a response to environmental change. Proceedings of the National Academy of Sciences 112:4885-4890.

Ito T, Kanzler H, Duramad O, Cao W, Liu Y. 2006. Specialization, kinetics, and repertoire of type 1 interferon responses by human plasmacytoid predendritic cells. Blood 107:2423-2431.

Ivanov D, Tsodikov O, Kasanov J, Ellenberger T, Wagner G, Collins T. 2007. Domain-swapped dimerization of the HIV-1 capsid C-terminal domain. Proceedings of the National Academy of Sciences 104:4353-4358.

Jacks T, Power M, Masiarz F, Luciw P, Barr P, Varmus H. 1988. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. Nature 331:280-283.

Jamieson B, Yang O, Hultin L, Hausner M, Hultin P, Matud J, Kunstman K,

Killian S, Altman J, Kommander K, Korber B, Giorgi J, Wolinsky S. 2003. Epitope Escape Mutation and Decay of Human Immunodeficiency Virus Type 1-Specific CTL Responses. The Journal of Immunology 171:5372-5379.

Jiménez V, Wit F, Joerink M, Maurer I, Harskamp A, Schouten J, Prins M, van Leeuwen E, Booiman T, Deeks S, Reiss P, Kootstra N. 2016. T-Cell Activation Independently Associates With Immune Senescence in HIV-Infected Recipients of Long-term Antiretroviral Treatment. Journal of Infectious Diseases 214:216-225.

Jani I, Meggi B, Mabunda N, Vubil A, Sitoe N, Tobaiwa O, Quevedo J, Lehe J, Loquiha O, Vojnov L, Peter T. 2014. Accurate Early Infant HIV Diagnosis in Primary Health Clinics Using a Point-of-Care Nucleic Acid Test. JAIDS Journal of Acquired Immune Deficiency Syndromes 67:e1-e4.

Jetzt A, Yu H, Klarmann G, Ron Y, Preston B, Dougherty J. 2000. High Rate of Recombination throughout the Human Immunodeficiency Virus Type 1 Genome. Journal of Virology 74:1234-1240.

Jiang J, Ablan S, Derebail S, Hercík K, Soheilian F, Thomas J, Tang S, Hewlett I, Nagashima K, Gorelick R, Freed E, Levin J. 2011. The interdomain linker region of HIV-1 capsid protein is a critical determinant of proper core assembly and stability. Virology 421:253-265.

John G and Kreiss J. 1996. Mother-to-Child Transmission of Human Immunodeficiency Virus Type 1. Epidemiologic Reviews 18:149-157.

John M, Heckerman D, James I, Park L, Carlson J, Chopra A, Gaudieri S, Nolan D, Haas D, Riddler S, Haubrich R, Mallal S. 2010. Adaptive Interactions between HLA and HIV-1: Highly Divergent Selection Imposed by HLA Class I Molecules with Common Supertype Motifs. The Journal of Immunology 184:4368-4377.

Jones N, Wei X, Flower D, Wong M, Michor F, Saag M, Hahn B, Nowak M, Shaw G, Borrow P. 2004. Determinants of Human Immunodeficiency Virus Type 1 Escape from the Primary CD8+Cytotoxic T Lymphocyte Response. The Journal of Experimental Medicine 200:1243-1256.

Jones M and Núñez M. 2012. Liver Toxicity of Antiretroviral Drugs. Seminars in Liver Disease 32:167-176.

Jost S, Bernard M, Kaiser L, Yerly S, Hirschel B, Samri A, Autran B, Goh L, Perrin L. 2002. A Patient With HIV-1 Superinfection. Infectious Diseases in Clinical Practice 11:260.

Jourdain G, Mary J, Coeur S, Ngo-Giang-Huong N, Yuthavisuthi P, Limtrakul A, Traisathit P, McIntosh K, Lallemant M. 2007. Risk Factors for In Utero or Intrapartum Mother-to-Child Transmission of Human Immunodeficiency Virus Type 1 in Thailand. The Journal of Infectious Diseases 196:1629-1636.

Juarez-Molina C, Payne R, Soto-Nava M, Avila-Rios S, Valenzuela-Ponce H, Adland E, Leitman E, Brener J, Muenchhoff M, Branch S, Landis C, Reyes-Teran G, Goulder P. 2014. Impact of HLA Selection Pressure on HIV Fitness at a Population Level in Mexico and Barbados. Journal of Virology 88:10392-10398.

Juarez-Molina C, Valenzuela-Ponce H, Avila-Rios S, Garrido-Rodriguez D, Garcia-Tellez T, Soto-Nava M, Garcia-Morales C, Goulder P, Reyes-Teran G. 2014. Impact of HLA-B*35 subtype differences on HIV disease outcome in Mexico. AIDS 28:1687-1690.

Jung A, Maier R, Vartanian J, Bocharov G, Jung V, Fischer U, Meese E, Wain-Hobson S, Meyerhans A. 2002. Recombination: Multiply infected spleen cells in HIV patients. Nature 418:144-144.

Juno J, van Bockel D, Kent S, Kelleher A, Zaunders J, Munier C. 2017. Cytotoxic CD4 T Cells—Friend or Foe during Viral Infection?. Frontiers in Immunology 8.

Kao S, Calman A, Luciw P, Peterlin B. 1987. Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. Nature 330:489-493.

Khan MA. 2010. HLA and spondyloarthropathies. In Mehra N (Ed). The HLA Complex in Biology and Medicine Jaypee Brothers Medical Publishers 422-44.

Kaslow R, Carrington M, Apple R, Park L, Muñoz A, Saah A, Goedert J, Winkler C, O'Brien S, Rinaldo C, Detels R, Blattner W, Phair J, Erlich H, Mann D. 1996. Influence of combinations of human major histocompatibility complex genes on the course of HIV–1 infection. Nature Medicine 2:405-411.

Kaufmann D, Bailey P, Sidney J, Wagner B, Norris P, Johnston M, Cosimi L, Addo M, Lichterfeld M, Altfeld M, Frahm N, Brander C, Sette A, Walker B, Rosenberg E. 2004. Comprehensive Analysis of Human Immunodeficiency Virus Type 1-Specific CD4 Responses Reveals Marked Immunodominance of gag and nef and the Presence of Broadly Recognized Peptides. Journal of Virology 78:4463-4477.

Kawashima Y, Pfafferott K, Frater J, Matthews P, Payne R, Addo M, Gatanaga H, Fujiwara M, Hachiya A, Koizumi H, Kuse N, Oka S, Duda A, Prendergast A, Crawford H, Leslie A, Brumme Z, Brumme C, Allen T, Brander C, Kaslow R, Tang J, Hunter E, Allen S, Mulenga J, Branch S, Roach T, John M, Mallal S, Ogwu A, Shapiro R, Prado JG, Fidler S, Weber J, Pybus OG, Klenerman P, Ndung'u T, Phillips R, Heckerman D, Harrigan PR, Walker BD, Takiguchi M, Goulder P. 2009. Adaptation of HIV-1 to human leukocyte antigen class I. Nature. 458(7238):641–645.

Kawashima Y, Kuse N, Gatanaga H, Naruto T, Fujiwara M, Dohki S, Akahoshi T, Maenaka K, Goulder P, Oka S, Takiguchi M. 2010. Long-Term Control of HIV-1 in Hemophiliacs Carrying Slow-Progressing Allele HLA-B*5101. Journal of Virology 84:7151-7160.

Keele B, Giorgi E, Salazar-Gonzalez J, Decker J, Pham K, Salazar M, Sun C, Grayson T, Wang S, Li H, Wei X, Jiang C, Kirchherr J, Gao F, Anderson J, Ping L, Swanstrom R, Tomaras G, Blattner W, Goepfert P, Kilby J, Saag M, Delwart E, Busch M, Cohen M, Montefiori D, Haynes B, Gaschen B, Athreya G, Lee H, Wood N, Seoighe C, Perelson A, Bhattacharya T, Korber B, Hahn B, and Shaw G. 2008. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection PNAS 105(21):7552-7557.

Kelleher A, Long C, Holmes E, Allen R, Wilson J, Conlon C, Workman C, Shaunak S, Olson K, Goulder P, Brander C, Ogg G, Sullivan J, Dyer W, Jones I, McMichael A, Rowland-Jones S, Phillips R. 2001. Clustered Mutations in HIV-1 Gag Are Consistently Required for Escape from Hla-B27–Restricted Cytotoxic T Lymphocyte Responses. The Journal of Experimental Medicine 193:375-386.

Keller M, Lu Y, Lalonde R, Klein M. 2009. Impact of HIV-1 viral subtype on CD4+ T-cell decline and clinical outcomes in antiretroviral naive patients receiving universal healthcare. AIDS 23(6):731-737.

Kiepiela P, Leslie A, Honeyborne I, Ramduth D, Thobakgale C, Chetty S, Rathnavalu P, Moore C, Pfafferott K, Hilton L, Zimbwa P, Moore S, Allen T, Brander C, Addo M, Altfeld M, James I, Mallal S, Bunce M, Barber L, Szinger J, Day C, Klenerman P, Mullins J, Korber B, Coovadia H, Walker B, Goulder P. 2004. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. Nature 432:769-775.

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

Killian M, Fujimura S, Hecht F, Levy J. 2006. Similar changes in plasmacytoid dendritic cell and CD4 T-cell counts during primary HIV-1 infection and treatment. AIDS 20:1247-1252.

Kinouchi Y, Matsumoto K, Negoro K, Takagi S, Takahashi S, Hiwatashi N, Shimosegawa T. 2003. HLA-B genotype in Japanese patients with Crohn's disease. Disease Colon Rectum. 46(10 Suppl):S10-14.

Kisselev A, Akopian T, Woo K, Goldberg A. 1999. The Sizes of Peptides Generated from Protein by Mammalian 26 and 20 S Proteasomes: Implications For Understanding The Degradative Mechanism And Antigen Presentation. Journal of Biological Chemistry 274:3363-3371.

Kløverpris H, Adland E, Koyanagi M, Stryhn A, Harndahl M, Matthews P, Shapiro R, Walker B, Ndung'u T, Brander C, Takiguchi M, Buus S, Goulder P. 2014. HIV Subtype Influences HLA-B*07:02-Associated HIV Disease Outcome. AIDS Research and Human Retroviruses 30:468-475.

Kløverpris H, Leslie A, Goulder P. 2016. Role of HLA Adaptation in HIV Evolution. Front. Immunol. 2016;6:665.

Kobayashi M, Igarashi H, Takeda A, Kato M, Matano T. 2005. Reversion In Vivo after Inoculation of a Molecular Proviral DNA Clone of Simian Immunodeficiency Virus with a Cytotoxic-T-Lymphocyte Escape Mutation. Journal of Virology 79:11529-11532.

Koehler R, Walsh A, Saathoff E, Tovanabutra S, Arroyo M, Currier J, Maboko L,

Hoelsher M, Robb M, Michael N, McCutchan F, Kim J, Kijak G. 2010. Class I HLA-A*7401 Is Associated with Protection from HIV-1 Acquisition and Disease Progression in Mbeya, Tanzania. The Journal of Infectious Diseases 202:1562-1566.

Koenig S, Gendelman H, Orenstein J, Dal Canto M, Pezeshkpour G, Yungbluth M, Janotta F, Aksamit A, Martin M, Fauci A. 1986. Detection of AIDS virus in macrophages in brain tissue from AIDS patients with encephalopathy. Science 233:1089-1093.

Könnyű B, Sadiq S, Turányi T, Hírmondó R, Müller B, Kräusslich H, Coveney P, Müller V. 2013. Gag-Pol Processing during HIV-1 Virion Maturation: A Systems Biology Approach. PLoS Computational Biology 9:e1003103.

Korber B, Muldoon M, Theiler J, Gao F, Gupta R, Lapedes A, Hahn BH, Wolinsky S, Bhattacharya T. 2000. Timing the ancestor of the HIV-1 pandemic strains. Science 288: 1789–1796.

Koup R. 1994. Virus escape from CTL recognition. Journal of Experimental Medicine 180:779-782.

Koup R, Safrit J, Cao Y, Andrews C, McLeod G, Borkowsky W, Farthing C, Ho D. 1994. 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):4650-4655.

Krebs F, Hogan T, Quiterio S, Gartner S, Wigdahl B. 2001. Lentiviral LTR-directed Expression, Sequence Variation, and Disease Pathogenesis. In Kuiken, C; Foley, B; B; Marx, P; McCutchan, F; Mellors, JW; Wolinsky, S; Korber, B. *HIV Sequence Compendium 2001*. Los Alamos, NM: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory 29–70.

Kulkarni S, Savan R, Qi Y, Gao X, Yuki Y, Bass S, Martin M, Hunt P, Deeks S, Telenti A, Pereyra F, Goldstein D, Wolinsky S, Walker B, Young H, Carrington M. 2011. Differential microRNA regulation of HLA-C expression and its association with HIV control. Nature 472:495-498.

Lapenta C, Santini S, Proietti E, Rizza P, Logozzi M, Spada M, Parlato S, Fais S, Pitha P, Belardelli F. 1999. Type I Interferon Is a Powerful Inhibitor of in Vivo HIV-1 Infection and Preserves Human CD4+ T Cells from Virus-Induced Depletion in SCID Mice Transplanted with Human Cells. Virology 263:78-88.

Lambotte O, Ferrari G, Moog C, Yates N, Liao H, Parks R, Hicks C, Owzar K, Tomaras G, Montefiori D, Haynes B, Delfraissy J. 2009. Heterogeneous neutralizing antibody and antibody-dependent cell cytotoxicity responses in HIV-1 elite controllers. AIDS 23:897-906.

Larder B, Darby G, Richman D. 1989. HIV with reduced sensitivity to zidovudine (AZT) isolated during prolonged therapy. Science 243:1731-1734.

Lazaryan A, Song W, Lobashevsky E, Tang J, Shrestha S, Zhang K, McNicholl J, Gardner L, Wilson C, Klein R, Rompalo A, Mayer K, Sobel J, Kaslow R. 2011. The influence of human leukocyte antigen class I alleles and their population frequencies on human immunodeficiency virus type 1 control among African Americans. Human Immunology 72:312-318.

Leitman E, Hurst J, Mori M, Kublin J, Ndung'u T, Walker B, Carlson J, Gray G, Matthews P, Frahm N, Goulder P. 2016. Lower Viral Loads and Slower CD4+T-Cell Count Decline in MRKAd5 HIV-1 Vaccinees Expressing Disease-Susceptible HLA-B*58:02. Journal of Infectious Diseases 214:379-389.

Leslie A, Pfafferott K, Chetty P, Draenert R, Addo M, Feeney M, Tang Y, Holmes E, Allen T, Prado J, Altfeld M, Brander C, Dixon C, Ramduth D, Jeena P, Thomas S, John A, Roach T, Kupfer B, Luzzi G, Edwards A, Taylor G, Lyall H, Tudor-Williams G, Novelli V, Martinez-Picado J, Kiepiela P, Walker B, Goulder P. 2004. HIV evolution: CTL escape mutation and reversion after transmission. Nature Medicine 10:282-289.

Leslie A, Kavanagh D, Honeyborne I, Pfafferott K, Edwards C, Pillay T, Hilton L, Thobakgale C, Ramduth D, Draenert R, Le Gall S, Luzzi G, Edwards A, Brander C, Sewell A, Moore S, Mullins J, Moore C, Mallal S, Bhardwaj N, Yusim K, Phillips R, Klenerman P, Korber B, Kiepiela P, Walker B, Goulder P. 2006. Transmission and accumulation of CTL escape variants drive negative associations between HIV polymorphisms and HLA. J Exp Med. 201:891–902.

Leslie A, Matthews P, Listgarten J, Carlson J, Kadie C, Ndung'u T, Brander C, Coovadia H, Walker B, Heckerman D, Goulder P. 2010. Additive Contribution of HLA Class I Alleles in the Immune Control of HIV-1 Infection. Journal of Virology 84:9879-9888.

Levesque M, Moody M, Hwang K, Marshall D, Whitesides J, Amos J, Gurley T, Allgood S, Haynes B, Vandergrift N, Plonk S, Parker D, Cohen M, Tomaras G, Goepfert P, Shaw G, Schmitz J, Eron J, Shaheen N, Hicks C, Liao H, Markowitz M, Kelsoe G, Margolis D, Haynes B. 2009. Polyclonal B Cell Differentiation and Loss of Gastrointestinal Tract Germinal Centers in the Earliest Stages of HIV-1 Infection. PLoS Medicine 6:e1000107.

Llano A, Williams A, Olvera A, Silva-Arrieta S, Brander C. 2013. Best-Characterized HIV-1 CTL Epitopes: The 2013 Update. HIV Molecular Immunology 2013: 3–25.

Li S, Hill P, Sundquist W, Finch T. 2000. Image reconstructions of helical assemblies of the HIV-1 CA protein. Nature 407:409–413.

Li S, Jiao H, Yu X, Strong A, Shao Y, Sun Y, Altfeld M, Lu Y. 2007. Human Leukocyte Antigen Class I and Class II Allele Frequencies and HIV-1 Infection Associations in a Chinese Cohort. JAIDS Journal of Acquired Immune Deficiency Syndromes 44:121-131.

Li Q, Estes J, Schlievert P, Duan L, Brosnahan A, Southern P, Reilly C, Peterson M, Schultz-Darken N, Brunner K, Nephew K, Pambuccian S, Lifson J, Carlis J, Haase A. 2009. Glycerol monolaurate prevents mucosal SIV transmission. Nature 458:1034-1038.

Lichterfeld M, Kaufmann D, Yu X, Mui S, Addo M, Johnston M, Cohen D, Robbins G, Pae E, Alter G, Wurcel A, Stone D, Rosenberg E, Walker B, Altfeld M. 2004. Loss of HIV-1–specific CD8+T Cell Proliferation after Acute HIV-1 Infection and Restoration by Vaccine-induced HIV-1–specific CD4+T Cells. The Journal of Experimental Medicine 200:701-712.

Lichterfeld M, Kavanagh D, Williams K, Moza B, Mui S, Miura T, Sivamurthy R, Allgaier R, Pereyra F, Trocha A, Feeney M, Gandhi R, Rosenberg E, Altfeld M, Allen T, Allen R, Walker B, Sundberg E, Yu X. 2007. A viral CTL escape mutation leading to immunoglobulin-like transcript 4–mediated functional inhibition of myelomonocytic cells. The Journal of Experimental Medicine 204:2813-2824.

Lichterfeld M and Yu X. 2011. The emerging role of leukocyte immunoglobulin-like receptors (LILRs) in HIV-1 infection. Journal of Leukocyte Biology 91:27-33.

Limou S, Le Clerc S, Coulonges C, Carpentier W, Dina C, Delaneau O, Labib T, Taing L, Sladek R, Deveau C, Ratsimandresey R, Montes M, Spadoni J, Lelièvre J, Lévy Y, Therwath A, Schächter F, Matsuda F, Gut I, Froguel P, Delfraissy J, Hercberg S, Zagury J. 2009. Genomewide Association Study of an AIDS-Nonprogression Cohort Emphasizes the Role Played byHLAGenes (ANRS Genomewide Association Study 02). The Journal of Infectious Diseases 199:419-426.

Lin P, Lai C, Yang J, Huang H, Huang M, Tsai M, Yang C, Cheng C, Su Y, Chang S, Liu W, Hung C, Chang S, Chang S. 2013. Slow immunological progression in HIV-1 CRF07_BC-infected injecting drug users. Emerging Microbes & Infections 2:e83.

Lin Z, Kuroki K, Kuse N, Sun X, Akahoshi T, Qi Y, Chikata T, Naruto T, Koyanagi M, Murakoshi H, Gatanaga H, Oka S, Carrington M, Maenaka K, Takiguchi M. 2016. HIV-1 Control by NK Cells via Reduced Interaction between KIR2DL2 and HLA-C∗12:02/C∗14:03. Cell Reports 17:2210-2220.

Lindenbach D. 2009. Measuring HCV infectivity produced in cell culture and in vivo. Methods Molecular Biology 510:329-336.

Liu M, Hawkins N, Ritchie A, Ganusov V, Whale V, Brackenridge S, Li H, Pavlicek J, Cai F, Rose-Abrahams M, Treurnicht F, Hraber P, Riou C, Gray C, Ferrari G, Tanner R, Ping L, Anderson J, Swanstrom R, B C, Cohen M, Karim S, Haynes B, Borrow P, Perelson A, Shaw G, Hahn B, Williamson C, Korber B, Gao F, Self S, McMichael A, Goonetilleke N. 2012. Vertical T cell immunodominance and epitope entropy determine HIV-1 escape. Journal of Clinical Investigation 123(1):380-393.

Liu D, Zuo T, Hora B, Song H, Kong W, Yu X, Goonetilleke N, Bhattacharya T, Perelson A, Haynes B, McMichael A, Gao F. 2014. Preexisting compensatory amino acids compromise fitness costs of a HIV-1 T cell escape mutation. Retrovirology 11.

Ljunggren H and Kärre K. 1990. In search of the ‘missing self’: MHC molecules and NK cell recognition. Immunology Today 11:237-244.

Lozupone C, Rhodes M, Neff C, Fontenot A, Campbell T, Palmer B. 2014. HIV-induced alteration in gut microbiota. Gut Microbes 5:562-570.

Lu K, Heng X, Summers M. 2011. Structural Determinants and Mechanism of HIV-1 Genome Packaging. Journal of Molecular Biology 410:609-633.

Lugada E, Mermin J, Kaharuza F, Ulvestad E, Were W, Langeland N, Asjo B,

Malamba S, Downing R. 2004. Population-Based Hematologic and Immunologic Reference Values for a Healthy Ugandan Population. Clinical and Vaccine Immunology 11:29-34.

Luzuriaga K, Holmes D, Hereema A, Wong J, Panicali D, Sullivan J. 1995. HIV-1-specific cytotoxic T lymphocyte responses in the first year of life. Journal of Immunology. 154(1):433–443.

Luzuriaga K, Gay H, Ziemniak C, Sanborn K, Somasundaran M, Rainwater-Lovett K, Mellors J, Rosenbloom D, Persaud D. 2015. Viremic Relapse after HIV-1 Remission in a Perinatally Infected Child. New England Journal of Medicine 372:786-788.

Lyles R, Muñoz A, Yamashita T, Bazmi H, Detels R, Rinaldo C, Margolick J, Phair J, Mellors J. 2000. Natural History of Human Immunodeficiency Virus Type 1 Viremia after Seroconversion and Proximal to AIDS in a Large Cohort of Homosexual Men. The Journal of Infectious Diseases 181:872-880.

Marcello A. 2006. Latency: the hidden HIV-1 challenge. Retrovirology 16;3:7.

Marinda E, Humphrey J, Iliff P, Mutasa K, Nathoo K, Piwoz E, Moulton L, Salama P, Ward B. 2007. Child Mortality According to Maternal and Infant HIV Status in Zimbabwe. The Pediatric Infectious Disease Journal 26:519-526.

Marsh S, Parham P, Barber L. 2000. The HLA factsbook, 1st ed. Academic Press, San Diego. 52-56, 218-219.

Martin M, Gao X, Lee J, Nelson G, Detels R, Goedert J, Buchbinder S, Hoots K, Vlahov D, Trowsdale J, Wilson M, O'Brien S, Carrington M. 2002. Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS. Nature Genetics 31(4):429-434.

Martin M, Qi Y, Gao X, Yamada E, Martin J, Pereyra F, Colombo S, Brown E, Shupert W, Phair J, Goedert J, Buchbinder S, Kirk G, Telenti A, Connors M, O'Brien S, Walker B, Parham P, Deeks S, McVicar D, Carrington M. 2007. Innate partnership of HLA-B and KIR3DL1 subtypes against HIV-1. Nature Genetics 39:733-740.

Martin-Gayo E, Buzon M, Ouyang Z, Hickman T, Cronin J, Pimenova D, Walker B, Lichterfeld M, Yu X. 2015. Potent Cell-Intrinsic Immune Responses in Dendritic Cells Facilitate HIV-1-Specific T Cell Immunity in HIV-1 Elite Controllers. PLOS Pathogens 11:e1004930.

Martinez-Picado J, Prado J, Fry E, Pfafferott K, Leslie A, Chetty S, Thobakgale C, Honeyborne I, Crawford H, Matthews P, Pillay T, Rousseau C, Mullins J, Brander C, Walker B, Stuart D, Kiepiela P, Goulder P. 2006. Fitness Cost of Escape Mutations in p24 Gag in Association with Control of Human Immunodeficiency Virus Type 1. Journal of Virology 80:3617-3623.

Matthews P, Prendergast A, Leslie A, Crawford H, Payne R, Rousseau C, Rolland M, Honeyborne I, Carlson J, Kadie C, Brander C, Bishop K, Mlotshwa N, Mullins J, Coovadia H, Ndung'u T, Walker B, Heckerman D, Goulder P. 2008. Central Role of Reverting Mutations in HLA Associations

with Human Immunodeficiency Virus Set Point. Journal of Virology 82:8548-8559.

Matthews P, Koyanagi M, Kloverpris H, Harndahl M, Stryhn A, Akahoshi T, Gatanaga H, Oka S, Juarez Molina C, Valenzuela Ponce H, Avila Rios S, Cole D, Carlson J, Payne R, Ogwu A, Bere A, Ndung'u T, Gounder K, Chen F, Riddell L, Luzzi G, Shapiro R, Brander C, Walker B, Sewell A, Reyes Teran G, Heckerman D, Hunter E, Buus S, Takiguchi M, Goulder P. 2012. Differential Clade-Specific HLA-B*3501 Association with HIV-1 Disease Outcome Is Linked to Immunogenicity of a Single Gag Epitope. Journal of Virology 86:12643-12654.

Mammano F, Ohagen A, Höglund S, Göttlinger G. 1994. Role of the major homology region of human immunodeficiency virus type 1 in virion morphogenesis. Journal of Virology 68(8):4927-4936.

Maness N, Yant L, Chung C, Loffredo J, Friedrich T, Piaskowski S, Furlott J, May G, Soma T, Leon E, Wilson N, Piontkivska H, Hughes A, Sidney J, Sette A, Watkins D. 2008. Comprehensive Immunological Evaluation Reveals Surprisingly Few Differences between Elite Controller and Progressor Mamu-B*17-Positive Simian Immunodeficiency Virus-Infected Rhesus Macaques. Journal of Virology 82:5245-5254.

Mann J, Byakwaga H, Kuang X, Le A, Brumme C, Mwimanzi P, Omarjee S, Martin E, Lee G, Baraki B, Danroth R, McCloskey R, Muzoora C, Bangsberg D, Hunt P, Goulder P, Walker B, Harrigan P, Martin J, Ndung'u T, Brockman M, Brumme Z. 2013. Ability of HIV-1 Nef to downregulate CD4 and HLA class I differs among viral subtypes. Retrovirology 10:100.

Mann J, Barton J, Ferguson A, Omarjee S, Walker B, Chakraborty A, Ndung'u T. 2014. The Fitness Landscape of HIV-1 Gag: Advanced Modeling Approaches and Validation of Model Predictions by In Vitro Testing. PLoS Computational Biology 10:e1003776.

Mann J and Ndung'u T. 2015. HIV-1 vaccine immunogen design strategies. Virology Journal 12:3.

Mansky L and Temin H. 1995. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. Journal of Virology 69:5087-5094.

Martin M, Qi Y, Gao X, Yamada E, Martin J, Pereyra F, Colombo S, Brown E, Shupert W, Phair J, Goedert J, Buchbinder S, Kirk G, Telenti A, Connors M, O'Brien S, Walker B, Parham P, Deeks S, McVicar D, Carrington M. 2007. Innate partnership of HLA-B and KIR3DL1 subtypes against HIV-1. Nature Genetics 39:733-740.

Mateu M. 2009. The capsid protein of human immunodeficiency virus: intersubunit interactions during virus assembly. FEBS Journal 276:6098-6109.

Mathieu A, Paladini F, Vacca A, Cauli A, Fiorillo M, Sorrentino R. 2009. The interplay between the geographic distribution of HLA-B27 alleles and their role in infectious and autoimmune diseases: A unifying hypothesis. Autoimmunity Reviews 8:420-425.

McElrath M, De Rosa S, Moodie Z, Dubey S, Kierstead L, Janes H, Defawe O, Carter D, Hural J, Akondy R, Buchbinder S, Robertson M, Mehrotra D, Self S, Corey L, Shiver J, Casimiro D. 2008. HIV-1 vaccine-induced immunity in the test-of-concept Step Study: a case–cohort analysis. The Lancet 372:1894-1905.

McKinnon L, Capina R, Peters H, Mendoza M, Kimani J, Wachihi C, Kariri A, Kimani M, Richmond M, Kiazyk S, Fowke K, Jaoko W, Luo M, Ball T, Plummer F. 2009. Clade-Specific Evolution Mediated by HLA-B*57/5801 in Human Immunodeficiency Virus Type 1 Clade A1 p24. Journal of Virology 83:12636-12642.

McNamara L and Collins K. 2011. Hematopoietic stem/precursor cells as HIV reservoirs. Current Opinion in HIV and AIDS 6:43-48.

McMichael A, Borrow P, Tomaras G, Goonetilleke N, Haynes B. 2009. The immune response during acute HIV-1 infection: clues for vaccine development. Nature Reviews Immunology 10:11-23.

Mellors J, Rinaldo C, Gupta P, White R, Todd J, Kingsley L. 1996. Prognosis in HIV-1 Infection Predicted by the Quantity of Virus in Plasma. Science 272:1167-1170.

Migueles S, Sabbaghian M, Shupert W, Bettinotti M, Marincola F, Martino L, Hallahan C, Selig S, Schwartz D, Sullivan J, Connors M. 2000. HLA B*5701 is highly associated with restriction of virus replication in a subgroup of HIV-infected long term nonprogressors. Proceedings of the National Academy of Sciences 97:2709-2714.

Migueles S, Laborico A, Shupert W, Sabbaghian M, Rabin R, Hallahan C, Baarle D, Kostense S, Miedema F, McLaughlin M, Ehler L, Metcalf J, Liu S, Connors M. 2002. HIV-specific CD8+ T cell proliferation is coupled to perforin expression and is maintained in nonprogressors. Nature Immunology 3:1061-1068.

Migueles S, Osborne C, Royce C, Compton A, Joshi R, Weeks K, Rood J, Berkley A, Sacha J, Cogliano-Shutta N, Lloyd M, Roby G, Kwan R, McLaughlin M, Stallings S, Rehm C, O'Shea M, Mican J, Packard B, Komoriya A, Palmer S, Wiegand A, Maldarelli F, Coffin J, Mellors J, Hallahan C, Follman D, Connors M. 2008. Lytic Granule Loading of CD8+ T Cells Is Required for HIV-Infected Cell Elimination Associated with Immune Control. Immunity 29:1009-1021.

Milicic A, Price D, Zimbwa P, Booth B, Brown H, Easterbrook P, Olsen K, Robinson N, Gileadi U, Sewell A, Cerundolo V, Phillips R. 2005. CD8+ T Cell Epitope-Flanking Mutations Disrupt Proteasomal Processing of HIV-1 Nef. The Journal of Immunology 175:4618-4626.

Miura T, Brockman M, Schneidewind A, Lobritz M, Pereyra F, Rathod A, Block B, Brumme Z, Brumme C, Baker B, Rothchild A, Li B, Trocha A, Cutrell E, Frahm N, Brander C, Toth I, Arts E, Allen T, Walker B. 2009a. HLA-B57/B*5801 Human Immunodeficiency Virus Type 1 Elite Controllers Select for Rare Gag Variants Associated with Reduced Viral Replication Capacity

and Strong Cytotoxic T-Lymphocyte Recognition. Journal of Virology 83:5961-5961.

Miura T, Brumme C, Brockman M, Brumme Z, Pereyra F, Block B, Trocha A, John M, Mallal S, Harrigan P, Walker B. 2009b. HLA-Associated Viral Mutations Are Common in Human Immunodeficiency Virus Type 1 Elite Controllers. Journal of Virology 83:3407-3412.

Momany C, Kovari L, Prongay A, Keller W, Gitti R, Lee B, Gorbalenya A, Tong L, McClure J, Ehrlich L, Summers M, Carter C, Rossmann M. 1996. Crystal structure of dimeric HIV-1 capsid protein. Nature Structural and Molecular Biology. 3:763–770.

Monteiro P, Gosselin A, Wacleche V, El-Far M, Said E, Kared H, Grandvaux N, Boulassel M, Routy J, Ancuta P. 2011. Memory CCR6+CD4+ T Cells Are Preferential Targets for Productive HIV Type 1 Infection Regardless of Their Expression of Integrin 7. The Journal of Immunology 186:4618-4630.

Mori M, Wichukchinda N, Miyahara R, Rojanawiwat A, Pathipvanich P, Maekawa T, Miura T, Goulder P, Yasunami M, Ariyoshi K, Sawanpanyalert P. 2014. HLA-B*35: 05 is a protective allele with a unique structure among HIV-1 CRF01_AE-infected Thais, in whom the B*57 frequency is low. AIDS 28:959-967.

Mothe B, Llano A, Ibarrondo J, Daniels M, Miranda C, Zamarreño J, Bach V, Zuniga R, Pérez-Álvarez S, Berger C, Puertas M, Martinez-Picado J, Rolland M, Farfan M, Szinger J, Hildebrand W, Yang O, Sanchez-Merino V, Brumme C, Brumme Z, Heckerman D, Allen T, Mullins J, Gómez G, Goulder P, Walker B, Gatell J, Clotet B, Korber B, Sanchez J, Brander C. 2011. Definition of the viral targets of protective HIV-1-specific T cell responses. Journal of Translational Medicine 9:208.

Motulsky A, Vandepitte J, Fraser G. Population genetic studies in the Congo. 1. Glucose-6-phosphate dehydrogenase deficiency, hemoglobin S, and malaria. 1966. American Journal of Human Genetics 18;514–537.

Mphatswe W, Blanckenberg N, Tudor-Williams G, Prendergast A, Thobakgale C, Mkhwanazi N, McCarthy N, Walker B, Kiepiela P, Goulder P. 2007. High frequency of rapid immunological progression in African infants infected in the era of perinatal HIV prophylaxis. AIDS 21:1253-1261.

Mudd P, Matins M, Ericsen A, Tully D, Power K, Bean A, Piaskowski S, Duan L, Seese A, Gladden A, Weisgrau K, Furlott J, Kim Y, Veloso de Santana M, Rakasz E, Capuano S III, Wilson N, Bonaldo M, Galler R, Allison D, Piatak M, Haase A, Lifson J, Allen T, Watkins D. 2012. Vaccine induced CD8+ T cells control AIDS virus replication. Nature 491: 129-133.

Muenchhoff M, Prendergast A, Goulder P. 2014. Immunity to HIV in Early Life. Frontiers in Immunology 5:391.

Muenchhoff M, Adland E, Karimanzira O, Crowther C, Pace M, Csala A, Leitman E, Moonsamy A, McGregor C, Hurst J, Groll A, Mori M, Sinmyee S, Thobakgale C, Tudor-Williams G, Prendergast A, Kloverpris H, Roider J, Leslie A, Shingadia D, Brits T, Daniels S, Frater J, Willberg C, Walker B, Ndungu T, Jooste P, Moore P, Morris L, Goulder P. 2016. Nonprogressing

HIV-infected children share fundamental immunological features of nonpathogenic SIV infection. Science Translational Medicine 8:358ra125.

Müller C, Engler-Blum G, Gekeler V, Steiert I, Weiss E, Schmidt H. 1989. Genetic and serological heterogeneity of the supertypic HLA-B locus specificities Bw4 and Bw6. Immunogenetics 30:200-207.

Murakoshi H, Akahoshi T, Koyanagi M, Chikata T, Naruto T, Maruyama R, Tamura Y, Ishizuka N, Gatanaga H, Oka S, Takiguchi M. 2015. Clinical Control of HIV-1 by Cytotoxic T Cells Specific for Multiple Conserved Epitopes. Journal of Virology 89:5330-5339.

Murakoshi H, Koyanagi M, Chikata T, Rahman M, Kuse N, Sakai K, Gatanaga H, Oka S, Takiguchi M. 2016. Accumulation of Pol mutations selected by HLA-B*52:01-C*12:02 protective haplotype-restricted CTLs causes low plasma viral load due to low viral fitness of mutant viruses. Journal of Virology JVI.02082-16.

Murphy E, Korber B, Georges-Courbot M, You B, Pinter A, Cook D, Kieny M, Georges A, Mathiot C, Barré-Sinoussi F, Girard M. 1993. Diversity of V3 Region Sequences of Human Immunodeficiency Viruses Type 1 from the Central African Republic. AIDS Research and Human Retroviruses 9:997-1006.

Murphy K, Travers P, Walport M, Janeway C. 2007. Janeway's immunobiology, 7th ed. Garland Science, New York 196-199, 536.

Nabel G and Baltimore D. 1987. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. Nature 326:711-713.

Nahmias A, Weiss J, Yao X, Lee F, Kodsi R, Schanfield M, Matthews T, Bolognesi D, Durack D, Motulsky A, Kanki P, Essex M. 1986. Evidence For Human Infection with an HTLV III/LAV-Like Virus In Central Africa, 1959. The Lancet 327:1279-1280.

Naruto T, Gatanaga H, Nelson G, Sakai K, Carrington M, Oka S, Takiguchi M. 2012. 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 86:10870-10872.

Ndhlovu Z, Kamya P, Mewalal N, Kløverpris H, Nkosi T, Pretorius K, Laher F, Ogunshola F, Chopera D, Shekhar K, Ghebremichael M, Ismail N, Moodley A, Malik A, Leslie A, Goulder P, Buus S, Chakraborty A, Dong K, Ndung'u T, Walker B. 2015. Magnitude and Kinetics of CD8+ T Cell Activation during Hyperacute HIV Infection Impact Viral Set Point. Immunity 43:591-604.

Neefjes J, Momburg F, Hammerling G. 1993. Selective and ATP-dependent translocation of peptides by the MHC-encoded transporter. Science 261:769-771.

Ngumbela K, Day C, Mncube Z, Nair K, Ramduth D, Thobakgale C, Moodley E, Reddy S, de Pierres C, Mkhwanazi N, Bishop K, van der Stok M, Ismail N, Honeyborne I, Crawford H, Kavanagh D, Rousseau C, Nickle D, Mullins J, Heckerman D, Korber B, Coovadia H, Kiepiela P, Goulder P,

Walker B. 2008. Targeting of a CD8 T Cell Env Epitope Presented by HLA-B*5802 Is Associated with Markers of HIV Disease Progression and Lack of Selection Pressure. AIDS Research and Human Retroviruses 24:72-82.

Nguyen D and Hildreth J. 2000. Evidence for Budding of Human Immunodeficiency Virus Type 1 Selectively from Glycolipid-Enriched Membrane Lipid Rafts. Journal of Virology 74:3264-3272.

Noviello C, Lopez C, Kukull B, McNett H, Still A, Eccles J, Sloan R, Barklis E. 2011. Second-Site Compensatory Mutations of HIV-1 Capsid Mutations. Journal of Virology 85:4730-4738.

Nurzia E, Panimolle F, Cauli A, Mathieu A, Magnacca A, Paladini F, Sorrentino R, Fiorillo M. 2010. CD8+ T-cell mediated self-reactivity in HLA-B27 context as a consequence of dual peptide conformation. Clinical Immunology 135:476-482.

Oberlin E, Amara A, Bachelerie F, Bessia C, Virelizier J, Arenzana-Seisdedos F, Schwartz O, Heard J, Clark-Lewis I, Legler D, Loetscher M, Baggiolini M, Moser B. 1996. The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1. Nature 382:833-835.

Okulicz J, Marconi V, Landrum M, Wegner S, Weintrob A, Ganesan A, Hale B, Crum-Cianflone N, Delmar J, Barthel V, Quinnan G, Agan B, Dolan M. 2009. Clinical Outcomes of Elite Controllers, Viremic Controllers, and Long-Term Nonprogressors in the US Department of Defense HIV Natural History Study. The Journal of Infectious Diseases 200:1714-1723.

Olinger G, Saifuddin M, Spear G. 2000. CD4-Negative Cells Bind Human Immunodeficiency Virus Type 1 and Efficiently Transfer Virus to T Cells. Journal of Virology 74:8550-8557.

Ondondo B, Murakoshi H, Clutton G, Abdul-Jawad S, Wee E, Gatanaga H, Oka S, McMichael A, Takiguchi M, Korber B, Hanke T. 2016. Novel Conserved-region T-cell Mosaic Vaccine With High Global HIV-1 Coverage Is Recognized by Protective Responses in Untreated Infection. Molecular Therapy 24:832-842.

Ortmann B, Copeman J, Lehner P, Sadasivan B, Herberg J, Grandea A, Riddell S, Tampé R, Spies T, Trowsdale J, Cresswell P. 1997. A Critical Role for Tapasin in the Assembly and Function of Multimeric MHC Class I-TAP Complexes. Science 277:1306-1309.

Osmond D, Charlebois E, Lang W, Shiboski S, Moss A. 1994. Changes in AIDS Survival Time in Two San Francisco Cohorts of Homosexual Men, 1983 to 1993. JAMA: The Journal of the American Medical Association 271:1083.

Ott M, Geyer M, Zhou Q. 2011. The Control of HIV Transcription: Keeping RNA Polymerase II on Track. Cell Host & Microbe 10:426-435.

Oxenius A, Price D, Trkola A, Edwards C, Gostick E, Zhang H, Easterbrook P, Tun T, Johnson A, Waters A, Holmes E, Phillips R. 2004. Loss of Viral Control in Early HIV-1 Infection Is Temporally Associated with Sequential Escape from CD8+T Cell Responses and Decrease in HIV–1–Specific

CD4+and CD8+T Cell Frequencies. The Journal of Infectious Diseases 190:713-721.

Paiardini M, Cervasi B, Reyes-Aviles E, Micci L, Ortiz A, Chahroudi A, Vinton C, Gordon S, Bosinger S, Francella N, Hallberg P, Cramer E, Schlub T, Chan M, Riddick N, Collman R, Apetrei C, Pandrea I, Else J, Munch J, Kirchhoff F, Davenport M, Brenchley J, Silvestri G. 2011. Low levels of SIV infection in sooty mangabey central memory CD4+ T cells are associated with limited CCR5 expression. Nature Medicine 17:830-836.

Pantaleo G, Graziosi C, Fauci A. 1995. New Concepts in the Immunopathogenesis of HIV Infection. Annual Review of Immunology 13:487-512.

Pantaleo G and Fauci A. 1996. Immunopathogenesis of HIV Infection. Annual Review of Microbiology 50:1-854.

Parham P, Norman P, Abi-Rached L, Guethlein L. 2012. Human-specific evolution of killer cell immunoglobulin-like receptor recognition of major histocompatibility complex class I molecules. Philosophical Transactions of the Royal Society B: Biological Sciences 367:800-811.

Parsons M, Wren L, Isitman G, Navis M, Stratov I, Bernard N, Kent S. 2012. HIV Infection Abrogates the Functional Advantage of Natural Killer Cells Educated through KIR3DL1/HLA-Bw4 Interactions To Mediate Anti-HIV Antibody-Dependent Cellular Cytotoxicity. Journal of Virology 86:4488-4495.

Patarca R and Haseltine W. 1985. A major retroviral core protein related to EPA and TIMP. Nature 318:390-390.

Patterson S and Knight S. 1987. Susceptibility of Human Peripheral Blood Dendritic Cells to Infection by Human Immunodeficiency Virus. Journal of General Virology 68:1177-1181.

Pauwels R, De Clercq E, Desmyter J, Balzarini J, Goubau P, Herdewijn P, Vanderhaeghe H, Vandeputte M. 1987. Sensitive and rapid assay on MT-4 cells for detection of antiviral compounds against the AIDS virus. Journal of Virological Methods 16:171-185.

Payne R, Kløverpris H, Sacha J, Brumme Z, Brumme C, Buus S, Sims S, Hickling S, Riddell L, Chen F, Luzzi G, Edwards A, Phillips R, Prado J, Goulder P. 2010. Efficacious Early Antiviral Activity of HIV Gag- and Pol-Specific HLA-B*2705-Restricted CD8+T Cells. Journal of Virology 84:10543-10557.

Payne R, Muenchhoff M, Mann J, Roberts H, Matthews P, Adland E, Hempenstall A, Huang K, Brockman M, Brumme Z, Sinclair M, Miura T, Frater J, Essex M, Shapiro R, Walker B, Ndung'u T, McLean A, Carlson J, Goulder P. 2014. Impact of HLA-driven HIV adaptation on virulence in populations of high HIV seroprevalence. Proceedings of the National Academy of Sciences 111:E5393-E5400.

Perelson A, Neumann A, Markowitz M, Leonard J, Ho D. 1996. HIV-1 Dynamics in Vivo: Virion Clearance Rate, Infected Cell Life-Span, and Viral Generation Time. Science 271:1582-1586.

Pereyra F, Addo M, Kaufmann D, Liu Y, Miura T, Rathod A, Baker B, Trocha A, Rosenberg R, Mackey E, Ueda P, Lu Z, Cohen D, Wrin T, Petropoulos C, Rosenberg E, Walker B. 2008. Genetic and Immunologic Heterogeneity among Persons Who Control HIV Infection in the Absence of Therapy. The Journal of Infectious Diseases 197:563-571.

Pereyra F, Palmer S, Miura T, Block B, Wiegand A, Rothchild A, Baker B, Rosenberg R, Cutrell E, Seaman M, Coffin J, Walker B. 2009. Persistent Low-Level Viremia in HIV-1 Elite Controllers and Relationship to Immunologic Parameters. The Journal of Infectious Diseases 200:984-990.

Pereyra F, et al. The International HIV Controllers Study, Writing team. 2010. The Major Genetic Determinants of HIV-1 Control Affect HLA Class I Peptide Presentation. Science 330(6010):1551-1557.

Pereyra F, Heckerman D, Carlson J, Kadie C, Soghoian D, Karel D, Goldenthal A, Davis O, DeZiel C, Lin T, Peng J, Piechocka A, Carrington M, Walker B. 2014. HIV Control Is Mediated in Part by CD8+ T-Cell Targeting of Specific Epitopes. Journal of Virology 88:12937-12948.

Persaud D, Gay H, Ziemniak C, Chen Y, Piatak M, Chun T, Strain M, Richman D, Luzuriaga K. 2013. Absence of Detectable HIV-1 Viremia after Treatment Cessation in an Infant. New England Journal of Medicine 369:1828-1835.

Peterlin B and Trono D. 2003. Hide, shield and strike back: how HIV-infected cells avoid immune eradication. Nature Reviews Immunology 3:97-107.

Petrova M, van den Broek M, Balzarini J, Vanderleyden J, Lebeer S. 2013. Vaginal microbiota and its role in HIV transmission and infection. FEMS Microbiology Reviews 37:762-792.

Phillips R, Rowland-Jones S, Nixon D, Gotch F, Edwards J, Ogunlesi A, Elvin J, Rothbard J, Bangham C, Rizza C, McMichael A. 1991. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. Nature 354:453-459.

Pillay T, Zhang H, Drijfhout J, Robinson N, Brown H, Khan M, Moodley J, Adhikari M, Pfafferott K, Feeney M, St. John A, Holmes E, Coovadia H, Klenerman P, Goulder P, Phillips R. 2005. Unique Acquisition of Cytotoxic T-Lymphocyte Escape Mutants in Infant Human Immunodeficiency Virus Type 1 Infection. Journal of Virology 79:12100-12105.

Plantier J, Leoz M, Dickerson J, De Oliveira F, Cordonnier F, Lemée V, Damond F, Robertson D, Simon F. 2009. A new human immunodeficiency virus derived from gorillas. Nature Medicine 15:871-872.

Pollard V and Malim M. 1998. The HIV-1 Rev Protein. Annual Review of Microbiology 52:491-532.

Poropatich K and Sullivan D. 2009. Human immunodeficiency virus type 1 long-term non-progressors: the viral, genetic and immunological basis for disease non-progression. Journal of General Virology 92:247-268.

Prado J, Prendergast A, Thobakgale C, Molina C, Tudor-Williams G, Ndung'u T,

Walker B, Goulder P. 2009. Replicative Capacity of Human Immunodeficiency Virus Type 1 Transmitted from Mother to Child Is Associated with Pediatric Disease Progression Rate. Journal of Virology 84:492-502.

Prendergast A, Mphatswe W, Tudor-Williams G, Rakgotho M, Pillay V, Thobakgale C, McCarthy N, Morris L, Walker B, Goulder P. 2008. Early virological suppression with three-class antiretroviral therapy in HIV-infected African infants. AIDS 22:1333-1343.

Prendergast A, Klenerman P, Goulder P. 2012. The impact of differential antiviral immunity in children and adults. Nature Review Immunology 12:636-648.

Preston B, Poiesz B, Loeb L. 1988. Fidelity of HIV-1 reverse transcriptase. Science 242:1168-1171.

Price D, Goulder P, Klenerman P, Sewell A, Easterbrook P, Troop M, Bangham C, Phillips R. 1997. Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection. Proceedings of the National Academy of Sciences 94:1890-1895.

Price D. 2000. P-TEFb, a Cyclin-Dependent Kinase Controlling Elongation by RNA Polymerase II. Molecular and Cellular Biology 20:2629-2634.

Price A, Jacques D, McEwan W, Fletcher A, Essig S, Chin J, Halambage U, Aiken C, James L. 2014. Host Cofactors and Pharmacologic Ligands Share an Essential Interface in HIV-1 Capsid That Is Lost upon Disassembly. PLoS Pathogens 10:e1004459.

Prince J, Claiborne D, Carlson J, Schaefer M, Yu T, Lahki S, Prentice H, Yue L, Vishwanathan S, Kilembe W, Goepfert P, Price M, Gilmour J, Mulenga J, Farmer P, Derdeyn C, Tang J, Heckerman D, Kaslow R, Allen S, Hunter E. 2012. Role of Transmitted Gag CTL Polymorphisms in Defining Replicative Capacity and Early HIV-1 Pathogenesis. PLoS Pathogen 8:e1003041.

Poropatich K and Sullivan D. 2010. Human immunodeficiency virus type 1 long-term non-progressors: the viral, genetic and immunological basis for disease non-progression. Journal of General Virology 92:247-268.

Qi Y, Martin M, Gao X, Jacobson L, Goedert J, Buchbinder S, Kirk G, O'Brien S, Trowsdale J, Carrington M. 2006. KIR/HLA Pleiotropism: Protection against Both HIV and Opportunistic Infections. PLoS Pathogens 2:e79.

Quigley M, Pereyra F, Nilsson B, Porichis F, Fonseca C, Eichbaum Q, Julg B, Jesneck J, Brosnahan K, Imam S, Russell K, Toth I, Piechocka-Trocha A, Dolfi D, Angelosanto J, Crawford A, Shin H, Kwon D, Zupkosky J, Francisco L, Freeman G, Wherry E, Kaufmann D, Walker B, Ebert B, Haining W. 2010. Transcriptional analysis of HIV-specific CD8+ T cells shows that PD-1 inhibits T cell function by upregulating BATF. Nature Medicine 16:1147-1151.

Ragonnet-Cronin M, Shilaih M, Günthard H, Hodcroft E, Böni J, Fearnhill E, Dunn D, Yerly S, Klimkait T, Aubert V, Yang W, Brown A, Lycett S, Kouyos R, Brown A. 2016. A Direct Comparison of Two Densely Sampled HIV Epidemics: The UK and Switzerland. Scientific Reports 6:32251.

Rambaut A, Posada D, Crandall K, Holmes E. 2004. The causes and consequences of HIV evolution. Nature Reviews Genetics 5:52-61.

Ranasinghe S, Cutler S, Davis I, Lu R, Soghoian D, Qi Y, Sidney J, Kranias G, Flanders M, Lindqvist M, Kuhl B, Alter G, Deeks S, Walker B, Gao X, Sette A, Carrington M, Streeck H. 2013. Association of HLA-DRB1–restricted CD4+ T cell responses with HIV immune control. Nature Medicine 19:930-933.

Ranga U, Shankarappa R, Siddappa N, Ramakrishna L, Nagendran R, Mahalingam M, Mahadevan A, Jayasuryan N, Satishchandra P, Shankar S, Prasad V. 2004. Tat Protein of Human Immunodeficiency Virus Type 1 Subtype C Strains Is a Defective Chemokine. Journal of Virology 78:2586-2590.

Reed L and Muench H. 1938. A simple method of estimating fifty percent endpoints. American Journal of Epidemiology 27(3):493–497.

Reid S, McAdam S, Smith K, Klenerman P, O'Callaghan C, Harlos K, Jakobsen B, McMichael A, Bell J, Stuart D, Jones E. 1996. Antagonist HIV-1 Gag Peptides Induce Structural Changes in HLA B8. Journal of Experimental Medicine 184:2279-2286.

Reits E, Griekspoor A, Neijssen J, Groothuis T, Jalink K, van Veelen P, Janssen H, Calafat J, Drijfhout J, Neefjes J. 2003. Peptide Diffusion, Protection, and Degradation in Nuclear and Cytoplasmic Compartments before Antigen Presentation by MHC Class I. Immunity 18:97-108.

Ren S, Yang G, Liu C, Zhang C, Shou Q, Yu S, Li W, Su X. 2012. Association between HLA-A and -B polymorphisms and susceptibility to Henoch-Schönlein purpura in Han and Mongolian children from Inner Mongolia. Genetics and Molecular Research 11:221-228.

Ribas-Silva R, Ribas A, Ferreira E, Silveira T, Borelli S. 2015. Association between HLA-C*04 and American cutaneous leishmaniasis in endemic region of southern Brazil. Genetics and Molecular Research 14:14929-14935.

Richman D, Margolis D, Delaney M, Greene W, Hazuda D, Pomerantz R. 2009. The Challenge of Finding a Cure for HIV Infection. Science 323:1304-1307.

Rihn S, Wilson S, Loman N, Alim M, Bakker S, Bhella D, Gifford R, Rixon F, Bieniasz P. 2013. Extreme Genetic Fragility of the HIV-1 Capsid. PLoS Pathogens 9:e1003461.

Roach J, Glusman G, Smit A, Huff C, Hubley R, Shannon P, Rowen L, Pant K, Goodman N, Bamshad M, Shendure J, Drmanac R, Jorde L, Hood L, Galas D. 2010. Analysis of Genetic Inheritance in a Family Quartet by Whole-Genome Sequencing. Science 328:636-639.

Roberts H, Hurst J, Robinson N, Brown H, Flanagan P, Vass L, Fidler S, Weber J, Babiker A, Phillips R, McLean A, Frater J. 2015. Structured Observations Reveal Slow HIV-1 CTL Escape. PLOS Genetics 11:e1004914.

Robertson D, Anderson J, Bradac J. 2000. HIV-1 Nomenclature Proposal. Science

288:55d-55.

Robinson H. 2002. Vaccines: New Hope For An Aids Vaccine. Nature Reviews Immunology 2:239-250.

Robinson J, Mistry K, McWilliam H, Lopez R, Parham P, Marsh S. 2011. The IMGT/HLA database. Nucleic Acids Resserch. 39(Database issue):D1171-6.

Rock K, Gramm C, Rothstein L, Clark K, Stein R, Dick L, Hwang D, Goldberg A. 1994. Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules. Cell 78:761-771.

Rodrigo A, Shpaer E, Delwart E, Iversen A, Gallo M, Brojatsch J, Hirsch M, Walker B, Mullins J. 1999. Coalescent estimates of HIV-1 generation time in vivo. Proceedings of the National Academy of Sciences 96:2187-2191.

Rolland M, Manocheewa S, Swain J, Lanxon-Cookson E, Kim M, Westfall D, Larsen B, Gilbert P, Mullins J. 2013. HIV-1 Conserved-Element Vaccines: Relationship between Sequence Conservation and Replicative Capacity. Journal of Virology 87:5461-5467.

Sacha J, Chung C, Rakasz E, Spencer S, Jonas A, Bean A, Lee W, Burwitz B, Stephany J, Loffredo J, Allison D, Adnan S, Hoji A, Wilson N, Friedrich T, Lifson J, Yang O, Watkins D. 2007a. Gag-Specific CD8+ T Lymphocytes Recognize Infected Cells before AIDS-Virus Integration and Viral Protein Expression. The Journal of Immunology 178:2746-2754.

Sacha J, Chung C, Reed J, Jonas A, Bean A, Spencer S, Lee W, Vojnov L, Rudersdorf R, Friedrich T, Wilson N, Lifson J, Watkins D. 2007b. Pol-Specific CD8+ T Cells Recognize Simian Immunodeficiency Virus-Infected Cells Prior to Nef-Mediated Major Histocompatibility Complex Class I Downregulation. Journal of Virology 81:11703-11712.

Saez-Cirion A, Lacabaratz C, Lambotte O, Versmisse P, Urrutia A, Boufassa F, Barre-Sinoussi F, Delfraissy J, Sinet M, Pancino G, Venet A. 2007. HIV controllers exhibit potent CD8 T cell capacity to suppress HIV infection ex vivo and peculiar cytotoxic T lymphocyte activation phenotype. Proceedings of the National Academy of Sciences 104:6776-6781.

Sáez-Cirión A, Bacchus C, Hocqueloux L, Avettand-Fenoel V, Girault I, Lecuroux C, Potard V, Versmisse P, Melard A, Prazuck T, Descours B, Guergnon J, Viard J, Boufassa F, Lambotte O, Goujard C, Meyer L, Costagliola D, Venet A, Pancino G, Autran B, Rouzioux C. 2013. Post-Treatment HIV-1 Controllers with a Long-Term Virological Remission after the Interruption of Early Initiated Antiretroviral Therapy ANRS VISCONTI Study. PLoS Pathogens 9:e1003211.

Sakai K, Chikata T, Brumme Z, Brumme C, Gatanaga H, Oka S, Takiguchi M. 2015. Lack of a significant impact of Gag-Protease-mediated HIV-1 replication capacity on clinical parameters in treatment-naive Japanese individuals. Retrovirology 12:98-100.

Salazar-Gonzalez J, Salazar M, Keele B, Learn G, Giorgi E, Li H, Decker J, Wang

S, Baalwa J, Kraus M, Parrish N, Shaw K, Guffey M, Bar K, Davis K, Ochsenbauer-Jambor C, Kappes J, Saag M, Cohen M, Mulenga J, Derdeyn C, Allen S, Hunter E, Markowitz M, Hraber P, Perelson A, Bhattacharya T, Haynes B, Korber B, Hahn B, Shaw G. 2009. Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. Journal of Experimental Medicine 206(6):1273.

Saitoh T, Komano J, Saitoh Y, Misawa T, Takahama M, Kozaki T, Uehata T, Iwasaki H, Omori H, Yamaoka S, Yamamoto N, Akira S. 2012. Neutrophil Extracellular Traps Mediate a Host Defense Response to Human Immunodeficiency Virus-1. Cell Host & Microbe 12:109-116.

Salemi M. 2013. The intra-host evolutionary and population dynamics of human immunodeficiency virus type 1: a phylogenetic perspective. Infectious Disease Reports 5:3.

Samji H, Cescon A, Hogg R, Modur S, Althoff K, Buchacz K, Burchell A, Cohen M, Gebo K, Gill M, Justice A, Kirk G, Klein M, Korthuis P, Martin J, Napravnik S, Rourke S, Sterling T, Silverberg M, Deeks S, Jacobson L, Bosch R, Kitahata M, Goedert J, Moore R, Gange S. 2013. Closing the Gap: Increases in Life Expectancy among Treated HIV-Positive Individuals in the United States and Canada. PLoS ONE 8:e81355.

Sandler N, Bosinger S, Estes J, Zhu R, Tharp G, Boritz E, Levin D, Wijeyesinghe S, Makamdop K, del Prete G, Hill B, Timmer J, Reiss E, Yarden G, Darko S, Contijoch E, Todd J, Silvestri G, Nason M, Norgren Jr R, Keele B, Rao S, Langer J, Lifson J, Schreiber G, Douek D. 2014. Type I interferon responses in rhesus macaques prevent SIV infection and slow disease progression. Nature 511:601-605.

Sanjanwala B, Draghi M, Norman P, Guethlein L, Parham P. 2008. Polymorphic Sites Away from the Bw4 Epitope That Affect Interaction of Bw4+ HLA-B with KIR3DL1. The Journal of Immunology 181:6293-6300.

Santiago M, Rodenburg M, Kamenya S, Bibollet-Ruche F, Gao F, Bailes E, Meleth S, Soong S, Kilby J, Moldoveanu Z, Fahey B, Muller M, Ayouba A, Nerrienet E, McClure H, Heeney J, Pusey A, Collins D, Boesch C, Wrangham R, Goodall J, Sharp P, Shaw G, Hahn B. 2002. SIVcpz in Wild Chimpanzees. Science 295:465-465.

Sather D, Armann J, Ching L, Mavrantoni A, Sellhorn G, Caldwell Z, Yu X, Wood B, Self S, Kalams S, Stamatatos L. 2008. Factors Associated with the Development of Cross-Reactive Neutralizing Antibodies during Human Immunodeficiency Virus Type 1 Infection. Journal of Virology 83:757-769.

Schacker T, Hughes J, Shea T, Coombs R, Corey L. 1998. Biological and Virologic Characteristics of Primary HIV Infection. Annals of Internal Medicine 128:613.

Scheid J, Mouquet H, Feldhahn N, Seaman M, Velinzon K, Pietzsch J, Ott R, Anthony R, Zebroski H, Hurley A, Phogat A, Chakrabarti B, Li Y, Connors M, Pereyra F, Walker B, Wardemann H, Ho D, Wyatt R, Mascola J, Ravetch J, Nussenzweig M. 2009. Broad diversity of neutralizing

antibodies isolated from memory B cells in HIV-infected individuals. Nature 458:636-640.

Schmid B, Keşmir C, de Boer R. 2009. Correction: The Specificity and Polymorphism of the MHC Class I Prevents the Global Adaptation of HIV-1 to the Monomorphic Proteasome and TAP. PLoS ONE 4.

Schmitz J, Kuroda M, Santra S, Sasseville V, Simon M, Lifton M, Racz P, Tenner-Racz K, Dalesandro M, Scallon B, Ghrayeb J, Forman M, Montefiori D, Rieber E, Letvin N, Reimann K. 1999. Control of Viremia in Simian Immunodeficiency Virus Infection by CD8+ Lymphocytes. Science 283:857-860.

Schneidewind A, Brockman M, Yang R, Adam R, Li B, Le Gall S, Rinaldo C, Craggs S, Allgaier R, Power K, Kuntzen T, Tung C, LaBute M, Mueller S, Harrer T, McMichael A, Goulder P, Aiken C, Brander C, Kelleher A, Allen T. 2007. Escape from the Dominant HLA-B27-Restricted Cytotoxic T-Lymphocyte Response in Gag Is Associated with a Dramatic Reduction in Human Immunodeficiency Virus Type 1 Replication. Journal of Virology 81:12382-12393.

Schneidewind A, Brockman M, Sidney J, Wang Y, Chen H, Suscovich T, Li B, Adam R, Allgaier R, Mothe B, Kuntzen T, Oniangue-Ndza C, Trocha A, Yu X, Brander C, Sette A, Walker B, Allen T. 2008. Structural and Functional Constraints Limit Options for Cytotoxic T-Lymphocyte Escape in the Immunodominant HLA-B27-Restricted Epitope in Human Immunodeficiency Virus Type 1 Capsid. Journal of Virology 82:5594-5605.

Scott-Algara D, Truong L, Versmisse P, David A, Luong T, Nguyen N, Theodorou I, Barre-Sinoussi F, Pancino G. 2003. Cutting Edge: Increased NK Cell Activity in HIV-1-Exposed but Uninfected Vietnamese Intravascular Drug Users. The Journal of Immunology 171:5663-5667.

Serwold T, Gonzalez F, Kim J, Jacob R, Shastri N. 2002. ERAAP customizes peptides for MHC class I molecules in the endoplasmic reticulum. Nature 419:480-483.

Sharp P, M Robertson, D Gao, Hahn, B. 1994. Origins and diversity of human immunodeficiency viruses. AIDS 8:S27–S42.

Shearer W, Rosenblatt H, Gelman R, Oyomopito R, Plaeger S, Stiehm E, Wara D, Douglas S, Luzuriaga K, McFarland E, Yogev R, Rathore M, Levy W, Graham B, Spector S. 2003. Lymphocyte subsets in healthy children from birth through 18 years of age. Journal of Allergy and Clinical Immunology 112:973-980.

Sheehan N. 2010. HLA-B*27: what's new? Rheumatology (Oxford) 49(4): 621-631.

Shisana O, Connolly C, Rehle T, Mehtar S, Dana P. 2008. HIV risk exposure among South African children in public health facilities. AIDS Care 20:755-763.

Siekevitz M, Josephs S, Dukovich M, Peffer N, Wong-Staal F, Greene W. 1987. Activation of the HIV-1 LTR by T cell mitogens and the trans-activator protein of HTLV-I. Science 238:1575-1578.

Sidney J, Peters B, Frahm N, Brander C, Sette A. 2008. HLA class I supertypes: a revised and updated classification. BMC Immunology 9:1.

Silvestri G, Sodora D, Koup R, Paiardini M, O'Neil S, McClure H, Staprans S, Feinberg M. 2003. Nonpathogenic SIV Infection of Sooty Mangabeys Is Characterized by Limited Bystander Immunopathology Despite Chronic High-Level Viremia. Immunity 18:441-452.

Simmons R, Scully E, Groden E, Arnold K, Chang J, Lane K, Lifson J, Rosenberg E, Lauffenburger D, Altfeld M. 2013. HIV-1 infection induces strong production of IP-10 through TLR7/9-dependent pathways. AIDS 27:2505-2517.

Sluis-Cremer N, Wainberg M, Schinazi R. 2015. Resistance to reverse transcriptase inhibitors used in the treatment and prevention of HIV-1 infection. Future Microbiology 10:1773-1782.

Sodora D, Gettie A, Miller C, Marx P. 1998. Vaginal transmission of SIV: assessing infectivity and hormonal influences in macaques inoculated with cell-free and cell-associated viral stocks. AIDS Reserarch Human Retroviruses 14:S119–S123.

Soghoian D, Jessen H, Flanders M, Sierra-Davidson K, Cutler S, Pertel T, Ranasinghe S, Lindqvist M, Davis I, Lane K, Rychert J, Rosenberg E, Piechocka-Trocha A, Brass A, Brenchley J, Walker B, Streeck H. 2012. HIV-Specific Cytolytic CD4 T Cell Responses During Acute HIV Infection Predict Disease Outcome. Science Translational Medicine 4:123ra25.

Sokolskaja E, Sayah D, Luban J. 2004. Target Cell Cyclophilin A Modulates Human Immunodeficiency Virus Type 1 Infectivity. Journal of Virology 78:12800-12808.

Song Y, Meng Z, Xing H, Ruan Y, Li X, Xin R, Ma P, Peng H, Shao Y. 2007. Analysis of HIV-1 CRF07_BC gag p6 sequences indicating novel deletions in the central region of p6. Archives of Virology 152:1553-1558.

Soogoor M and Daar E. 2005. Primary HIV-1 infection: Diagnosis, pathogenesis, and treatment. Current Infectious Disease Reports 7:147-153.

Soto-Vega E, García-Muñoz R, Richaud-Patin Y, Zúñiga-Ramos J, Crispín J, Díaz-Jouanen E, Flores-Suárez L, Llorente L, Granados J. 2004. Class I and class II MHC polymorphisms in Mexican patients with Behçet's disease. Immunology Letters 93:211-215.

Spiegel H, Herbst H, Niedobitek G, Foss HD, Stein H. 1992. Follicular Dendritic Cells Are A Major Reservoir For Human Immunodeficiency Virus Type 1 In Lymphoid Tissues Facilitating Infection Of CD4+ T-Helper Cells. Amarican Journal Pathology. 140(1):15–22.

Sperduto A, Bryson Y, Chen I. 1993. Increased Susceptibility of Neonatal Monocyte/Macrophages to HIV-1 Infection. AIDS Research and Human Retroviruses 9:1277-1285.

Stacey A, Norris P, Qin L, Haygreen E, Taylor E, Heitman J, Lebedeva M,

DeCamp A, Li D, Grove D, Self S, Borrow P. 2009. Induction of a Striking Systemic Cytokine Cascade prior to Peak Viremia in Acute Human Immunodeficiency Virus Type 1 Infection, in Contrast to More Modest and Delayed Responses in Acute Hepatitis B and C Virus Infections. Journal of Virology 83:3719-3733.

Stephens H, Klaythong R, Sirikong M, Vaughn D, Green S, Kalayanarooj S, Endy T, Libraty D, Nisalak A, Innis B, Rothman A, Ennis F, Chandanayingyong D. 2002. HLA-A and -B allele associations with secondary dengue virus infections correlate with disease severity and the infecting viral serotype in ethnic Thais. Tissue Antigens 60:309-318.

Strack B, Calistri A, Accola M, Palu G, Gottlinger H. 2000. A role for ubiquitin ligase recruitment in retrovirus release. Proceedings of the National Academy of Sciences 97:13063-13068.

Strambio-de-Castillia C and Hunter E. 1992. Mutational analysis of the major homology region of Mason-Pfizer monkey virus by use of saturation mutagenesis. Journal of Virology 66(12): 7021-7032.

Sun H, Hung C, Lin P, Chang S, Yang C, Chang S, Chang S. 2007. Incidence of abacavir hypersensitivity and its relationship with HLA-B*5701 in HIV-infected patients in Taiwan. Journal of Antimicrobial Chemotherapy 60:599-604.

Sunshine J, Kim M, Carlson J, Heckerman D, Czartoski J, Migueles S, Maenza J, McElrath M, Mullins J, Frahm N. 2014. Increased Sequence Coverage through Combined Targeting of Variant and Conserved Epitopes Correlates with Control of HIV Replication. Journal of Virology 88:1354-1365.

Tang J, Costello C, Keet I, Rivers C, Leblanc S, Karita E, Allen S, Kaslow R. 1999. HLA Class I Homozygosity Accelerates Disease Progression in Human Immunodeficiency Virus Type 1 Infection. AIDS Research and Human Retroviruses 15:317-324.

Tang C, Ndassa Y, Summers M. 2002. Structure of the N-terminal 283-residue fragment of the immature HIV-1 Gag polyprotein. Nature Structural Biology. 9:537–543.

Tang J, Malhotra R, Song W, Brill I, Hu L, Farmer P, Mulenga J, Allen S, Hunter E, Kaslow R. 2010. Human Leukocyte Antigens and HIV Type 1 Viral Load in Early and Chronic Infection: Predominance of Evolving Relationships. PLoS ONE 5:e9629.

Teixeira S, de Sá N, Campos D, Coelho A, Guimarães M, Leite T, Veloso V, Morgado M. 2014. Association of the HLA-B*52 allele with non-progression to AIDS in Brazilian HIV-1-infected individuals. Genes and Immunity 15:256-262.

Tenzer S, Crawford H, Pymm P, Gifford R, Sreenu V, Weimershaus M, de Oliveira T, Burgevin A, Gerstoft J, Akkad N, Lunn D, Fugger L, Bell J, Schild H, van Endert P, Iversen A. 2014. HIV-1 Adaptation to Antigen Processing Results in Population-Level Immune Evasion and Affects Subtype Diversification. Cell Reports 7:448-463.

Terao C, Yoshifuji H, Mimori T. 2014. Recent advances in Takayasu arteritis.

International Journal of Rheumatic Diseases 17:238-247.

Terao C, Matsumura T, Yoshifuji H, Kirino Y, Maejima Y, Nakaoka Y, Takahashi M, Amiya E, Tamura N, Nakajima T, Origuchi T, Horita T, Matsukura M, Kochi Y, Ogimoto A, Yamamoto M, Takahashi H, Nakayamada S, Saito K, Wada Y, Narita I, Kawaguchi Y, Yamanaka H, Ohmura K, Atsumi T, Tanemoto K, Miyata T, Kuwana M, Komuro I, Tabara Y, Ueda A, Isobe M, Mimori T, Matsuda F. 2015. Brief Report: Takayasu Arteritis and Ulcerative Colitis: High Rate of Co-Occurrence and Genetic Overlap. Arthritis & Rheumatology 67:2226-2232.

LANL HIV Sequence Database (LosAlamos National Laboratory) In https://www.hiv.lanl.gov/. Accessed on 24 January 2017.

The Allele Frequency Net Database. Allele, haplotype and genotype frequencies in Worldwide Populations. In www.allelefrequencies.net. Accessed on 21 January 2017.

Thobakgale C, Ramduth D, Reddy S, Mkhwanazi N, de Pierres C, Moodley E, Mphatswe W, Blanckenberg N, Cengimbo A, Prendergast A, Tudor-Williams G, Dong K, Jeena P, Kindra G, Bobat R, Coovadia H, Kiepiela P, Walker B, Goulder P. 2007. Human Immunodeficiency Virus-Specific CD8+ T-Cell Activity Is Detectable from Birth in the Majority of In Utero-Infected Infants. Journal of Virology 81:12775-12784.

Thomas R, Apps R, Qi Y, Gao X, Male V, O'hUigin C, O'Connor G, Ge D, Fellay J, Martin J, Margolick J, Goedert J, Buchbinder S, Kirk G, Martin M, Telenti A, Deeks S, Walker B, Goldstein D, McVicar D, Moffett A, Carrington M. 2009. HLA-C cell surface expression and control of HIV/AIDS correlate with a variant upstream of HLA-C. Nature Genetics 41:1290-1294.

Tobin N and Aldrovandi G. 2013. Immunology of pediatric HIV infection. Immunological Reviews 254:143-169.

Tomiyama H, Miwa K, Shiga H, Moore Y, Oka S, Iwamoto A, Kaneko Y, Takiguchi M. 1997. Evidence of presentation of multiple HIV-1 cytotoxic T lymphocyte epitopes by HLA-B*3501 molecules that are associated with the accelerated progression of AIDS. Journal of Immunology 158(10):5026-5034.

Towers G, Hatziioannou T, Cowan S, Goff S, Luban J, Bieniasz P. 2003. Cyclophilin A modulates the sensitivity of HIV-1 to host restriction factors. Nature Medicine 9:1138-1143.

Trapani J and Smyth M. 2002. Functional significance of the perforin/granzyme cell death pathway. Nature Reviews Immunology 2:735-747.

Troyer R, McNevin J, Liu Y, Zhang S, Krizan R, Abraha A, Tebit D, Zhao H, Avila S, Lobritz M, McElrath M, Le Gall S, Mullins J, Arts E. 2009. Variable Fitness Impact of HIV-1 Escape Mutations to Cytotoxic T Lymphocyte (CTL) Response. PLoS Pathogens 5:e1000365.

Tsuchiya N, Pathipvanich P, Wichukchinda N, Rojanawiwat A, Auwanit W, Ariyoshi K, Sawanpanyalert P. 2014. Incidence and predictors of regimen-modification from first-line antiretroviral therapy in Thailand: a cohort study. BMC Infectious Diseases 14.

Tschachler E, Groh V, Popovic M, Mann D, Konrad K, Safai B, Eron L, DiMarzo Veronese F, Wolff K, Stingl G. 1987. Epidermal Langerhans Cells-A Target for HTLV-III/LAV Infection. Journal of Investigative Dermatology 88:233-237.

Turville S, Cameron P, Handley A, Lin G, Pöhlmann S, Doms R, Cunningham A. 2002. Diversity of receptors binding HIV on dendritic cell subsets. Nature Immunology 3:975-983.

UNAIDS 90-90-90, 2014. In http://www.unaids.org/en/resources/documents/2014/90-90-90. Accessed on 1 February 2017.

UNAIDS AIDS by the number, 2016. In http://www.unaids.org/en/resources/documents/2016/AIDS-by-the-numbers. Accessed on 1 February 2017.

UNAIDS Prevention Gap Report, 2016. In http://www.unaids.org/en/resources/documents/2016/prevention-gap. Accessed on 1 February 2017.

Valentine L and Watkins D. 2008. Relevance of studying T cell responses in SIV-infected rhesus macaques. Trends in Microbiology 16:605-611.

Vallari A, Holzmayer V, Harris B, Yamaguchi J, Ngansop C, Makamche F, Mbanya D, Kaptue L, Ndembi N, Gurtler L, Devare S, Brennan C. 2011. Confirmation of Putative HIV-1 Group P in Cameroon. Journal of Virology 85:1403-1407.

van Endert P, Saveanu L, Hewitt E, Lehner P. 2002. Powering the peptide pump: TAP crosstalk with energetic nucleotides. Trends in Biochemical Sciences 27:454-461.

van Kooten Niekerk N. 2006. The First 5 Years of the Family Clinic for HIV at Tygerberg Hospital: Family Demographics, Survival of Children and Early Impact of Antiretroviral Therapy. Journal of Tropical Pediatrics 52:3-11.

Vaz P, Pedro A, Le Bozec S, Macassa E, Salvador S, Biberfeld G, Blanche S, Andersson S. 2010. Nonvertical, nonsexual transmission of human immunodeficiency virus in children. Pediatric Infectious Disease Journal 29:271–274.

Vieillard V, Strominger J, Debre P. 2005. NK cytotoxicity against CD4+ T cells during HIV-1 infection: A gp41 peptide induces the expression of an NKp44 ligand. Proceedings of the National Academy of Sciences 102:10981-10986.

Vieillard V, Fausther-Bovendo H, Samri A, Debré P. 2010. Specific Phenotypic and Functional Features of Natural Killer Cells From HIV-Infected Long-Term Nonprogressors and HIV Controllers. JAIDS Journal of Acquired Immune Deficiency Syndromes 1 53(5):564–573.

Vijaya Lakshmi V, Rakh S, Anu Radha B, Hari Sai Priya V, Pantula V, Jasti S, Suman Latha G, Murthy K. 2006. Role of HLA-B51 and HLA-B52 in susceptibility to pulmonary tuberculosis. Infection, Genetics and Evolution 6:436-439.

Vingert B, Perez-Patrigeon S, Jeannin P, Lambotte O, Boufassa F, Lemaître F, Kwok W, Theodorou I, Delfraissy J, Thèze J, Chakrabarti L. 2010. HIV Controller CD4+ T Cells Respond to Minimal Amounts of Gag Antigen Due to High TCR Avidity. PLoS Pathogens 6:e1000780.

von Schwedler U, Stray K, Garrus J, Sundquist W. 2003. Functional Surfaces of the Human Immunodeficiency Virus Type 1 Capsid Protein. Journal of Virology 77:5439-5450.

Wang Y, Lv Z, Chu Y. 2015. HIV protease inhibitors: a review of molecular selectivity and toxicity. HIV/AIDS - Research and Palliative Care 95.

Walker B and Yu X. 2013. Unravelling the mechanisms of durable control of HIV-1. Nature Reviews Immunology 13:487-498.

Wei X, Ghosh S, Taylor M, Johnson V, Emini E, Seutsch P, Lifson J, Bonhoeffer S, Nowak M, Hahn B, Saag M, Shaw G. 1995. Viral dynamics in human immunodeficiency virus type 1 infection. Nature 373:117-122.

Wei P, Garber M, Fang S, Fischer W, Jones K. 1998. A Novel CDK9-Associated C-Type Cyclin Interacts Directly with HIV-1 Tat and Mediates Its High-Affinity, Loop-Specific Binding to TAR RNA. Cell 92:451-462.

Wherry E and Kurachi M. 2015. Molecular and cellular insights into T cell exhaustion. Nature Reviews Immunology 15:486-499.

Wills J and Craven R. 1991. Form, function, and use of retroviral Gag proteins. AIDS 5:639-654.

Wilen C, Tilton J, Doms R. 2012. HIV: Cell Binding and Entry. Cold Spring Harbor Perspectives in Medicine 2:a006866-a006866.

Wilson W, Braddock M, Adams S, Rathjen P, Kingsman S, Kingsman A. 1988. HIV expression strategies: Ribosomal frameshifting is directed by a short sequence in both mammalian and yeast systems. Cell 55:1159-1169.

Wilson J, Ogg G, Allen R, Davis C, Shaunak S, Downie J, Dyer W, Workman C, Sullivan J, McMichael A, Rowland-Jones S. 2000. Direct visualization of HIV-1-specific cytotoxic T lymphocytes during primary infection. AIDS 14:225-233.

Wright J, Brumme Z, Carlson J, Heckerman D, Kadie C, Brumme C, Wang B, Losina E, Miura T, Chonco F, van der Stok M, Mncube Z, Bishop K, Goulder P, Walker B, Brockman M, Ndung'u T. 2010. Gag-Protease-Mediated Replication Capacity in HIV-1 Subtype C Chronic Infection: Associations with HLA Type and Clinical Parameters. Journal of Virology 84:10820-10831.

Wright J, Naidoo V, Brumme Z, Prince J, Claiborne D, Goulder P, Brockman M, Hunter E, Ndung'u T. 2012. Impact of HLA-B*81-Associated Mutations in HIV-1 Gag on Viral Replication Capacity. Journal of Virology 86:3193-3199.

WHO. 2010. Consolidated ART guidelines 2010. June 2010. In http://www.who.int/hiv/pub/guidelines/arv2010/intro/executivesummary/en. Accessed on 23 January 2017.

WHO. 2013. Consolidated ART guidelines 2013. June 2013. In http://www.who.int/hiv/pub/guidelines/arv2013/intro/executivesummary/en. Accessed on 23 January 2017.

WHO. 2016. Consolidated guidelines on HIV prevention, diagnosis, treatment and care for key populations. July 2016. http://www.who.int/hiv/pub/guidelines/keypopulations-2016/en/. Accessed on 15 March 2017.

WHO. 2017. MTCT. In http://www.who.int/hiv/topics/mtct/about/en/. Accessed on 3 February 2017.

Yagita Y, Kuse N, Kuroki K, Gatanaga H, Carlson J, Chikata T, Brumme Z, Murakoshi H, Akahoshi T, Pfeifer N, Mallal S, John M, Ose T, Matsubara H, Kanda R, Fukunaga Y, Honda K, Kawashima Y, Ariumi Y, Oka S, Maenaka K, Takiguchi M. 2012. Distinct HIV-1 Escape Patterns Selected by Cytotoxic T Cells with Identical Epitope Specificity. Journal of Virology 87:2253-2263.

Yang K, Chen S, Shyr M, Lin P. 2009. High-resolution human leukocyte antigen (HLA) haplotypes and linkage disequilibrium of HLA-B and -C and HLA-DRB1 and -DQB1 alleles in a Taiwanese population. Human Immunology 70:269-276.

Yewdell J, Reits E, Neefjes J. 2003. Making sense of mass destruction: quantitating MHC class I antigen presentation. Nature Reviews Immunology 3:952-961.

Yokomaku Y, Miura H, Tomiyama H, Kawana-Tachikawa A, Takiguchi M, Kojima A, Nagai Y, Iwamoto A, Matsuda Z, Ariyoshi K. 2004. Impaired Processing and Presentation of Cytotoxic-T-Lymphocyte (CTL) Epitopes Are Major Escape Mechanisms from CTL Immune Pressure in Human Immunodeficiency Virus Type 1 Infection. Journal of Virology 78:1324-1332.

Yu X, Lichterfeld M, Chetty S, Williams K, Mui S, Miura T, Frahm N, Feeney M, Tang Y, Pereyra F, LaBute M, Pfafferott K, Leslie A, Crawford H, Allgaier R, Hildebrand W, Kaslow R, Brander C, Allen T, Rosenberg E, Kiepiela P, Vajpayee M, Goepfert P, Altfeld M, Goulder P, Walker B. 2007. Mutually Exclusive T-Cell Receptor Induction and Differential Susceptibility to Human Immunodeficiency Virus Type 1 Mutational Escape Associated with a Two-Amino-Acid Difference between HLA Class I Subtypes. Journal of Virology 81:1619-1631.

Zhao G, Perilla J, Yufenyuy E, Meng X, Chen B, Ning J, Ahn J, Gronenborn A, Schulten K, Aiken C, Zhang P. 2013. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. Nature 497:643-646.

Zhefeng M, Huiliang H, Chao Q, Jun S, Jianxin L, Xiaoyan Z, Jianqing X. 2011. Transmission of new CRF07_BC Strains with 7 amino acid deletion in Gag p6. Virology Journal 8:60.

Zhu T, Korber B, Nahmias A, Hooper E, Sharp P, Ho D. 1997. An African HIV-1 sequence from 1959 and implications for the origin of the epidemic. Nature 391; 594-597.

Zimmerli S, Harari A, Cellerai C, Vallelian F, Bart P, Pantaleo G. 2005. HIV-1-specific IFN- /IL-2-secreting CD8 T cells support CD4-independent proliferation of HIV-1-specific CD8 T cells. Proceedings of the National Academy of Sciences 102:7239-7244.

Zuniga R, Lucchetti A, Galvan P, Sanchez S, Sanchez C, Hernandez A, Sanchez H, Frahm N, Linde C, Hewitt H, Hildebrand W, Altfeld M, Allen T, Walker B, Korber B, Leitner T, Sanchez J, Brander C. 2006. Relative Dominance of Gag p24-Specific Cytotoxic T Lymphocytes Is Associated with Human Immunodeficiency Virus Control. Journal of Virology 80:3122-3125.