THE MULTIPLE ROLES OF HLA IN HIV IMMUNITY AND TREATMENT

Shayne Loubser

A thesis submitted to the faculty of Health Sciences, School of Pathology, University of Witwatersrand, Johannesburg, South Africa in fulfilment of the requirements for the Degree of Doctor of Philosophy in Virology

DECLARATION

I, Shayne Loubser, hereby declare that:

a) The research reported in this thesis, except where otherwise indicated is my original work.
b) This thesis has not been submitted for any degree or examination to any other University
c) This thesis does not contain other people's data, pictures, graphs or information, unless specifically stated as being of sourced from other people.
d) This thesis does not contain another person's writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then their words have been re-written but the general information attributed to them has been referenced.
e) Where I have reproduced a publication of which I am an author or co-author, I have indicated in detail which part of the publication was written by myself alone and have fully referenced such publications.
f) This thesis does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged, and the source being listed in the thesis.

Signed_____ Date_____

DEDICATION

I dedicate this manuscript to my son,

Liam Shaun Loubser

(born 16 September 2002)

Walk the path of Science, but

don’t be afraid to stop and admire the flowers

ACKNOWLEDGEMENTS

To the study participants, healthy or bearing the burden of HIV-1, my deepest thanks to you for allowing me the opportunity to present this work.

This work is based on the research supported by the South African Research Chairs Initiative of the Department of Science and Technology and the National Research Foundation of South Africa. Student support was received from the Poliomyelitis Research Foundation (PRF). Student and project support was received from the Kwa-Zulu Natal Research Institute for Tuberculosis and HIV (K-RITH).

To my supervisor, Prof. Caroline T. Tiemessen, thank you for adopting me as your student and for living up to all expectations as a committed, resourceful, capable and approachable mentor.

A special note of thanks to Prof. Clive M. Gray, Director of Immunopaedia, for nurturing my interest in Immunology and for initiating my PhD aspirations.

To my long time friend and colleague, Dr. Maria “Lucia” Paximadis, thank you for being a believer in me and for playing your invaluable part in my scientific journey, “the long, dark road to enlightenment”.

To Heather “HH” Hong, “Bellatrix” Picton, Sir Leonard Damelin, Di Schramm and “Sharry” Shalekoff, my “senior” colleagues, for assistance, support and conversations.

To my fellow PhDs, special thanks to Nikki “Nix” Gentle for sharing her work and to “Mama” Ria Lassauniere and Heather “HH” Hong, for student support and counseling.

To my colleagues at the laboratory, Avani, “Gemsquash” and Refilwe for spiritual upliftment.

Lastly, to the new recruits, Bianca, Yashini, Sizanani and Ntando, thank you for sharing the experience.

ABSTRACT

South Africa has a national HIV-1 prevalence of 12.2% and due to the large population size of 52.9 million, this equates to a 6.5 million people living with HIV-1. The high HIV-1 prevalence has warranted the scaling up of the national antiretroviral treatment program with over 2 million people accessing treatment since 2012. The population of South Africa is genetically diverse and consists of South African Black (SAB), South African Mixed ancestry (SAM), South African Caucasian (SAC) and South African Indian (SAI) populations that constitute 79.8%, 9.0%, 8.7% and 2.5% of the total population. The reasons for the disproportionate HIV-1 prevalence rates estimated for the different populations (15.0%, 3.1%, 0.3% and 0.8% for the SAB, SAM, SAC and SAI populations, respectively) are unknown; however, host genetic differences may be a contributory factor. *Human leukocyte antigen (HLA) class I* genes play a crucial role in the antiviral innate and adaptive immune response, since HLA class I proteins present antigenic peptides to $CD8^+$ Cytotoxic T cells (CTLs) of the adaptive immune system and are also involved in the innate immune response via interaction with killer-cell immunoglobulin-like receptors (KIRs) expressed by Natural Killer cells (NK cells).

Different *HLA class I* alleles associate with HIV-1 acquisition and disease progression and some alleles can precipitate adverse antiretroviral drug reactions. *HLA class I* allele frequencies in the SAC and SAB populations are known, however, limited information is available for the SAI and SAB populations. An immunogenetic characterisation study in healthy SAI (n=50) and SAM (n=50) individuals was performed to establish the prevalence of *HLA class I* alleles and other host genetic factors including *CCR5Δ32, KIR* genes, *HLA-*$Bw4^{80I}$*/*$Bw4^{80T}$*/C1/C2* ligands, HLA-B p2 signal peptide variants and *HLA-C 3'untranslated region +263* insertion/deletion variants. The results were compared to the available gene frequencies in healthy SAC and SAB populations and as expected the populations showed distinct differences in the prevalence of several of the genes studied. The comparative analyses are presented in Chapter 3.

In Chapter 4, a real-time AS-PCR assay was developed to detect specific *HLA class I* alleles associated with immune hypersensitivity reaction (IHR) to Abacavir (ABC) and Nevirapine (NVP). IHR can be fatal and develops in *HLA-B*57:01* carriers exposed to ABC

and in *HLA-B*35:05*, *-C*04* or *-C*08* carriers exposed to NVP. SAI and SAM populations were screened for these alleles and the frequencies compared to SAC and SAB populations. *HLA-B*57:01* was prevalent in 8.0%, 12.0% and 8.2% of SAM, SAI and SAC individuals, respectively and absent in the SAB population, while *HLA-B*35:05* was present in 6.0% and 1.0% in SAM and SAC populations, respectively and absent in SAC and SAB populations. *HLA-C*04* was present at frequencies of 22.0%, 24.0%, 18.6% and 23.7% in SAM, SAI, SAC and SAB populations, respectively, while *HLA-C*08* was present at frequencies of 10.0%, 8.0%, 6.9% and 10.2% in SAM, SAI, SAC and SAB populations, respectively. The number of HIV-1 infected South Africans potentially at risk of developing ABC IHR and NVP IHR, was estimated at 14,235 and 2,212,700 people, respectively. Since treatment of HIV-1 infected individuals is underway on a large scale in South Africa and both ABC and NVP are in current use, it was concluded that a substantial number of individuals could benefit from a pre-screening test before treatment is started.

The possibility that differences in the prevalence of *HLA class I* genes and other host genetic factors that have been shown associate with HIV-1 acquisition/transmission might influence HIV-1 prevalence rates in the four populations was explored in Chapter 5. Using new data generated for the SAI and SAM populations and available data for the SAC and SAB populations, a cumulative analysis of genes associated with HIV-1 acquisition and viral load setpoint (VLS) was performed and compared between the populations. Significant results regarding HIV-1 acquisition suggested that more SAB individuals may be susceptible compared to SAI individuals (42.8% vs. 14.0%, respectively; $p=0.0002$), while more SAI individuals may be resistant compared to SAB individuals (68.0% vs. 31.4%, respectively; $p<0.0001$). Similarly, significant results regarding genotypes associated with VLS suggested that more SAB individuals may have genotypes associated with high VLS compared to SAI individuals (59.7% vs. 30.0%, respectively; $p=0.0003$), while more SAI individuals may have a genotypes associated with low VLS compared to SAB individuals (48.0% vs. 13.2%, respectively; $p<0.0001$). Although not statistically significant, SAM and SAC individuals may also be less susceptible and more resistant to HIV-1 infection compared to SAB individuals. More SAM and SAC individuals have overall genotypes associated with a low VLS compared to SAB individuals, but the frequency of SAM and SAC individuals with overall genotypes associated with a high VLS was similar compared to SAB individuals. Based on extrapolation from these data analysed in healthy individuals, it was concluded that host genetic differences may plausibly influence HIV-1 prevalence in the four populations.

Next we examined the impact of viral mutagenesis on the role of HLA class I molecules in antiviral immunity studied in HIV-1 infected SAB individuals (Chapter 6). Viral evolution in the *gag* gene was studied in the context of two common *HLA class I* alleles present in the SAB population, *HLA-B*58:01* and *-B*58:02,* that associate with opposite clinical outcomes. *HLA-B*58:01* is a protective allele associated with low viral load and higher $CD4^+$ T cell count. *HLA-B*58:01* (n=5) and *HLA-B*58:02* (n=13) carriers, classified as progressors due to $CD4^+$ T cell counts declining to <350 cells/µl, were selected for the study. No significant differences were found between the two groups regarding viral load, $CD4^+$ T cell count or rates of $CD4^+$ T cell decline. Sequences of the *gag* gene were generated from viral RNA extracted from plasma samples at baseline, 6 months and 12 months that were available prior to initiation of treatment. Deduced Gag protein sequences were examined and CTL escape mutations were identified in HLA-B*58:01-restricted epitopes, while no probable HLA-B*58:02-restricted epitopes were identified. Escape mutations in other HLA class I restricted epitopes, compensatory mutations, NK cell escape mutations and mutations associated with reduced TRIM5α sensitivity were detected in both *HLA-B*58:01* and *HLA-B*58:02* carriers. PTAP duplications in $p6^{Gag}$ known to enhance Tsg101-mediated viral budding were detected in some sequences and a rare PYRE insertion in $p6^{Gag}$ known to mediate viral budding via an alternative ALIX-mediated pathway. As other investigators have shown, disease progression in *HLA-B*58:01* carriers associates with CTL escape mutations in *HLA-B*58:01*-restricted Gag epitopes and compensatory mutations, while in contrast, *HLA-B*58:02* carriers did not target Gag. However, additional escape from other antiviral mechanisms such as the additive effect of other *HLA class I* molecules, NK cell activation and TRIM5α sensitivity were observed. Furthermore, increased virus production mediated by compensatory mutations and insertions that enhance viral budding, infectivity and replication kinetics were identified. The study concluded that disease progression can occur in both *HLA-B*58:01* and *-B*58:02* carriers as a consequence of both immune escape mutations and enhancement of virus production.

Overall, the findings presented in this thesis regarding the crucial roles played by HLA class I molecules and other immunogenetic factors in mediating adaptive and innate immune responses to HIV-1, provides insight into the complex and variable nature of host genetics in different populations that govern viral susceptibility, resistance and infectivity as well as in treatment-related immune hypersensitivity reactions. Viral mutagenesis was shown to mediate the failure of HLA class I molecules to maintain antiviral immune responses.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

αEβ7	Alpha chain E beta chain 7
µl	microlitre
µm	micrometre (10^{-6}m)
ABC	Abacavir
AIDS	Acquired Immunodeficiency Syndrome
ALB	Albumin
ALIX	ALG-2-interacting protein X
APOBEC3G	Apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3G
ARV	AIDS-associated retrovirus
AS-PCR	Allele-specific polymerase chain reaction
BGB	*Beta-globin gene*
c	Centromeric region
CA	Capsid protein
CCR5	C-C chemokine receptor type 5
CCR5Δ32	C-C chemokine receptor type 5 32-bp deletion variant
CD	Cluster of differentiation
CD62L	Cluster of differentiation 62 ligand
CDC	USA Centers for Disease Control and Prevention
cDNA	Complementary DNA
CEPH	Centre de'Etude du Polymorphism Humain
CMV	Cytomegalovirus
CNS	Central nervous system
CRF	Circulating recombinant form
cryo-EM	Cryogenic electron microscopy
CTL	$CD8^+$ Cytotoxic T cell
CXCR4	C-X-C chemokine receptor type 4
Da	Dalton
DAP12	DNAX activation protein of 12 kDa
DC-SIGN	Dendritic Cell-Specific Intercellular Adhesion Molecule-3-Grabbing Non-Integrin
Del	Deletion
dGTP	Deoxy-guanosine triphosphate
DNA	Deoxy-ribonucleic acid
dNTP	Deoxy-nucleotide triphosphate
DOH	South African Department of Health
EM	Electron Microscopy
Env	Envelope protein
ESCRT	Endosomal sorting complexes required for transport
ESKOM	Elecricity supply commision
f	Allele frequency (n/2 x total no. of individuals)
FasL	Fas ligand
G	Guanine
G-A	Guanine to adenine
Gag	Group associated antigen

GALC	Galactosylceramidase gene
gp120	Envelope glycoprotein of 120 kDa
gp160	Envelope polyprotein of 160 kDa
gp41	Envelope glycoprotein of 41 kDa
HCP5	HLA complex protein 5
HCV	Hepatitis C virus
HIV-1	Human Immunodeficiency Virus Type 1
HIV-2	Human Immunodeficiency Virus Type 2
HLA	Human leukocyte antigen
HLA II-DR	human leukocyte antigen class II DR
HTLV	Human T-cell Leukaemia Virus
HTLV-III	Human T-cell Leukaemia Virus Type 3
ICAM-1	Intercellular adhesion molecule 1
ICTV	International Committee on Taxonomy of Viruses
IDAV	Immunodeficiency-associated Virus
IgM	Immunoglobulin M
IHR	Immune hypersensitivity reaction
IHWG	International Histocompatibility Working Group
IL-1	Interleukin 1
IL-6	Interleukin 6
IN	Integrase
Ins	Insertion
IPD	Immuno Polymorphism Database
IRIS	Immune reconstitution inflammatory syndrome
ITAM	lmmunoreceptor tyrosine-based activation motif
kb	Kilobase
kDa	KiloDalton
Kg	kilogram
KIR	Killer-cell immunoglobulin-like receptor
LAV	Lympadenopathy-associated Virus
LAV-II	Lympadenopathy-associated Virus Type 2
LEDGF	Lens epithelium-derived growth factor
LFA-1	Lymphocyte function-associated antigen 1
LNA	Locked nucleic acid
LRC	Leukocyte receptor complex
LTNP	Long-term non-progressor
LTR	Long terminal repeat
M	Methionine
MA	Matrix protein
MGB	Minor groove binding protein
MHC	Major histocompatibility complex
ml	Millilitre
MMWR	Morbidity and Mortality Weekly Report
mRNA	Messenger RNA
N	Asparagine
NC	Nucleoprotein

Nef	Negative regulatory factor
NFAT	Nuclear factor of activated T cells
NF-κB	Nuclear factor kappa B
NK cells	Natural Killer cells
N-linked	Asparagine-linked
nm	Nanometre (10^{-9}m)
NNRTI	Non-nucleoside reverse transcriptase inhibitor
NRTI	Nucleoside reverse transcriptase inhibitor
NVP	Nevirapine
ORF	Open reading frame
p.i.	Post-infection
PCP	*Pneumocystis carinii pneumonia*
PIC	Pre-integration complex
pol	Polymerase gene
PR	p10 protease
Pr	precursor protein
PR	protease
$Pr160^{Gag-Pol}$	Gag-Pol fusion polyprotein
$Pr5^{Gag}$	Large Gag precursor polyprotein
PTAP	Proline-threonine-alanine-proline
PYRE	Proline-tyrosine-arginine-glutamic acid
Rev	Regulator of virion expression
RNA	Ribonucleic acid
RT	Reverse transcriptase
RTC	Reverse transcription complex
SAB	South African Black
SAC	South African Caucasian
SAI	South African Indian
SAM	South African Mixed ancestry
SAMHD1	Sterile alpha motif domain and hydrolase domain 1
SBT	Sequence based typing
sHLA-G	Soluble HLA-G
SIV	Simian Immunodeficiency Virus
SIVcpz	Simian Immunodeficiency Virus Chimpanzee isolate
SIVgor	Simian Immunodeficiency Virus Gorilla isolate
SU	gp120 surface glycoprotein
t	Telomeric region
T	Threonine
Tat	Trans-activator of transcription
TCR	T cell receptor
TM	gp41 transmembrane protein
TNF-α	Tumour necrosis factor alpha
TRAIL	Tumour necrosis factor-related apoptosis-inducing ligand
TRIM5α	Tripartite motif-5 α
tRNA	Transfer RNA
tRNALys3	Transfer RNA for lysine

TSG101	Tumour susceptibility gene 101
UNAIDS	Joint United Nations programme on HIV and AIDS
USA	United States of America
UTR	Untranslated region
V1-V2	Variable domain 1 to variable domain 2
V3 loop	Variable domain 3
Vif	Viral infectivity factor
VLS	Viral load setpoint
Vpr	Viral protein R
Vpu	Viral protein U
Vpx	Viral protein X

PUBLICATIONS

Loubser S, Paximadis M, Gentle N, Puren A, Gray CM, Tiemessen CT, Frequencies of immune hypersensitivity reaction-associated *HLA class I* alleles in healthy South African Indian and mixed ancestry populations determined by a novel real-time PCR assay, Tissue Antigens. 2014 Oct;84(4):389-97.

Hong HA, **Loubser AS**, de Assis Rosa D, Naranbhai V, Carr W, Paximadis M, Lewis DA, Tiemessen CT, Gray CM., Killer-cell immunoglobulin-like receptor genotyping and HLA killer-cell immunoglobulin-like receptor-ligand identification by real-time polymerase chain reaction, Tissue Antigens. 2011 Sep;78(3):185-94.

PRESENTATIONS

Hong HA, **Loubser AS**, de Assis Rosa D and Gray CM, Measurement of killer-cell immunoglobulin-like receptors (KIR) genotypes in South Africans using real-time PCR, Annual Conference of the South African Immunological Society (SAIS), Vineyard Hotel, Newlands, Cape Town, South Africa, Dec 9-11, 2009.

Hong HA, **Loubser AS**, de Assis Rosa D, Naranbhai V, Carr W, Paximadis M, Lewis DA, Tiemessen CT and Gray CM, Real-time PCR assay for the detection of killer-dell immunoglobulin-like receptor (KIR) ligands (HLA-Bw4, -C1 and -C2), Annual conference of the South African Immunological Society (SAIS), Ingwenya Country Escape, Muldersdrift, Johannesburg, South Africa, Dec 5-8, 2010.

Loubser S, Paximadis M, Puren A and Tiemessen CT, *KIR* genotyping, *KIR3DL1/S1* copy number variation and truncated/full-length *KIR2DS4* determination in South African Indian and Coloured population groups, EMBO Global Exchange Lecture Course on Innate Immunity – evolution and advances in clinical medicine, Stay City, Johannesburg, South Africa, Sep 2-5, 2012.

Loubser S, Paximadis M, Puren A and Tiemessen CT, *KIR* and *KIR*-ligand genotyping, *KIR3DL1/S1* copy number variation and truncated/full-length *KIR2DS4* determination in South African Indian and Coloured population groups., The 8th conference of the Federation of African Immunological Societies (FAIS), Southern Sun Elangeni Hotel, Durban, South Africa, Dec 3-5, 2012.

Loubser S, Paximadis M, Puren A and Tiemessen CT, Frequencies of *HLA class I* alleles associated with abacavir and nevirapine immune hypersensitivity reactions in South African Indian and Coloured populations, 15th Biennial Congress of the Southern African Society for Human Genetics (SASHG), Sandton, Johannesburg, South Africa, Oct 6-9, 2013.

Ngqobe R, **Loubser S**, Paximadis M, Puren A and Tiemessen CT, Frequency of *CCR5Δ32* deletion mutants in South African Indian and Coloured populations., 15th Biennial Congress of the Southern African Society for Human Genetics (SASHG), Sandton, Johannesburg, South Africa, Oct6-9, 2013.

Chapter 1:
Literature Review

1.1. An Overview of HIV-1

1.1.1. Discovery of HIV-1

On June 5, 1981, an article appeared in the Morbidity and Mortality Weekly Report (MMWR), published by the USA Centers for Disease Control and Prevention (CDC), highlighting five unusual cases of *Pneumocystis carinii* pneumonia (PCP) as well as Cytomegalovirus (CMV) co-infection in young homosexual men (29-36 years) living in the Los Angeles area, California (Centers for Disease, 1981). The previously healthy men, who were not acquainted with each other or others with the same symptoms, had been treated for PCP during the period October 1980 to May 1981. Two deaths had occurred subsequently and three of the patients demonstrated poor *in vitro* lymphocyte proliferation responses to mitogens and antigens, suggestive of an underlying immunodeficiency. The report concluded that "All the above observations suggest the possibility of a cellular-immune dysfunction related to a common exposure that predisposes individuals to opportunistic infections such as pneumocystosis and candidiasis." In retrospect, this article was to become the first clinical report to describe what was later to be termed "Acquired Immunodeficiency Syndrome" (AIDS).

An initiative to identify the causative agent of AIDS began in December 1982, when French clinicians contacted the laboratory of Dr. Luc Montagnier at the Pasteur Institute, Paris, and provided a lymph node biopsy, donated by a patient with lymphadenopathy and early symptoms of AIDS, for analysis (Barre-Sinoussi, 2009). The laboratory isolation and characterisation of a previously unknown human retrovirus was successfully achieved by a post-doctoral virologist, Dr. Francoise Barre-Sinoussi, who was able to demonstrate reverse transcriptase activity in T lymphocytes cultured from the sample (Barre-Sinoussi, 2009). These investigations determined that the cultured virus was closely related to the human T-lymphotrophic leukaemia virus (HTLV), the only known human retrovirus at the time. The novel retrovirus was named lymphadenopathy-associated virus (LAV) and details of its discovery were published in 1983, although the authors were cautious in extrapolating a direct link to AIDS (Barre-Sinoussi et al., 1983). LAV was later confirmed by other investigators to be the causative agent of AIDS (Gallo et al., 1984, Levy et al., 1984); although in these reports the virus was termed HTLV-III or AIDS-associated retrovirus (ARV).

In 1985, a second putative human retrovirus, closely related but not identical to LAV, was identified by serological evidence in an infected man in Senegal, West Africa (Barin et al., 1985). The virus was later isolated and characterised and due to the similarity to LAV, given the name LAV-II (Clavel et al., 1986). In 1986, due to confusion in the scientific community regarding virus nomenclature, the International Committee on Taxonomy of Viruses (ICTV) recommended the re-naming of LAV, also referred to as HTLV-III, ARV or immunodeficiency-associated virus (IDAV), to human immunodeficiency virus type 1 (HIV-1) (Case, 1986). Following on from this LAV-II, was subsequently termed HIV-2.

1.1.2. Taxonomy of HIV-1

The 2012 virus taxonomy release (http://ictvonline.org), compiled by the ICTV, classifies HIV-1 and HIV-2 as separate viral species, which together with seven other viral species (i.e. Bovine immunodeficiency virus, Caprine arthritis encephalitis virus, Equine infectious anaemia virus, Feline immunodeficiency virus, Puma lentivirus, Simian immunodeficiency virus and Visna/Maedi virus) are grouped under the genus Lentivirus. The Lentivirus genus, along with five other genera (i.e. Alpharetrovirus, Betaretrovirus, Deltaretrovirus, Epsilonretrovirus and Gammaretrovirus) is grouped under the virus subfamily Orthoretrovirinae. The two subfamilies, Orthoretrovirinae and Spumaretrovirinae, are grouped under the virus family Retroviridae, which is one of 72 other families of viruses.

1.1.3. Biology of HIV-1

1.1.3.1. Structure of HIV-1

Using cryogenic electron microscopy (cryo-EM) techniques, characterisation of the outer surface of HIV-1 determined that mature viral particles (virions) are lipid enveloped with a spherical morphology with a diameter of 110±8 nm (Zhu et al., 2006). Viral spike proteins protrude from the viral envelope and are composed of heterodimers of trans-membrane gp41 molecule and an ectodomain non-covalently associated with a gp120 molecule. The gp120 proteins are extensively glycosylated and contain 27 or more N-linked glycosylation sites. The spike proteins associate into trimeric structures (trimers) that are randomly distributed on the viral surface (14±7 per virion), sometimes in clusters, with no evidence of even spacing (**Figure 1.1**). High resolution cryo-EM imaging of trimeric

structures (**Figure 1.2**) measured the height and width at 12 nm and 15.5 nm, respectively (Mao et al., 2012). Since, the viral lipid envelope is derived from the infected host cell plasma membrane, this leads to the presence of human cellular proteins detectable on the viral surface including ICAM-1 (CD54), LFA-1, HLA II-DR, galectin-1, tetraspanin-14, CD40L, CD62L, CD9, CD46, CD53, CD55, CD59, CD63, CD80, CD81, CD82 and CD86 (Ott, 2008).

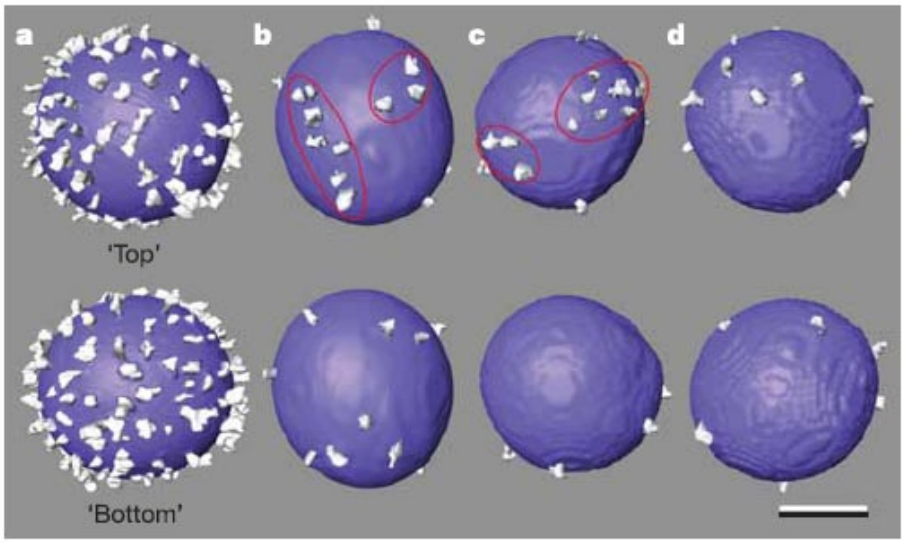

Figure 1.1. Surface-rendered model of SIV and HIV-1 virions (Panel a and b, respectively), determined by cryo-EM techniques, showing individual trimers (white) and clusters of 3 or more trimers (circled in red).Scale bar represents 50 nm. Taken from Zhu et al., 2006 **.**

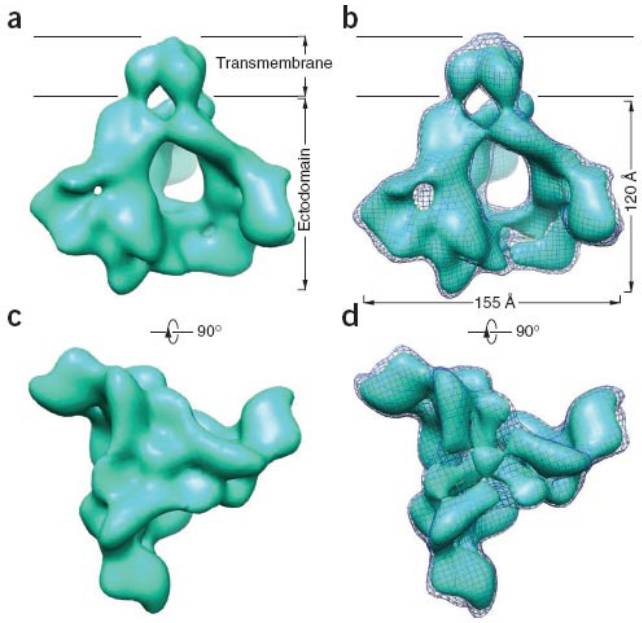

Figure 1.2. Surface-rendered models of membrane-bound HIV-1 envelope trimer, determined by cryo-EM techniques at 11-Å resolution. Side views (panels a and b) and top views (panels c and d) are depicted. Taken from Mao et al., 2012 .

The internal constituents of the mature virion (**Figure 1.3**) are composed of a cone-shaped viral core (capsid), assembled from multiple p24 capsid (CA) proteins and p17 matrix (MA) proteins tethered to the inner surface of the viral envelope (Ganser-Pornillos et al., 2012). The physical structure of the core was recently imaged at high resolution by cryo-EM techniques (**Figure 1.4**) and revealed that the characteristic bullet-shape is based on a "fullerene cone model" consisting of 216 CA hexamers and 12 CA pentamers (5 pentamers at the narrow end and 7 at the wider end) totaling 1,356 CA monomers (Zhao et al., 2013). The core encloses a non-covalently associated dimeric, positive-sense, 5'-capped, 3'-poly-adenylated RNA genome of ~9,000-10,000 nucleotides that is bound by nucleocapsid (NC) proteins and is also annealed to a host cell derived $tRNA^{Lys3}$ molecule (Watts et al., 2009). Viral enzymes reverse transcriptase (RT), protease (PR) and integrase (IN) are also incorporated into the capsid as well as the viral accessory protein, Nef, and a range of cellular proteins that vary according to the virus-producing cell-type (Santos et al., 2012).

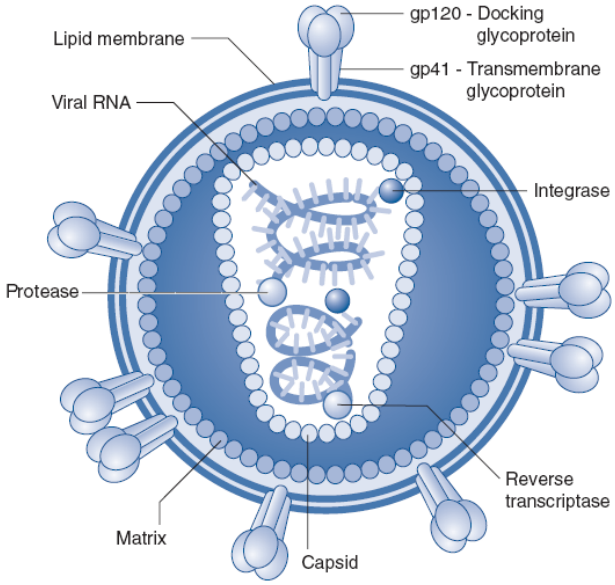

Figure 1.3. Structure of mature HIV-1 virion**.**Taken from Peters et al., 2013 .

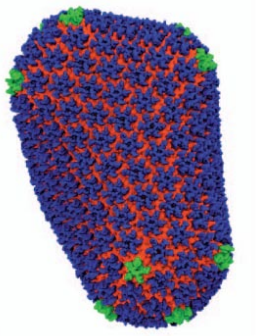

Figure 1.4. HIV-1 capsid structure based on molecular dynamics equilibrated all-atom model comprising 216 p24 hexamers (blue) and 12 p24 pentamers (green). Taken from Zhao et al., 2013 **.**

1.1.3.2. Genomic organisation and viral protein processing

The HIV-1 genome has 9 open reading frames (ORFs) and encodes 15 proteins (Watts et al., 2009). In addition, there are two untranslated regions (UTRs) at the 5' and 3' ends of the genome (**Figure 1.5**). The first ORF encodes the group associated antigen (*gag*) gene which is located after the 5'UTR. Unspliced full-length mRNA transcripts of ~9kb are used to translate the *gag* ORF into a large precursor polyprotein ($Pr5^{Gag}$) which is cleaved by viral protease into 4 mature proteins: p17 matrix protein (MA), p24 capsid protein (CA), p7 nucleoprotein (NC) and p6 (Bell and Lever, 2013). The second ORF following *gag* is the polymerase (*pol*) gene that encodes viral enzymes. The *pol* gene product is translated as a Gag-Pol fusion polyprotein ($Pr160^{Gag\text{-}Pol}$) due to the presence of a -1 ribosomal frameshift signal located upstream of the *pol* start codon and ~5% of transcripts are translated as fusion proteins (Bell and Lever, 2013). $Pr160^{Gag\text{-}Pol}$ is processed into 4 mature proteins by viral protease including p10 protease (PR), p66/p51 reverse transcriptase (RT) heterodimer and p31 integrase (IN). The third ORF encodes the envelope (*env*) glycoproteins. Initially 2kb spliced mRNA transcripts are translated to yield the two regulatory proteins, p19 regulator of virion expression (Rev) and p14 trans-activator of transcription (Tat) and the accessory protein, p27 negative regulatory factor (Nef). A 4kb spliced mRNA transcript is translated to yield the viral envelope polyprotein, gp160, and accessory proteins, p16 viral protein U (Vpu), p23 viral infectivity factor (Vif) and p15 viral protein R (Vpr) (Bell and Lever, 2013). The gp160 protein is translated and concomitantly glycosylated and later cleaved by a cellular protease, Furin, in the Golgi body, to generate the gp41 transmembrane (TM) envelope glycoproteins and the ectodomain gp120 surface (SU) glycoproteins (Checkley et al., 2011).

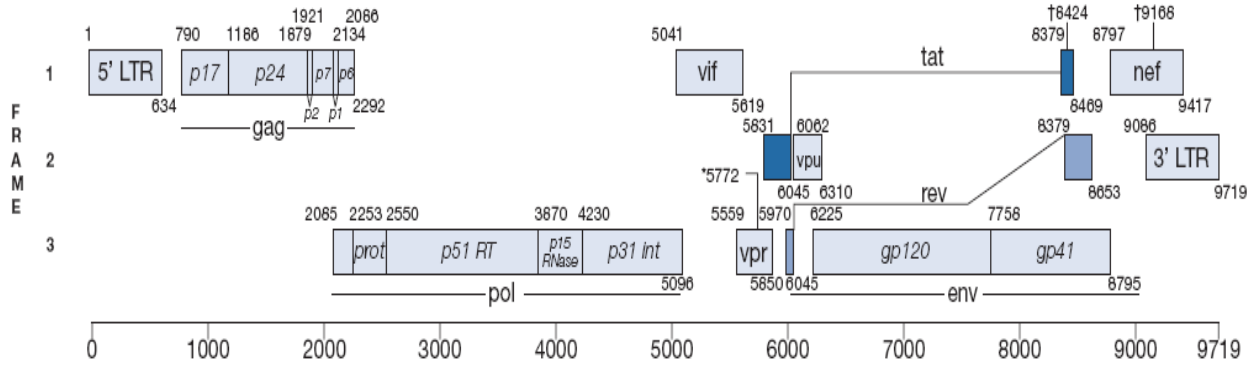

Figure 1.5. Genomic organisation of HIV-1.Taken from Peters et al., 2013 .

1.1.3.3. HIV-1 replication cycle

Mature, infectious HIV-1 virions preferentially infect $CD4^+$ T cells and cells of the macrophage/monocyte lineage and this tropism is largely determined by the envelope glycoproteins (Checkley et al., 2011). Viral entry is initiated by recognition of cell surface CD4 molecules by gp120, which induces a conformational change that exposes the binding site for a co-receptor molecule, either C-C chemokine receptor type 5 (CCR5) or C-X-C chemokine receptor type 4 (CXCR4). The entry inhibitor, Maraviroc, a CCR5 antagonist, binds and distorts the external structural domains of the co-receptor thereby preventing recognition of the amino-terminus and second extracellular loop by gp120 (Menendez-Arias, 2013). Reduced infectivity of target cells is mediated by a naturally occurring mutant of the *CCR5* gene that encodes a truncated protein that is not expressed at the cell surface due to a 32-bp deletion (*CCR5Δ32*) (Liu et al., 1996). Binding of gp120 to CCR5 promotes insertion of the fusion peptide, located at the gp41 amino-terminus, into the host cell membrane, followed by interaction between the heptad repeat domains 1 and 2 of trimeric gp41 to form a six-helical bundle that facilitates fusion of the viral envelope with the plasma membrane (Engelman and Cherepanov, 2012). Membrane fusion introduces the viral core into the cytoplasm, however, contrary to previous conceptions, the core does not uncoat immediately post-fusion, but undergoes intracellular trafficking towards the nuclear membrane using cytoplasmic actin and microtubule networks (Arhel et al., 2007). The mechanism of action of the intrinsic retroviral restriction factor, TRIM5α, is direct binding to the capsid which promotes premature uncoating and exposure of viral RNA to intracellular innate antiviral sensing mechanisms (Bieniasz, 2012).

Reverse transcription takes place inside the intact core, referred to as the reverse transcription complex (RTC), since the core is permeable to dNTPs required for DNA polymerisation, and following completion of reverse transcription, uncoating is triggered by the partially double stranded DNA flap structure (Arhel et al., 2007). A second antiretroviral intrinsic restriction factor, APOBEC3G, a cytidine deaminase enzyme, that is incorporated into virions and converts cytidine to uracil in viral cDNA leading to G-A hypermutation in the viral genome, although this can be counteracted by Vif that prevents encapsidation of APOBEC3G (Bieniasz, 2012). A recently identified third antiretroviral intrinsic restriction factor, SAMHD1, a dGTP-regulated deoxynucleotide triphosphohydrolase enzyme that reduces cytoplasmic dNTP levels in myeloid cells, has been shown to inhibit reverse

transcription and is counteracted by HIV-2 Vpx proteins (Bieniasz, 2012). The NRTI and NNRTI antiretroviral drug classes also inhibit reverse transcription by respective chain-terminating or enzyme active-site blocking mechanisms (Menendez-Arias, 2013). Uncoating of the core leads to the formation of the pre-integration complex (PIC), composed of linear blunt-ended proviral double-stranded DNA, IN, RT, MA protein and cellular proteins, such as LEDGF, that facilitates translocation of proviral DNA into the nucleus via a nucleopore (Engelman and Cherepanov, 2012, Coiras et al., 2009).

At this stage two forms of viral latency can occur, the formation of unintegrated circularised viral genomes containing one or two LTRs, or the integration of the proviral genome into host chromosomal DNA that remains quiescent, i.e. no viral gene transcription takes place (Coiras et al., 2009). Immune activation of latently infected cells promotes import of transcription factors, such as NF-κB and NFAT, into the nucleus, that binds to the proviral 5'UTR and recruits RNA polymerase II that is required to initiate viral mRNA transcription (Coiras et al., 2009). The first full-length viral mRNA transcripts are spliced to 2 kb transcripts encoding the regulatory proteins Rev and Tat as well as Nef accessory proteins (Bell and Lever, 2013). These transcripts are translated in the cytoplasm and Rev and Tat is imported back into the nucleus where Tat increases the rate of viral mRNA transcription, while Rev facilitates export of unspliced 4 kb and 9 kb viral mRNA transcripts. The 4 kb transcripts are used to translate envelope gp160 polyproteins, Vpu, Vpr, Vif and Tat, while Gag and Gag-Pol polyproteins, $pr55^{Gag}$ and $pr160^{Gag\text{-}Pol}$, are translated from the 9 kb transcripts. Envelope gp160 is cleaved by cellular furin, to generate gp41 and gp120 which associates into trimers and is exported to the plasma membrane using the cellular secretory pathway (Ganser-Pornillos et al., 2012, Checkley et al., 2011). Assembly of a new virus particle begins at the plasma membrane with $pr160^{Gag\text{-}Pol}$ binding to dimeric viral genomic RNA complexes present in the cytoplasm via the NC domain (Ganser-Pornillos et al., 2012). This complex is tethered to the inner surface of the plasma membrane via the myristoylated amino-terminus of the MA domain (Checkley et al., 2011). The MA domain also interacts with the cytoplasmic tail of transmembrane gp41 to facilitate recruitment of trimers (Ganser-Pornillos et al., 2012). Additional recruitment of ~2,400-2,500 $pr55^{Gag}$ molecules leads to formation of a lattice and outward curvature of the plasma membrane until a narrow connecting stalk remains (Lee et al., 2012, Ganser-Pornillos et al., 2012). Approximately 120 molecules of $pr160^{Gag\text{-}Pol}$ (Lee et al., 2012) as well as accessory proteins, Nef, and in some cases Vif and Vpr are also incorporated into the virion (Sundquist and Krausslich, 2012).

Maturation of the newly budded virus is dependent on viral protease activity. The interaction of two $pr160^{Gag\text{-}Pol}$ molecules leads to an autocatalytic process that generates a functional protease enzyme that in turn cleaves $pr55^{Gag}$ into p17 MA, p24 CA, p7 NC and p6 proteins and $pr160^{Gag\text{-}Pol}$ into p51/p66 RT, p10 PR and p31 IN (Lee et al., 2012). The abscission of the connecting stalk is facilitated by $p6^{Gag}$ in concert with cellular proteins TSG101 and other proteins of the endosomal sorting complexes required for transport (ESCRT) (Ganser-Pornillos et al., 2012). The intrinsic antiretroviral restriction factor, tetherin, captures budding viruses and prevents their release, but this function is inhibited by Vpu (Bieniasz, 2012). The immature virion is non-infectious at this stage and protease activity continues until sufficient mature viral proteins have been generated to re-organise the internal structure of the virion characterised by a cone-shaped core (Zhao et al., 2013), a condensed genomic RNA structure in association with NC and viral enzymes and MA proteins that remain associated with the inner envelope. The protease inhibitor class of antiretroviral drugs interferes with essential downstream protease functioning and thereby inhibits maturation of the virion (Menendez-Arias, 2013).

1.2. Epidemiology of HIV-1

1.2.1. Global

Epidemiological statistics for 2013 (**Figure 1.6**) released by the Joint United Nations HIV and AIDS programme (www.unaids.org) estimated that 35.0 million [33.2-37.2] people were living with HIV-1. Of these, 31.8 million [30.1-33.7] were adults, 16.0 million [15.2-16.9] were women and 3.2 million [2.9-3.5] were children under the age of 15 years. New infections were estimated at 2.1 million [1.9-2.4] of which 1.9 million [1.7-2.1] were adults and 240,000 [210,000-280,000] were children under the age of 15 years. AIDS deaths were 1.5 million [1.4-1.7] of which 1.3 million [1.2-1.5] were adults and 190,000 [170,000-220,000] were children under the age of 15 years.

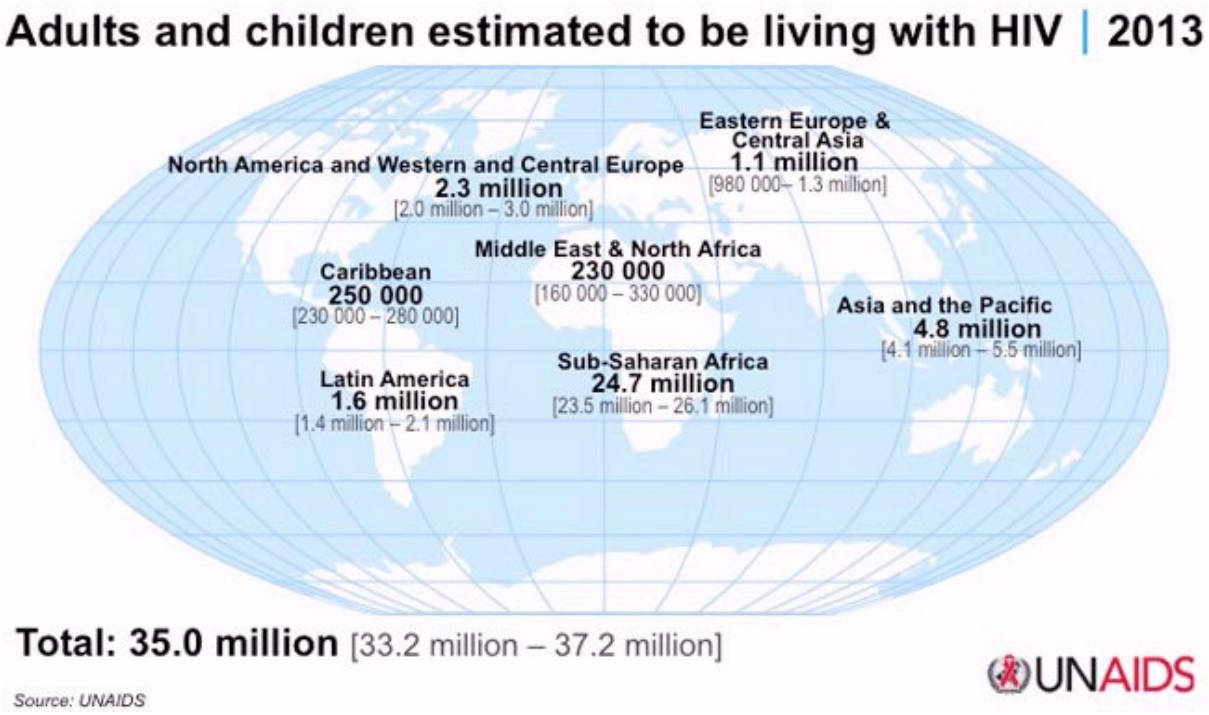

Figure 1.6. Global estimates of people living with HIV-1 in 2013.Taken from www.unaids.org.

1.2.2. Sub-Saharan Africa

Sub-Saharan Africa remains the region of highest HIV-1 burden with an estimated 24.7 million [23.5-26.1] persons living with HIV-1 of which 2.9 million [2.6-3.2] were children. A total of 1.5 million [1.3-1.6] new infections were estimated to have occurred in this region and 1.1 million [1.0-1.3] adult and child deaths were due to AIDS. The HIV-1 prevalence in the 15-49 year adult age group was estimated at 4.7% [4.4-4.9].

1.2.3. South Africa

Updated survey figures released by Statistics South Africa (www.statssa.gov.za) for mid-year 2013 estimated the national HIV-1 prevalence at ~10% and the total population was estimated at 52.982 million which equates to 5.298 million people living with HIV-1. This is an increase in the number of people living with HIV-1 in 2002, when 4 million people were estimated to be HIV-1 positive. In adults aged 15-49 the HIV-1 prevalence is estimated at 15.9%. The number of individuals in this age group was 29,218,479 which equates to 4,645,738 HIV-1-positive young adults. The local South African (SA) population is ethnically diverse, consisting of SA Black (SAB), SA mixed ancestry (SAM), SA Caucasian (SAC) and SA Indian (SAI) individuals (www.statssa.gov.za).The most recent national census (2013) provides estimates of the number of individuals per population as follows:

SAB (n=42,284,100), SAM (n=4,766,200), SAC (n=4,602,400) and SAI (n=1,329,300). Updated HIV prevalence rates stratified by population are not available, however, a national survey conducted in 2012 estimated rates of 15.0%, 0.3%, 3.1% and 0.8% in SAB, SAC, SAM and SAI populations, respectively (www.hsrc.ac.za/en/research-outputs/view/6871). Based on these HIV-1 prevalence rates, this equates to the number of individuals living with HIV-1 for SAB (n=6,342,615), SAM (n=147,752), SAC (n=13,807) and SAI (n=10,364) populations.

1.3. Origin and genetic diversity of HIV-1

Human adaptation of Simian immunodeficiency viruses (SIV) following zoonotic transmission from non-human primate species is thought to have lead to the introduction of HIV-1 and HIV-2 into humans, most likely in the Western and Central African regions via contact with infected animals through bushmeat hunting practices or keeping animals as pets (Hemelaar, 2012). More than 40 non-human primate species that inhabit the African continent (**Figure 1.7**) are known to carry species-specific SIV strains (Chahroudi et al., 2012). HIV-1 genotypes are sub-divided into 4 groups (group M, N, O and P) with group M being the largest group and responsible for the pandemic with millions of infections globally (Hemelaar, 2012). Group O HIV-1 infections are in the order of tens of thousands (~1% of all HIV-1 infections) and are mostly restricted to Western and Central Africa, while groups N and P contain a very small number of viral sequences (n=15 and n=2, respectively) identified in individuals from Cameroon (Mourez et al., 2013).

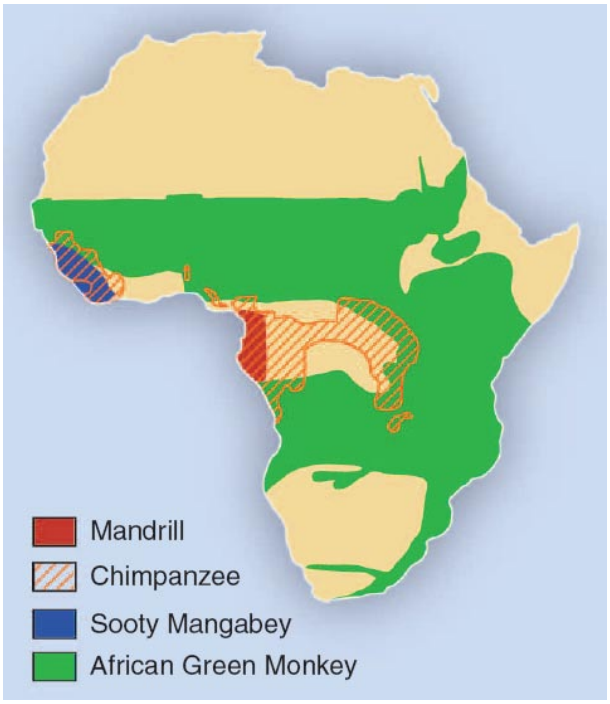

Figure 1.7. Natural habitats of some primate species.Taken from Chahroudi et al., 2012 .

The group M and N viruses are closely related to SIV strains circulating in chimpanzees (SIVcpz), specifically the species *Pan troglodytes troglodytes*, populating Western and Central Africa, and it is thought that two independent cross-species transmissions gave rise to these groups of viruses (Hemelaar, 2012). In contrast, group O and P viruses are more closely related to SIV strains circulating in gorillas (SIVgor), specifically the species *Gorilla gorilla gorilla*, populating the Western lowland of Cameroon, however, it is not clear from phylogenetic analysis whether these viruses originated from a direct zoonotic transmission from gorillas or if an intermediate host species was involved (Hemelaar, 2012). HIV-1 group M viruses have diversified considerably in humans and genetic variants can be classified into 9 subtypes (A-D, F-H, J and K) based on sequence variability. Several circulating recombinant forms (CRFs) of HIV-1 have arisen through recombination between viral strains belonging to different clades with a total of 61 HIV-1 CRFs currently documented by June 2014 (http://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html). Similarly, HIV-1 group O viruses are genetically diverse with 5 putative subtypes identified (subtypes 1-5) (Yamaguchi et al., 2002). HIV-2 is closely related to SIV strains circulating in the sooty mangabey monkey, species *Cercocebus atys*, populating the Western African region, and is also subdivided into groups A-H. Notably, these group variants have resulted from several

independent cross-species transmission events while HIV-2 CRFs were subsequently generated by genetic recombination in humans (Peeters et al., 2013).

The argument for the origin of HIV-1 in central Africa is supported by evidence such as the location of primate species (chimpanzee and gorilla) naturally infected with SIV strains closely-related to HIV-1, the presence of all 4 HIV-1 groups (M, N, O and P), the presence of all HIV-1 subtypes (A-D, F-H, J and K), higher degrees of intersample sequence diversity, a larger representation of CRFs, the location of earliest known HIV-1 positive samples dating to 1959 and 1960 (Hemelaar, 2012). HIV-1 prevalence has also declined to stable levels in this region, suggestive of an older epidemic. The earliest known HIV-1 strains have been detected in a plasma sample from an infected male in 1959 and a lymph node biopsy sample from an infected female in 1960, both from the Democratic Republic of Congo. Subsequent genotyping determined that these sequences were 12% divergent and belonging to subtype D and subtype A, respectively, demonstrating that HIV-1 had already diversified at this early time-point during the HIV-1 epidemic (Worobey et al., 2008, Zhu et al., 1998).

The timing of the origin of group M HIV-1 in humans was estimated to have occurred between 1915 and 1941, based on extrapolation from mutational rates determined for HIV-1 sequences available between 1980 and 2000 (Korber et al., 2000). This analysis also correctly estimated the age of the 1959 HIV-1 sequence to within a year or two, in support of the reliability of the method, however, the authors cannot be certain that the mutational rates was constant during the early stages of the epidemic. A recent epidemiological study of HIV-1 sequence data collected from central Africa has concluded that the epicenter of expansion of group M viruses is most likely Kinshasa in the Democratic Republic of Congo around the 1920s due to the culmination of several factors including extensive urban development and a lucrative sex trade industry (Faria et al., 2014). There is also speculation that exposure of humans to SIV/HIV may have occurred at much earlier time than current estimates. This is based on evidence of positive selection of resistance genes detected in members of the Biaka pygmie tribe resident in the Central African Republic (Zhao et al., 2012).

The spread of HIV-1 from central Africa to other regions of the world has lead to differential global subtype distributions (**Figure 1.8**), most likely affected by many factors including founder effects, population growth, urbanisation, cultural and sexual behavioural practices, transport and migration patterns. Notably, subtype C strains of HIV-1 represent

48% of global infections and these viruses predominate in the southern African regions as well as Ethiopia and India, possibly reflecting differences in viral or host genetics (Hemelaar, 2012). It has been postulated that the dominance of subtype C viruses may be related to viral and/or host genetic factors that increases susceptibility or promotes higher levels of viraemia in infected individuals, thus increasing transmission risk. This may be further exacerbated by poor antiretroviral coverage in the areas where subtype C viruses are circulating (Neogi et al., 2013). However, a recent publication did not find any significant differences in viral load setpoint in recent seroconverters from west and southern Africa infected with subtypes A, D and C (Campbell et al., 2013).

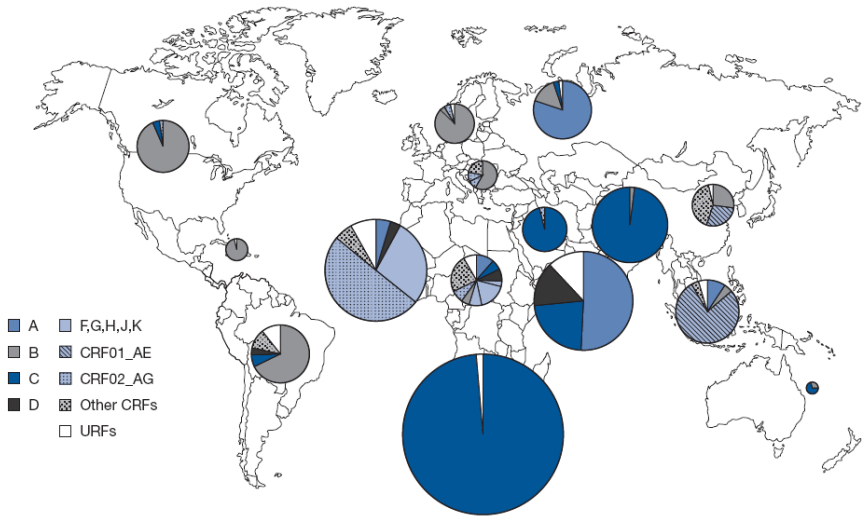

Figure 1.8. Global distribution of HIV-1 subtypes and circulating recombinant forms between 2004 and 2007**.** Taken from Peters et al., 2013 **.**

1.4. HIV-1 transmission, pathogenesis, immunity and treatment

1.4.1. Transmission of HIV-1

Infectious HIV-1 virions are present in semen, genital secretions, blood and breastmilk and can be transmitted by homosexual or heterosexual intercourse (which includes unprotected vaginal, anal and oral sex), parenteral inoculation (including needlestick injury and injecting drug users), exposure to infected blood or blood products (including transfusion and organ transplantation) and from mother-to-child (including *in utero*, intrapartum or post-natal transmission by breastfeeding) (Moss, 2013). The most common route of infection (~80%) is through mucosal surfaces (Cohen et al., 2011) that includes vaginal, penile and

rectal tissues, with the risk of acquiring HIV-1 through sexual contact being higher for receptive anal intercourse compared to heterosexual intercourse (Tebit et al., 2012). Most HIV-1 infections worldwide are through heterosexual contact with a disproportionately higher infection rate in women compared to men, partly due to a greater efficiency of HIV-1 transmission from male-to-female compared to female-to-male (Tebit et al., 2012).

The female genital tract (**Figure 1.9a**) is subdivided into endocervix and the transitional zone, composed of a single layer of simple columnar epithelium, and the vagina and ectocervix composed of multiple layers of stratified squamous epithelium (Carias et al., 2013, Hladik and McElrath, 2008). It was initially thought that it was more likely that HIV-1 will penetrate the mucosal barriers of the endocervix or transitional zone due to the relative thickness of the ectocervical squamous epithelium, however, in the endocervix most virus is efficiently trapped by mucus, while the larger surface area of the ectocervix (~15 times greater than the endocervix) provides a better opportunity for HIV-1 to breach mucosal barriers. New evidence also suggests that due to a weakening of tight-junctions, HIV-1 viral particles can passively diffuse between epithelial cells to an average depth of 2 µm (sometimes >10 µm), where contact with immune cells is more likely (Carias et al., 2013).

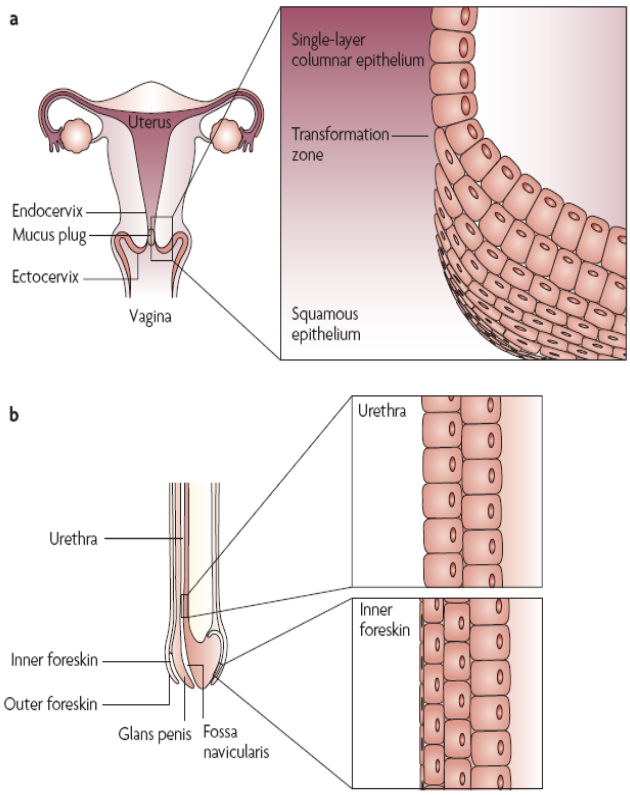

Figure 1.9. Basic cellular structure of female (a) and male (b) genital mucosal surfaces. Taken from Hladik and McElrath, 2008 .

It has been postulated that the first immune cells to contact HIV-1 are skin resident dendritic cells known as Langerhans' cells, which are abundant in the squamous epithelial cell layer and hence are likely to contact HIV-1 that has partially penetrated this epithelial layer (Hladik and McElrath, 2008). Langerhans' cells are also known to sample antigens present on the epithelial surface via cytoplasmic projections and may contact HIV-1 in this way (Hladik and Hope, 2009). Infection of males is thought to occur in a similar manner via the inner foreskin which also has multiple layers of squamous epithelial cells (**Figure 1.9b**) and tissue resident Langerhans' cells, while simple columnar epithelium is only present in the urethra and presents a very small surface area compared to the foreskin (Hladik and McElrath, 2008). The surgical removal of the foreskin by circumcision, which has been shown to reduce female-to-male HIV-1 transmission (Auvert et al., 2005), is most likely protective due to the absence of Langerhans' cells and increased keratinisation of the skin that reduces virus diffusion into the epithelium.

Langerhans' cells express C-type lectins on the cell surface called langerin (CD207) that bind carbohydrate moieties such as those present in gp120 viral envelope proteins. This facilitates internalisation of virus for degradation, however, in higher concentrations, HIV-1 can be trapped on the cell surface and transported to lymphoid tissues where $CD4^+$ T cells become infected (Tebit et al., 2012). Myeloid dendritic cells similarly trap HIV-1 on the cell surface with an alternative C-type lectin called DC-SIGN (CD209) (Geijtenbeek and van Kooyk, 2003). There are reports that Langerhans' cells can be productively infected with HIV-1 themselves (Kawamura et al., 2003), although these cells are CD4 negative but may become infected through langerin-mediated internalisation of HIV-1 (Hladik and Hope, 2009). Inflammation of the genital tract due to sexually transmitted infections increase the probability of HIV-1 acquisition, most likely due to activation of antigen presenting cells such a Langerhans' cells and recruitment of activated $CD4^+$ T cells to a closer proximity to infectious HIV-1 (Arien et al., 2011).

In heterosexual HIV-1 transmission, even though a heterogeneous viral population is usually present in the infected inoculum, such as semen, only a small number or most often a single virion establishes infection in the new host (Keele et al., 2008). There is uncertainty whether a single virus is selected at the mucosal site of entry or if multiple viruses breach the mucosal barrier but undergo positive selection in another compartment, nevertheless, in most circumstances the transmitted/founder virus has a particular phenotype that includes almost

exclusive dependence on CCR5 as the co-receptor for target cell entry, short V1-V2 loops in gp120 as well as fewer N-linked glycosylation sites and a higher neutralisation sensitivity compared to donor viruses (Arien et al., 2011).

1.4.2. Pathogenesis of HIV-1

The hallmark of HIV-1 pathogenesis in untreated individuals is the progressive depletion of $CD4^+$ T cells over time which is coupled to an increasing impairment of adaptive immune responses leading to increased susceptibility to opportunistic infections. In untreated individuals, the typical course of infection is subdivided into three stages: acute infection stage, asymptomatic period and AIDS onset (Qu et al., 2008). Based on laboratory testing for the presence of viral RNA, viral antigens or virus-specific antibodies, the Fiebig staging system can further expand virological and immunological changes during each stage of infection (Fiebig et al., 2003).

1.4.2.1. Acute infection stage

Heterosexual exposure to HIV-1 does not necessarily result in productive infection, since in low risk settings, HIV-1 transmission ranges from 2-50 per 10,000 exposures (Arien et al., 2011). Symptoms of primary infection may develop after 2-3 weeks and last up to 4 weeks (Qu et al., 2008). Exposure to HIV-1 at mucosal surfaces in susceptible hosts, permits virus to cross the epithelial cell barriers, most likely facilitated by Langerhans' cells that transport infectious virus to local lymphoid tissues; a process that takes ~2-6 hours (Cohen et al., 2011). Langerhans' cells in an activated state have been shown to mediate trans-infection of $CD4^+$ T cells (Fahrbach et al., 2007) which may be related to their professional role as antigen-presenting cells that migrate to lymphoid tissues where they stimulate antigen-specific T cells (Ayala-Garcia et al., 2005). There is also the possibility that Langerhans' cells may become productively infected although these dendritic cells are considered CD4 negative (Kawamura et al., 2003).

Trans-infection of $CD4^+CCR5^+$ T cells leads to local production of virus in regional draining lymph nodes of the genital tract and dissemination of infected effector T cells to mucosal tissues where uninfected T cells and macrophages can become infected as well as dissemination to other tissues. Activated antigen-presenting cells are known to secrete

cytokines such as IL-1, TNF-α and IL-6, known to cause fever. This symptom, including others, such as pharyngitis, myalgia, lymphadenopathy, headache, diarrhoea, nausea, meningismus, photophobia, rash, vomiting and oropharyngeal sores, are sometimes compared to "flu-like" or "mononucleosis-like" illness, although not all primary infections are associated with these symptoms (Celum et al., 2001). A skin rash (exanthem) is reported in up to 70% of primary HIV-1 infections. HIV-1 dissemination to the skin and productive infection of skin dendritic and Langerhans' cells has been reported in a human case study (Simonitsch et al., 2000) and investigations of the same phenomenon observed in the SIV-rhesus macaque monkey model revealed the presence of SIV and infiltrates of $CD4^+$ and $CD8^+$ T cells expressing αEβ7 integrins in skin biopsies (Sasseville et al., 1998). The integrin αEβ7 is specific for homing of effector T cells to the skin and this may be due to overlap of the cutaneous and genital mucosa with regard to immune responses involving skin-resident Langerhans' cells that are also present in both male and female genital squamous epithelium. Notably, lymphadenopathy caused by clonal expansion of antiviral $CD4^+$ and $CD8^+$ T cells leads to preferential infection of HIV-specific $CD4^+$ T cells (Douek et al., 2002) and likely reflects trans-infection of responding $CD4^+$ T cells by antigen-presenting cells that have captured viral antigens, such as Langerhans' cells expressing langerin or myeloid dendritic cells expressing DC-SIGN (Fahrbach et al., 2007).

SIV studies have shown rapid dissemination of virus to gut lymphoid tissues as early as 2 weeks after infection, while human studies have determined that this causes extensive $CD4^+$ T cell depletion in the gastrointestinal tract (Brenchley et al., 2004). Although the path of HIV-1 dissemination from the genital tract to the gut is unclear, there is evidence from murine studies that the mesenteric lymph node may facilitate a skin and gut homing overlap involving retinoic acid dependent activation of $langerin^+$ dendritic cells (Chang et al., 2008). This suggests that HIV-1 dissemination to the skin may mediate virus dissemination to the gut via trans-infection of gut homing $CD4^+$ T cells activated in the mesenteric lymph node. Systemic dissemination of virus also facilitates seeding of the central nervous system (CNS) with HIV-1 via infected monocytes or macrophages that differentiate into microglial cells or perivascular brain macrophages (Gras and Kaul, 2010). These cells become reservoirs of HIV-1 in the brain without virus production, since immune activation is suppressed in the CNS, however, failure to control peripheral HIV-1 replication at AIDS onset leads to a breakdown of immune privilege and viral replication in the CNS causes dementia.

During acute infection the appearance of biological markers of HIV-1 infection detectable by clinical laboratory testing procedures can be listed according to a staging system. According to Fiebig staging (Fiebig et al., 2003), during acute infection, plasma viral RNA is undetectable (eclipse phase) with ultrasensitive RNA testing assays (1-5 copies/ml) until ~10 days post-infection (p.i.); stage I (10-15 days p.i.) is when RNA becomes detectable in blood, but p24 antigen is undetectable; stage II (15-20 days p.i.) is when p24 antigen becomes detectable; stage III (20-24 days p.i.) is when HIV-specific IgM becomes detectable and it is during this seroconversion period when peak viraemia is attained and begins to decline; Stage IV (24-30 days p.i.) is the period where viral load continues to decline, p24 levels decline and Western blot results are indeterminate; Stage V (30-100 days p.i.) is when Western blot results are positive, excluding positivity for p31 integrase and viral load reach a low level but continue to decline slowly as do p24 levels; Stage VI (100 days p.i.) is when viral load becomes stable and is said to reach a "setpoint" , p24 becomes undetectable and the individual enters the early chronic infection period. The staging and results of clinical laboratory testing procedures is shown in **Figure 1.10** (Cohen et al., 2011).

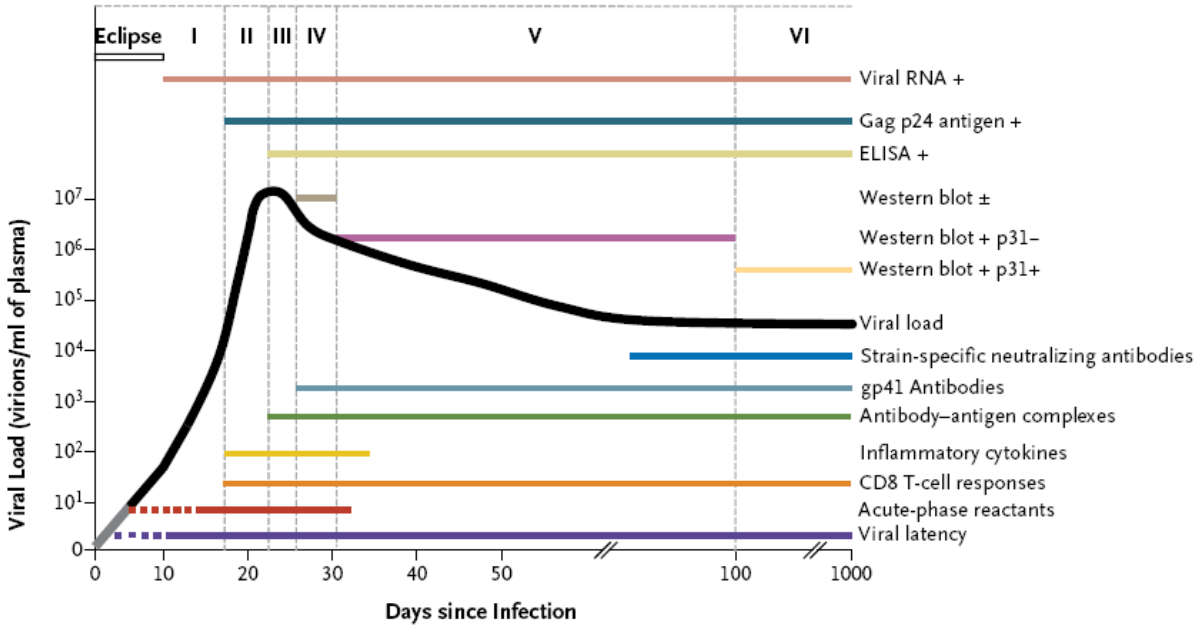

Figure 1.10. Staging, immune responses and clinical laboratory testing sensitivity during acute HIV-1 infection. Taken from Cohen et al., 2011 .

1.4.2.2. Asymptomatic period

The asymptomatic (also called chronic or clinical latency) period of HIV-1 infection is characterised by a viral load setpoint which remains relatively stable, although $CD4^+$ T cell count continues to decline and the viral load setpoint is positively correlated with disease progression (Geskus et al., 2007). Classification of untreated individuals based on the duration of the asymptomatic period defines three main groups: rapid progressors, typical progressors and long-term non-progressors (LTNPs) (Qu et al., 2008). Rapid progressors are characterised by an accelerated decline in $CD4^+$ T cell counts and progression to AIDS within 2-5 years, although this only occurs in 5-15% of individuals in the asymptomatic period, while typical progressors do not develop symptoms of AIDS for 8-10 years. Another 5-15% of individuals in the asymptomatic period remain clinically/immunologically stable beyond 10 years, albeit with detectable viraemia but usually at low levels, and these individuals are termed LTNPs. LTNPs, despite a slow decline in $CD4^+$ T cell count and low viraemia, do eventually develop AIDS (Qu et al., 2008). An additional subset of LTNPs have undetectable levels of viraemia (<50 copies/ml) and these individuals are termed "elite controllers", "elite suppressors" or "HIV controllers", however only 0.1%-1% of infected individuals represent this minority group (Deeks and Walker, 2007).

1.4.2.3. AIDS onset

Towards the end of the asymptomatic period viral load begins to increase and this precedes a more rapid decline in $CD4^+$ T cells count to <200 cells/µl that marks the onset of AIDS and the manifestation of opportunistic diseases (Qu et al., 2008). Due to immune dysregulation, opportunistic infections can also manifest before $CD4^+$ T counts decline to <200 cells/µl, however the severity of infections is increased at lower $CD4^+$ T cell counts. In untreated HIV-1 infection, opportunistic infections such as pulmonary tuberculosis, Pneumococcal pneumonia, Herpes zoster, Kaposi's sarcoma and oropharyngeal Candidiasis begin to appear at $CD4^+$ T cell counts of <500 cells/µl (Peters et al., 2013). At a $CD4^+$ T cell count of <200 cells/µl, extrapulmonary Tuberculosis, Pneumocystis pneumonia, clinical wasting and non-Hodgkin's lymphoma become more common. Cryptococcus, Candididal esophagitis and chronic Cryptosporidiosis occurs at $CD4^+$ T cell counts <100 cells/µl. Disseminated Cytomegalovirus and disseminated *Mycobacterium avium* complex are common opportunistic infections at $CD4^+$ T cell counts <50 cells/µl.

1.4.3. HIV-1 immunity

The human immune system has two defense arms known as innate immunity and adaptive immunity. The innate immune system is considered non-specific in the sense that defense against pathogens is generalised and targeting of common elements across invading organisms is often employed (pattern recognition). The innate immune system is also already present at birth and is only up or down regulated in response to infection. Conversely, the adaptive immune system has to be induced when infection occurs and the response is antigen-specific with respect to a particular pathogen. Adaptive immunity also has a memory component and must be more rigorously regulated to avoid autoimmune disease.

The HLA class I trans-membrane protein system plays a central role in both innate and adaptive immunity and is also implicated in immune hypersensitivity reactions to pharmaceutical drugs. HLA class I molecules are expressed on most nucleated cells, but can be absent or expressed at very low levels on erythrocytes, neurons, sperm/egg cells and villous trophoblast (Niederkorn and Wang, 2005, Sabir et al., 2013). In humans, there are three *HLA class I* gene loci (*HLA-A, -B* and *-C*) that map to the major histocompatibility complex region on chromosome 6p21.31 (**Figure 1.11**) (Klein and Sato, 2000). Due to extensive genetic polymorphism, *HLA class I* genes segregate into numerous allelic variants. To date 2,995 *HLA-A* alleles, 3,760 *HLA-B* alleles and 2,553 *HLA-C* alleles have been documented (http://www.ebi.ac.uk/ipd/imgt/hla/stats.html).The primary function of HLA class I molecules is to bind small immunogenic peptide antigens sampled from proteins present in the cytosol (**Figure 1.12**).The main difference between allelic variants of HLA class I molecules lies in the peptide-binding specificity of the protein which is largely determined by the amino-acid composition of the bound peptide, particularly residues 2 and 11 that anchor the peptide (**Figure 1.13**).

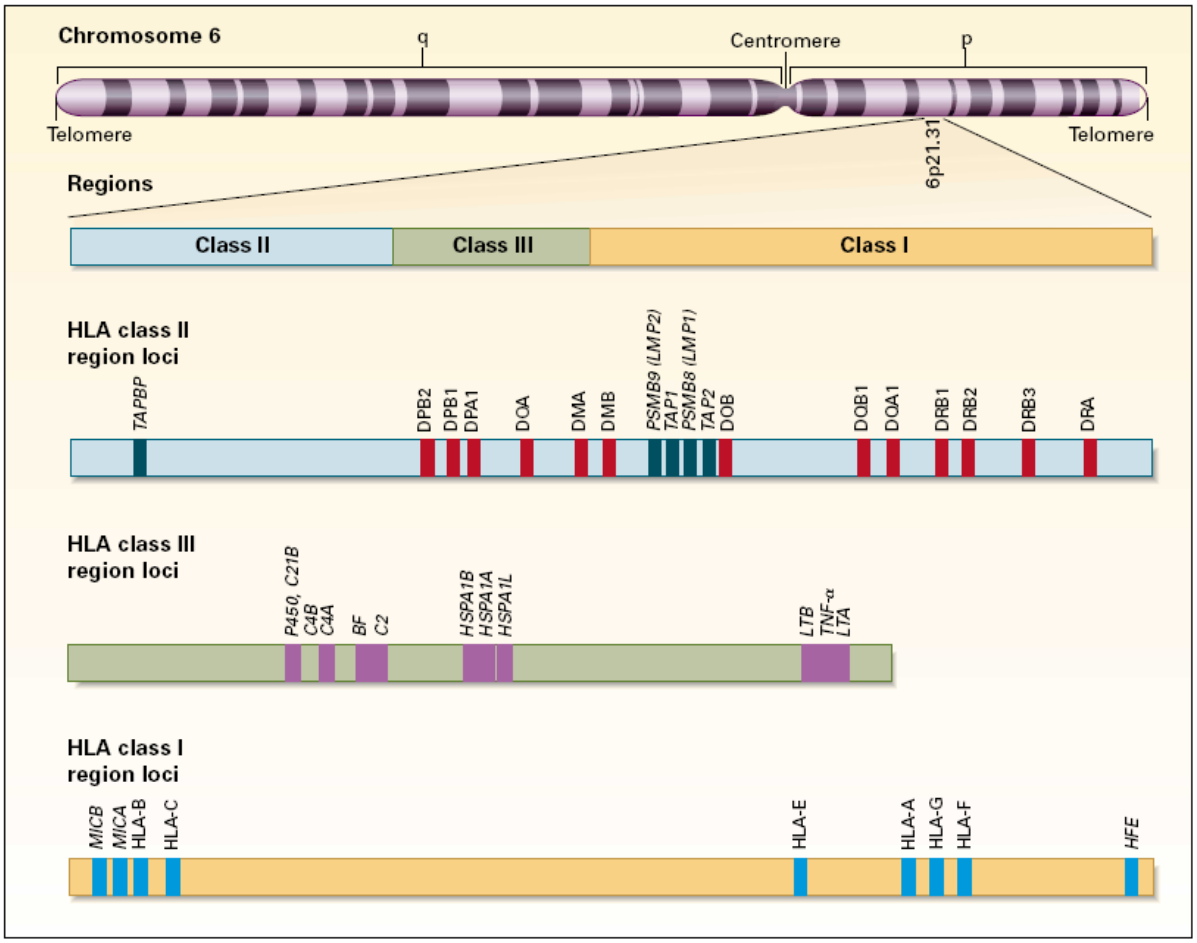

Figure 1.11. *HLA class I* loci mapping to chromosome 6p21.31**.** Taken from Klein and Sato, 2000 .

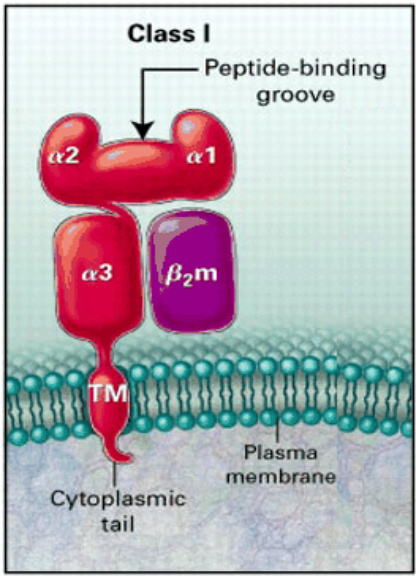

Figure 1.12. HLA class I and II molecule structures**.**Taken from Klein and Sato, 2000 .

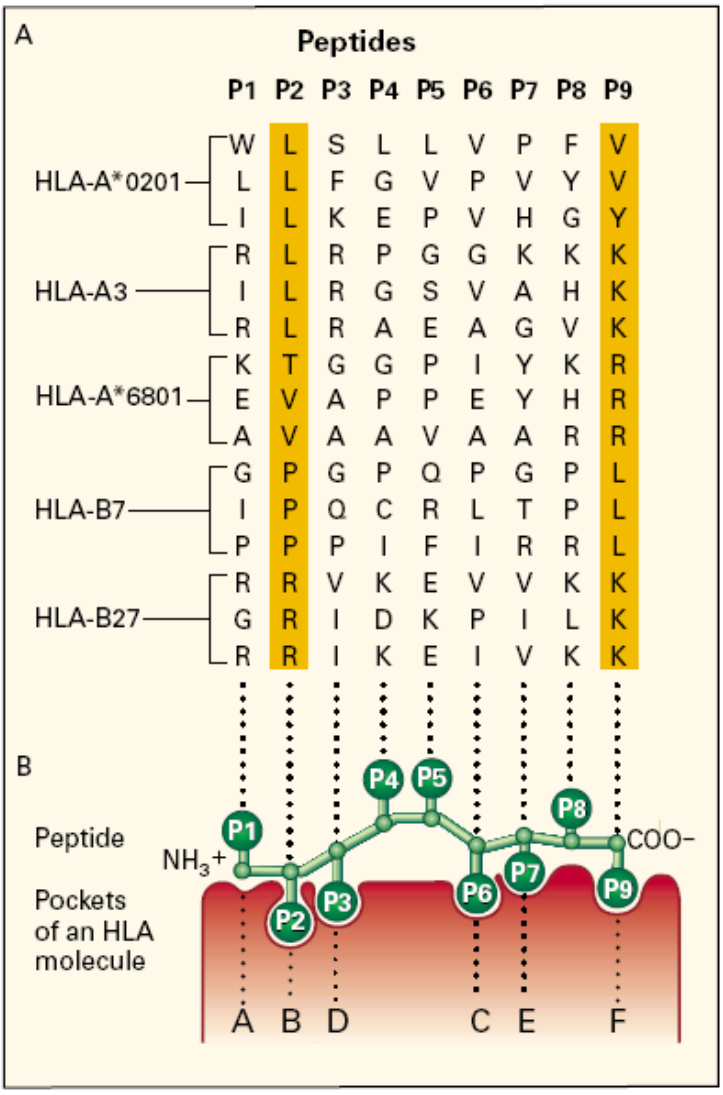

Figure 1.13. Peptide specificities for some HLA class I molecules (a) and orientation of amino acids of bound peptide (b). Taken from Klein and Sato, 2000 .

Many viruses, including HIV-1 and certain forms of malignancy interfere with the surface expression of HLA class I molecules. The innate immune system has evolved a mechanism to detect aberrations in cell surface expression of HLA class I molecules such as low abundance or absence. Lymphocyte subsets known as natural killer (NK) cells perform this duty and have the ability to kill target cells by perforin-dependent cell lysis (**discussed in greater detail in section 2.3.1.**). The adaptive immune system also makes use of HLA class I expression to detect the presence of intracellular pathogens, most commonly viruses. This is mediated by specialised immune cells known as $CD8^+$ cytotoxic T lymphocytes (CTLs) that express T cell receptors (TCR) that are capable of sensing foreign peptide antigens bound to HLA class I molecules (**discussed in greater detail in section 2.3.3.**).

1.4.3.1. Natural Killer cell biology

Natural killer (NK) cells constitute ~10% of circulating lymphocytes and mediate innate anti-viral/anti-tumour surveillance and facilitate adaptive immune responses via interaction with dendritic cells and pro-inflammatory cytokine production, such as interferon-γ and tumour necrosis factor-α (Campbell and Hasegawa, 2013). NK cells functionally kill targets by inducing apoptosis; either by release of lytic granules containing perforin and granzymes (degranulation) or by ligation of NK cell death ligands, such as Fas ligand (FasL) or tumour necrosis factor-related apoptosis-inducing ligand (TRAIL), with cognate death receptors (Deguine and Bousso, 2013, Long et al., 2013).

NK cell cytotoxicity and cytokine production is regulated by soluble factors as well as stimulation of cell surface activating/inhibitory receptors and co-stimulatory molecules (Long et al., 2013). An endosomal activation pathway via endocytosis of membrane receptor/ligand complexes that promotes cytokine production in the absence of cytotoxicity has also been described for uterine NK cells (Rajagopalan, 2010). The killer-cell immunoglobulin-like receptors (KIRs) are a diverse family of inhibitory/activating receptors which in combination with cognate ligands can associate with favourable/ deleterious outcomes in a range of infectious diseases including HIV, HCV, CMV, influenza, *Mycobacterium tuberculosis* and *Plasmodium falciparum* (Jost and Altfeld, 2013, Kulkarni et al., 2008). KIR/ligand interaction also influences susceptibility to many autoimmune/inflammatory conditions and certain types of cancer (Kulkarni et al., 2008). In addition, KIRs and their ligands are thought to play a role in human reproduction and can associate with susceptibility to pre-eclampsia, recurrent miscarriage and spontaneous abortion (Kulkarni et al., 2008, Parham and Moffett, 2013). KIR/ligand mismatched recipient NK cells contribute to solid organ transplant rejection (van Bergen et al., 2011), whereas in allogeneic stem cell transplantation, mismatched KIR/ligand donor NK cells expressing activating receptors reduce the risk of graft-versus-host disease and augment anti-tumour responses (Locatelli et al., 2013, Stern et al., 2010). In HIV discordant couples, mismatched KIR/ligand combinations between sexual partners associates with reduced risk of heterosexual transmission (Jennes et al., 2013) and conversely, discordance of *KIRs* between mother and infant associates with increased risk of vertical transmission of HIV (Hong. et al., 2013).

In humans, 14 *KIR* genes (*KIR2DL1-5, KIR2DS1-5, KIR3DL1-3* and *KIR3DS1*) and two pseudogenes (*KIR2DP1* and *KIR3DP1*) have been described and are located in the leukocyte receptor complex (LRC) on chromosome 19q13.4 (**Figure 1.14**) (Middleton and Gonzelez, 2010). KIRs are classified according to the presence of two (2D) or three (3D) extracellular immunoglobulin-like domains and either a long (L) or short (S) cytoplasmic tail (Uhrberg, 2005). With the exception of KIR2DL4, a long cytoplasmic tail transduces inhibitory signals (KIR2DL1-3, KIR2DL5 and KIR3DL1-3) mediated by two immune tyrosine-based inhibitory motifs (ITIMs), whereas a short cytoplasmic tail transduces activation signals (KIR2DS1-5 and KIR3DS1) mediated by a transmembrane positively charged lysine or arginine residue that recruits immune tyrosine-based activation motif (ITAM)-bearing DAP12 homodimers (Lanier et al., 1998). KIR2DL4 is an unusual receptor that shares structural features with both inhibitory and activating KIRs (i.e. a single ITIM in the cytoplasmic tail and a positively charged arginine in the transmembrane region) and functions as an activating receptor with inhibitory potential (Faure and Long, 2002).

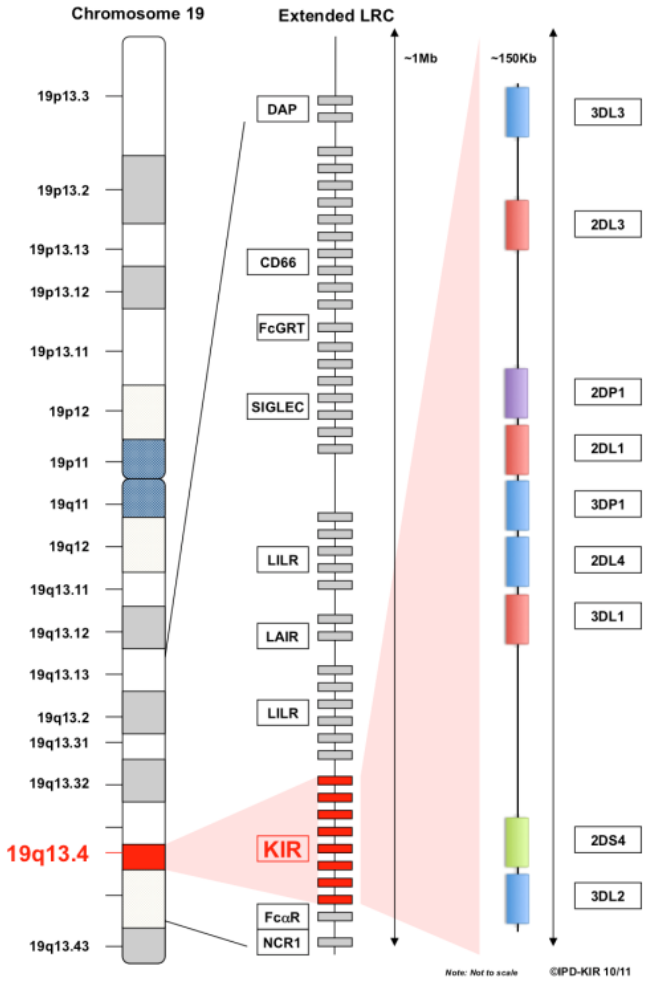

Figure 1.14. Map of the leukocyte receptor complex (LRC) depicting the *KIR* loci on chromosome 19q13.4**.** Taken from http://www.ebi.ac.uk/ipd/kir.

Classical HLA class I molecules serve as ligands for KIR2DS1/4, KIR2DL1-3 and KIR3DL1/2, while ligands for KIR2DS2/3/5, KIR2DL5, KIR3DS1 and KIR3DL3 are unknown (Parham et al., 2012). KIR3DL1 binds HLA-B molecules that express the Bw4 epitope defined by an isoleucine (HLA-Bw4^{80I}) or threonine (HLA-Bw4^{80T}) at amino acid position 80 (Cella et al., 1994, Gumperz et al., 1995). A subset of HLA-A molecules also express the HLA-Bw4^{80I} epitope (Stern et al., 2008). HLA-C molecules with an asparagine present at position 80 are defined as group C1 epitopes and are recognised by KIR2DL2 and KIR2DL3, whereas KIR2DL1 and KIR2DS1 recognise group C2 epitopes defined by the presence of a lysine at position 80 (Mandelboim et al., 1996). Interestingly, in comparison to other primate species, humans lack an activating KIR that recognises the C1 epitope (Moesta et al., 2010) and additionally HLA-B allotypes that express the group C1 epitope have been lost (Abi-Rached et al., 2010). Notably, a minority of *HLA-B* alleles do encode the C1 epitope, such as *HLA-B*07:13, -B*46:01:01, -B*46:01:02, -B*46:02-46:18* and *-B*67:02* as a result of recombination and *HLA-B*07:15, -B*08:15, -B*15:57, -B*18:06, -B*35:74, -B*39:27, -B*40:73, -B*55:03* and *-B*73:01* as a result of point mutations (Abi-Rached et al., 2010). KIR2DS4 recognises subsets of HLA-C1 and HLA-C2 allotypes as well as HLA-A*11 (Graef et al., 2009). KIR3DL2 has been shown to bind HLA-A*11 and HLA-A*03 allotypes as well as HLA-B*27 dimers and free heavy chains (Shaw and Kollnberger, 2012, Wong-Baeza et al., 2013). Recent reports have broadened the functional properties of KIR2DS4 and KIR3DL2 by demonstration of interaction with non-classical HLA-F molecules and peptide free HLA class I open conformers (Goodridge et al., 2013).

KIR gene expression is controlled by epistatic regulatory mechanisms, hence, KIRs are expressed stochastically on the surface of clonal subpopulations of NK cells (Chan et al., 2005). Conversely, KIR2DL4, is expressed by all NK cells and furthermore, constitutive internalisation of cell surface KIR2DL4 by endocytosis results in localisation of the majority of receptors in endosomes (Rajagopalan, 2010). The ligand for KIR2DL4 is soluble HLA-G (sHLA-G), a non-classical HLA class Ib molecule, which is abundantly produced at the maternal-infant interface (Carosella et al., 2008). Although KIR2DL4 has been shown to function as an activating receptor with inhibitory potential (Faure and Long, 2002), recent studies have demonstrated that endosomal KIR2DL4 transduces unique intracellular activation signals that mediates NK cell production of pro-inflammatory cytokines and pro-angiogenic factors, but does not induce cell-mediated cytotoxicity (Rajagopalan, 2010). This was suggestive that NK cells may be of importance in human reproduction, however, the

existence of truncated KIR2DL4 receptors, commonly present in Caucasians, that are not expressed on the cell surface (Kikuchi-Maki et al., 2003, Gedil et al., 2005) as well as normal pregnancies reported in women lacking KIR2DL4 is contradictory to this possibility (Gomez-Lozano et al., 2003, Nowak et al., 2011).

KIR genotypes are diverse at an individual and population level due to variability of the gene content of *KIR* haplotypes as well as allelic variation and copy number differences. Two *KIR* haplotype groups have been defined; haplotype A contains a fixed number of genes (n=9) that includes the "framework" genes; *KIR3DL3, -3DP1, -2DL4* and *-3DL2* present in the majority of humans (Middleton and Gonzelez, 2010) and *KIR2DL1, -2DL3, -3DL1, -2DS4* and *-2DP1*; whereas haplotype B is variable in gene content and includes the "framework" genes as well as one or more of the following *KIR* genes: *KIR2DL2, -2DL5, -2DS1, -2DS2, -2DS3, -2DS5* and *-3DS1* (Martin et al., 2004a). Due to a recombination breakpoint between *KIR3DP1* and *-2DL4*, *KIR* haplotypes can be further classified according to the genes encoded by the centromeric (c) region, flanked by *KIR3DL3* and *-3DP1,* and the telomeric (t) region, flanked by *KIR2DL4* and *-3DL2.* Analysis of *KIR* haplotypes (**Figure 1.15**) in Caucasian, Asian, African-American and Hispanic populations have defined four centromeric (cA01 and cB01-B03) and two telomeric (tA01 and tB01) motifs (Pyo et al., 2010, Jiang et al., 2012). *KIR2DL5* is encoded by two genetic loci present in the telomeric and centromeric region, termed *KIR2DL5A* and *KIR2DL5B*, respectively, however, *KIR2DL5B* is not transcriptionally active (Vilches et al., 2000). Similarly KIR2DS3 and KIR2DS5 can also be present in either the telomeric or centromeric region (Hou et al., 2012).

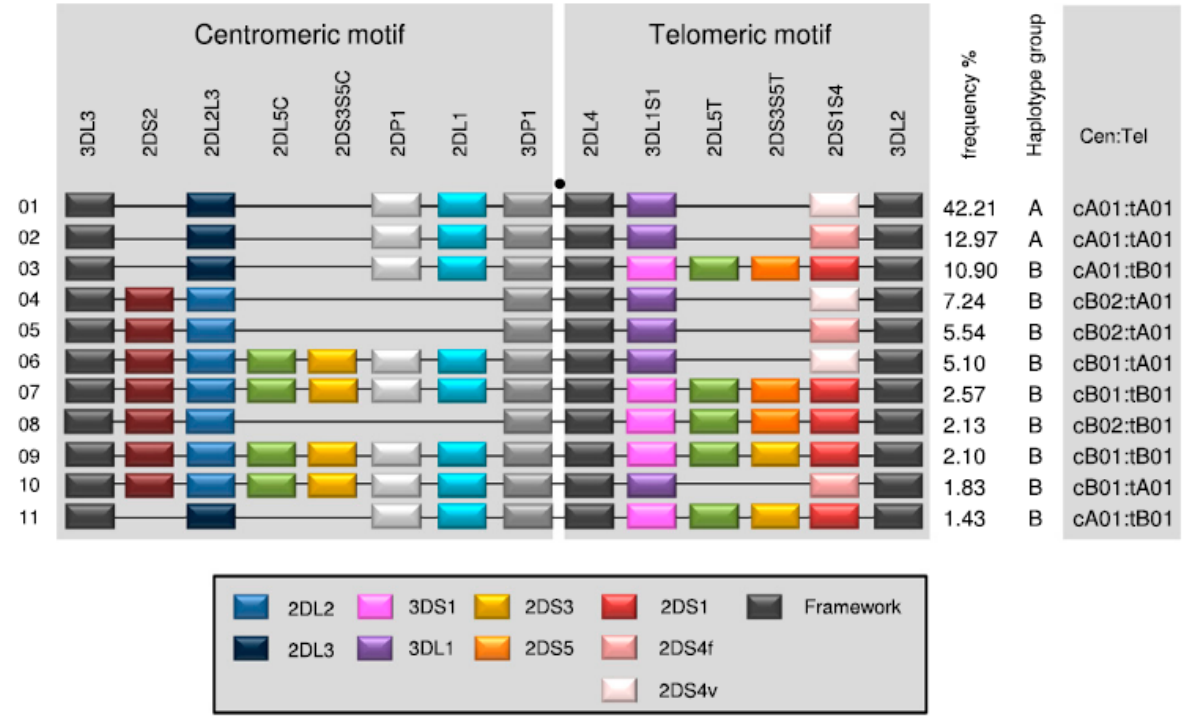

Figure 1.15. KIR haplotypes commonly found in Caucasians. Taken from Jiang et al., 2012 .

KIR genes have further diversified at an allelic level with the number of variants described as follows: *KIR2DL1* (n=48), *-2DL2* (n=30), *-2DL3* (n=55), *-2DL4* (n=52), *-2DL5* (n=48), *-2DS1* (n=16), *-2DS2* (n=22), *-2DS3* (n=15), *-2DS4* (n=31), *-2DS5* (n=18),*-R3DL1* (n=110), *-3DL2* (n=112), *-3DL3* (n=111), *-3DS1* (n=30), *-2DP1* (n=28) and *-3DP1* (n=27) (http://www.ebi.ac.uk/ipd/kir/stats.html). The documented *KIR2DL5* alleles can be further subdivided into *KIR2DL5A* (n=15), *-2DL5B* (n=28) and 5 unassigned allelic variants. Mutagenesis of several *KIR* allelic variants has lead to these genes encoding soluble proteins that are not expressed on the cell surface. For example, *KIR3DP1*004* encodes a secreted protein as a result of recombination between the promoter of *KIR2DL5A* and *-3DP*001* and is the only transcriptionally active pseudogene (Gomez-Lozano et al., 2005). Similarly, several allelic variants of *KIR2DL4* contain a -1 frameshift mutation in the coding region of the cytoplasmic tail that causes the expressed protein to be truncated and soluble (Goodridge et al., 2007). Loss of function due retention of proteins in the endoplasmic reticulum is attributed to the *KIR3DL1*004* allelic variant (Pando et al., 2003). A number of alleles of *KIR2DS4* encode truncated proteins that lack membrane anchoring domains due to a 22 base-pair deletion in exon 5 (Maxwell et al., 2002, Middleton et al., 2007).

In addition to diversity in *KIR* gene repertoires and allelic variation, individuals may carry variable copy numbers of *KIR* genes due to duplications/deletions. This has been shown previously for *KIR2DL4* (Williams et al., 2003), *KIR3DL1/S1* (Pelak et al., 2011, Williams et al., 2003) and more recently for other *KIR* genes in the Caucasian population (Jiang et al., 2012, Vendelbosch et al., 2013). The immunological significance of increased *KIR* copies has been demonstrated for *KIR2DL3* in clearance of HCV (Khakoo et al., 2004) and *KIR3DL1/S1* in the control of HIV-1 replication (Pelak et al., 2011).

1.4.3.2. CD8$^+$ cytotoxic T cell biology

Effector CTLs functionally kill infected targets in an antigen-specific manner via induction of apoptosis (**Figure 1.16**); either by release of lytic granules containing perforin and granzymes (degranulation) or by ligation of target cell death ligands, such as Fas ligand (FasL) or tumour necrosis factor-related apoptosis-inducing ligand (TRAIL) (Deguine and Bousso, 2013, Long et al., 2013). The mechanism of recognition of target cells involves the interaction between the T cell receptor (TCR) and HLA class I molecules expressed on the surface of the CTL and the target cell, respectively (Andersen et al., 2006). Cells productively infected with HIV-1 accumulate viral proteins in the cytoplasm and via constitutive proteolytic processing of cytosolic proteins, mediated by the proteosome, short peptide antigens (9-11 amino acids) are generated that bind to HLA class I molecules (**Figure 1.17**). Peptide-bound HLA class I molecules are then displayed on the cell surface for potential recognition by CTL's that express a unique TCR capable of recognising the peptide antigen in complex with HLA class I molecules. Recognition stimulates the effector functions of the CTL resulting in death of the target cell. CTL detection of virus-infected cells is dependent on presentation of antigenic peptides derived from viral proteins bound to HLA class I A, -B and -C molecules displayed on the surface of target cells.

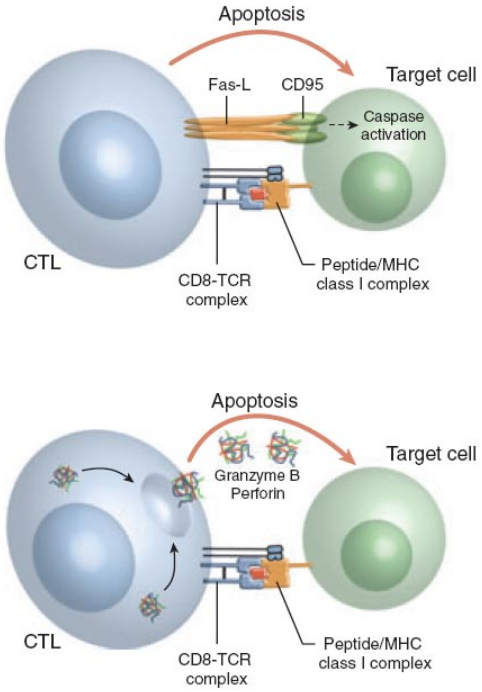

Figure 1.16. Induction of apoptosis in target cells by CTLs either by engagement of death receptor ligands (top) or release of granzyme B and perforin (bottom). Taken from Andersen et al., 2006 .

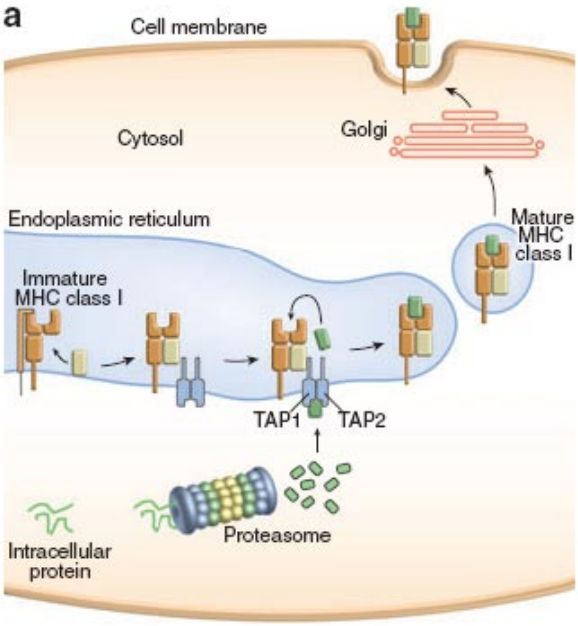

Figure 1.17. Intracellular protein processing and display of antigenic peptides on HLA class I molecules. Taken from Andersen et al., 2006 .

1.4.3.3. Virus/host interaction

1.4.3.3.1. NK cell immune evasion

The antiviral efficacy of NK cells are often underappreciated in HIV-1 infection, however, NK cell responses precede CTL responses during acute HIV-1 infection (Alter et al., 2007) and KIRs in combination with cognate HLA class I ligands, such as KIR3DL1 and HLA-Bw4^{80I}, associate with reduced viral load (Jiang et al., 2013, Pelak et al., 2011, Martin et al., 2007). The synergistic or epistatic interaction between KIR3DS1 and HLA-Bw4^{80I} also associates with low viral load (Pelak et al., 2011, Martin et al., 2002). *KIR3DS1* and *KIR3DL1* are similar in structure in the extracellular domains, however, functional studies have not shown direct interaction of KIR3DS1 with HLA-Bw4^{80I} ligands (Carr et al., 2007, Gillespie et al., 2007, O'Connor et al., 2007). Interestingly, CTLs also express KIRs, such as KIR3DL1 and KIR3DS1, and a recent report shows that KIR3DS1-positive CTLs are effective in targeting HIV-infected cells (Zipperlen et al., 2014).

Although contact between KIRs and HLA class I molecules does not span the peptide binding site and hence KIR-ligand interaction is essentially antigen-independent, there is an enhancing or inhibitory effect on binding affinity mediated by the peptide antigen (**Figure 1.18**). Mutations in viral proteins such as Gag, Vpr, Tat, Vpu, Env and Nef escape NK cell recognition by KIR2DS1/2/3/5, KIR2DL2/3 and KIR3DS1(Alter et al., 2011). Since KIRs can be activating or inhibitory, a plausible explanation for this finding is that mutations in HLA restricted viral epitopes either strengthen the inhibitory signal or dampen activation signals, thereby reducing NK cell activation potential. Sequence variation in the YL9 epitope in p24^{Gag} restricted by HLA-C*03:04 have been found to influence binding to KIR2DL2 inhibitory receptors (van Teijlingen et al., 2014), and similar studies in HLA-B*57 and KIR3DL1 have shown the same effects of amino acid substitutions in the p24^{Gag} TW10 epitope (Brackenridge et al., 2011).

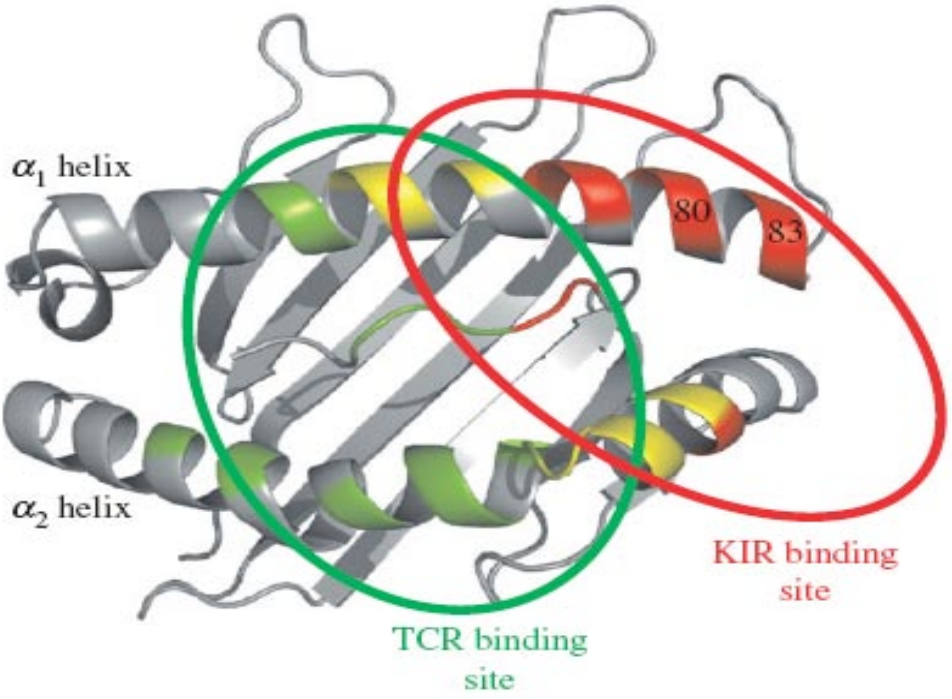

Figure 1.18. HLA class I molecule showing the KIR and TCR binding sites**.** Taken from Parham et al., 2012 .

1.4.3.3.2. CD8$^+$ cytotoxic T cell immune evasion

In HIV-1 infection, it is known that control of viral replication is primarily mediated by virus-specific CD8$^+$ cytotoxic T lymphocytes (CTLs) (Borrow et al., 1994). This is further supported by SIV studies showing that depletion of CTLs results in increased viraemia (Schmitz et al., 1999). Due to the extensive genetic variation in *HLA class I A*, *-B* and *-C* alleles, differences in HIV-1 antigenic peptide binding specificities, has lead to the observation that some HLA class I alleles can be protective or deleterious. Immunogenetic studies in HIV-1 subtype C infected Black South African individuals have established that particular *HLA class I* alleles, such as *HLA-B*57*, *-B*58:01*, *-B*81:01*, *-A*02:05*, *-C*04:01* and *-B*42:01* are significantly associated with lower viral loads, higher CD4$^+$ T cell counts and slower disease progression, whereas the opposite is true for deleterious alleles such as *HLA-B*45:01*, *-B*58:02*, *-B*18* and *-C*06:02* (Kiepiela et al., 2004). One explanation for this observation is that protective HLA class I molecules bind immunogenic peptides derived from conserved viral protein domains which are less tolerant of mutations and therefore a sustained antiviral CTL response maintains a low viral load. Deleterious HLA class I molecules on the other hand target variable protein domains that mutate and allow escape from CTL surveillance. This is supported by work showing that CTL responses directed towards conserved proteins, such as Gag, are associated with lower viral loads than responses

to epitopes derived from variable proteins such as Env (Kiepiela et al., 2007, Wang et al., 2009, Masemola et al., 2004a). It was demonstrated that Gag-specific HLA-B*58:01-restricted CTLs could be identified in individuals controlling HIV-1 replication, whereas, in non-controlling individuals, HLA-B*58:02-restricted CTLs were directed towards Env, and furthermore, selection pressure on viral epitopes was reduced in *HLA-B*58:02*, but not in *HLA-B*58:01* carriers (Ngumbela et al., 2008). Although similar in structure *HLA-B*58:01* is a protective allele, but *HLA-B*58:02* is a deleterious allele.

It has also been observed that the protective benefits of *HLA class I* alleles, such as *HLA-B*57* and *-B*58:01* can be due to fitness costs imposed on replicating virus when CTL escape mutations are introduced into conserved proteins (Crawford et al., 2007). This is further supported by the observation that escape mutations often revert to wild-type when virus is transmitted to HLA mismatched hosts (Leslie et al., 2004). Unfortunately, it has also been shown that the benefits of early escape mutations that reduce viral replicative capacities can be overcome by compensatory mutations that develop during the chronic stage of infection (Brockman et al., 2010, Gijsbers et al., 2013). This has been documented in *HLA-B*58:01* carriers with subtype C HIV-1 infection in South Africa (Chopera et al., 2011). Over time this effect may also become established at a population level as evidenced by a lack of the protective benefits of *HLA-B*57/58:01* in Botswana compared to South Africa due to the longer duration of the epidemic in Botswana (Payne et al., 2014).

Recent data has shown that the timing of epitope presentation to CTLs may be of importance for long-term control of HIV-1 replication. Protective HLA alleles not only target conserved viral proteins, but also target $p24^{Gag}$ epitopes derived from incoming capsid proteins delivered to the cytoplasm following envelope fusion (Kloverpris et al., 2013a). This was first noted in SIV studies where early CTL responses to Gag where observed prior to integration and viral protein expression (Sacha et al., 2007). Following integration of the HIV-1 genome, due to *Rev*-dependent regulation of transcription, Rev, Tat and Nef proteins are expressed early in the replication cycle, while Gag and Pol are expressed at late stages (Karn and Stoltzfus, 2012). Expression of Env is also late and additional RNA splicing events are required. Attempts to relocate Gag epitopes from late to early proteins in integrated forms of virus showed variability in the timing of epitope presentation, meaning that it was not possible to confirm that lower viral loads associate with CTL responses before Nef-mediated down-regulation of HLA-A and -B molecules (Balamurugan et al., 2013). It therefore seems

likely that CTL targeting of HIV-1 core proteins derived from infecting virions is critical for long term control of virus. Interestingly, HIV-1 Nef proteins are incorporated into the virion by interaction with the $p6^{Pol}$ transframe protein, which may be an attempt to interfere with HLA class I presentation of antigens derived from core components of the infecting virion (Costa et al., 2004).

In addition, an important finding in infected individuals carrying protective *HLA* alleles is that immune pressure by other HLA alleles in the individual's repertoire, are co-operative or additive in the overall control of viral replication (Matthews et al., 2012, Leslie et al., 2010). Particular haplotypes, such as *HLA-C*04:01-B*81:01, -C*12:03-B*39:10, -A*74:01-B*53:03* associate with improved viral control compared to the individual alleles (Leslie et al., 2010). Slow disease progression based on $CD4^+$ T cell count associates with haplotypes *HLA-A*74:01-B*15:03, -B*14:02-C*08:02* and *-A*74:01-B*15:03-C*02:02*, while accelerated disease progression associates with haplotypes *HLA-A*30:02-B*45:01, -A*30:02-C*16:01, -B*53:01-C*04:01, -B*15:10-C*03:04* and *-B*58:01-C*03:02* (Sampathkumar et al., 2014).

1.4.4. HIV-1 treatment

1.4.4.1. Antiretroviral drug therapy

The development of antiretroviral drugs (ARVs) for the treatment of HIV-1 infection has made the largest impact in the limiting of viral transmission. Although ARVs are only suppressive and do not cure HIV-1 infection, the development of different classes of compounds that target different aspects of the viral replication cycle (**Figure 1.19**) has provided a large repertoire of combinational therapies that has changed the status of HIV-1 infection from an untreatable fatal disease to a manageable chronic illness (Cohen et al., 2008). The first drug approved for HIV-1 treatment was zidovudine, a nucleoside reverse transcriptase inhibitor (NRTI), belonging to a class of compounds known as DNA chain terminators. To date there are 7 approved NRTIs including lamivudine (3TC), emtricitabine (FTC), stavudine (d4T), didanosine (ddI), abacavir (ABC), zidovudine (AZT) and zalcitabine (ddC), although ddC is no longer marketed.

These drugs require intracellular phosphorylation to generate a nucleotide triphosphate that serves as a substrate for reverse transcriptase. The first nucleotide reverse transcriptase inhibitor (NtRTI) to be developed, tenofovir (TDF), already contains the first phosphate group and hence only requires double phosphorylation for activity. Collectively there are 8 nucleotide/nucleoside reverse transcriptase inhibitors (nRTIs) currently approved by the FDA (http://www.fda.gov/ForPatients/Illness/HIVAIDS/Treatment/ucm118915.htm). The early use of NRTIs in monotherapeutic and dual therapy settings lead to the discovery of drug resistance mutations that rendered decreased sensitivity to therapy. Due to the error prone nature of reverse transcription, randomly generated nucleotide substitutions in the reverse transcriptase gene were able to select for amino acid substitutions in the enzyme capable of either discriminating analogue molecules from normal nucleotides or facilitating excision of incorporated analogues (Arts and Hazuda, 2012).

The development of other classes of antiretroviral therapy such as the non-nucleoside reverse transcriptase inhibitors (NNRTIs) and protease inhibitors (PIs) lead to the development of combinational therapy known as highly active antiretroviral therapy (HAART) that was able to suppress virus replication sufficiently in order to limit the emergence of resistant virus. Currently 5 NNRTIs are licensed for use including delavirdine (DLV), efavirenz (EFV), nevirapine (NVP), etravirine (ETR) and rilpivirine (RPV) (http://www.fda.gov/ForPatients/Illness/HIVAIDS/Treatment/ucm118915.htm). ETR and RPV are second generation NNRTIs that are insensitive to some of the resistance mutations common to the first generation drugs. Licensed PIs include lopinavir (LPV), amprenavir (APV), fosamprenavir (FOS-APV), ritonovir (RTV), atazanavir (ATV), indinivir (IDV), nelfinavir (NFV), saquinavir (SQV), darunavir (DRV) and tipranavir (TPV), although FOS-APV and APV are no longer marketed (http://www.fda.gov/ForPatients/Illness/HIVAIDS/Treatment/ucm118915.htm).

Other targets in the viral replication cycle include inhibition of proviral integration mediated by integrase and viral entry facilitated by the interaction of gp41 with host membrane and interaction of gp120 with the co-receptor CCR5. Two licensed integrase inhibitors; raltegravir (RAL) and dolutegravir (DTG) are currently approved (http://www.fda.gov/ForPatients/Illness/HIVAIDS/Treatment/ucm118915.htm). The entry inhibitor, enfuvirtide (T-20), targets the mechanism of action of gp41 mediated fusion of viral envelope with the cell membrane, while the co-receptor antagonist, Maraviroc, inhibits CCR5

co-receptor binding to gp120 spike proteins that precedes envelope fusion (http://www.fda.gov/ForPatients/Illness/HIVAIDS/Treatment/ucm118915.htm).

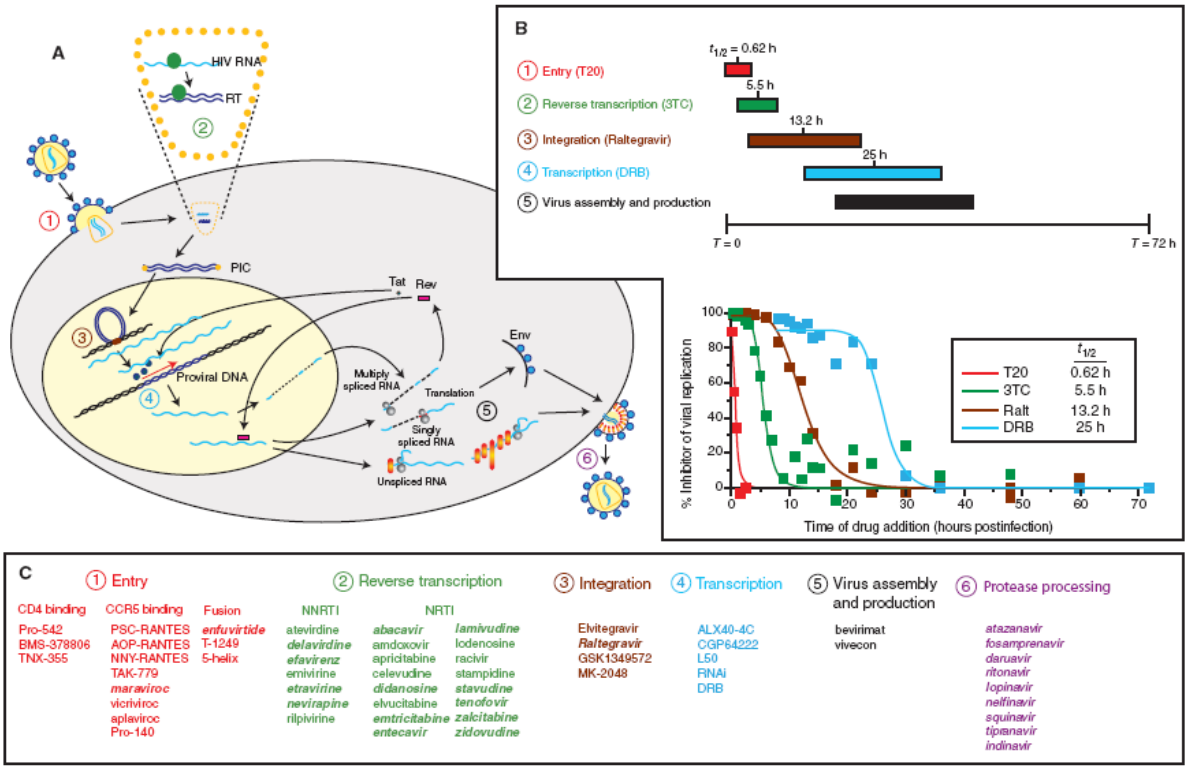

Figure 1.19. Drug targets in the viral replication cycle (a), relative time frame of drug action (b) and list of approved and investigational drugs targeting the viral life cycle (c). Taken Arts and Hazuda, 2012 .

1.4.4.2. Immune hypersensitivity reaction

Although ARV therapy has reduced morbidity and mortality associated with HIV-1 infection, treatment success can be compromised by toxicity, resistance and immune complications such as immune reconstitution inflammatory syndrome (IRIS) and immune hypersensitivity reaction (IHR). In susceptible individuals, IHR develops rapidly following administration of the ABC or NVP. IHR is potentially life-threatening due to excessive activation of autoimmune CTLs and increased production of pro-inflammatory cytokines (Pompeu et al., 2012). Symptoms include fever, rash, nausea, vomiting, diarrhoea, abdominal pain, dyspnoea, sore throat, cough and malaise (Yuen et al., 2008). In extreme circumstances, failure to withdraw the drug, or re-initiation of therapy, can lead to multiple organ failure and death. ABC IHR associates with the presence of *HLA-B*57:01*, although, not all carriers develop a response (Hetherington et al., 2002, Mallal et al., 2002, Martin et al., 2004b, Phillips et al., 2005). Similarly, NVP IHR associates with the presence of *HLA-B*35:05* in Thai populations (Chantarangsu et al., 2009, Yuan et al., 2011); *HLA-C*04* in Thai, Chinese

and African populations (Likanonsakul et al., 2009, Gao et al., 2012, Yuan et al., 2011); *HLA-C*08* in Japanese populations (Gatanaga et al., 2007) and *HLA-B*14-C*08* haplotypes in Caucasian populations (Littera et al., 2006). Recently, the *HLA-C*04:01* allelic variant was specifically associated with NVP IHR in an African population (Carr et al., 2013a). An association of *HLA class II* alleles, such as *HLA-DRB1*01* (Vitezica et al., 2008) and *HLA-DRB1*0101* (Martin et al., 2005), with cutaneous rash/hepatotoxicity in Caucasians treated with NVP, has also been described and more recently, a study in South African populations has linked NVP hepatotoxicity with *HLA-DRB1*0102* as well as *HLA-B*58:01* (Phillips et al., 2013).

The immunopathogenetic mechanism for ABC IHR is well characterised and demonstrates that drug-specific CTLs are activated in response to ABC exposure as shown experimentally by incubation of drug-treated HLA-B*57:01-expressing target cells with donor CTLs from *HLA-B*57:01* carriers (Chessman et al., 2008). Furthermore, key residues 114D and 116S, located in the peptide-binding pocket of HLA-B*57:01, determine the *HLA class I* allele specificity, since closely related HLAs, such as HLA-B*57:02 or HLA-B*57:03, contain 114N and 116Y residues and are insensitive to ABC (Chessman et al., 2008). Subsequent crystallographic studies revealed that ABC binds to HLA-B*57:01 via chemical interaction with residues 114D and 116S (**Figure 1.20**), with the latter position playing a dominant role (Illing et al., 2012, Norcross et al., 2012, Ostrov et al., 2012, Chessman et al., 2008). In what is termed the “altered peptide repertoire” mechanism, the presence of ABC in the HLA-B*57:01 peptide-binding pocket modifies the antigen specificity, leading to the presentation of self-peptides to autoimmune CTLs (Pompeu et al., 2012). An alternative and not necessarily mutually exclusive mechanism, known as the “hapten” model, suggests that drug metabolites covalently linked to host proteins leads to the presentation of drug-modified self-peptides restricted by HLA-B*57:01 (Pompeu et al., 2012). Indeed, ABC-modified proteins generated through aldehyde intermediates have been identified (Grilo et al., 2013) and furthermore, drug-specific CTL clones can be activated by overnight incubation of antigen presenting cells with ABC, but not following short-duration exposure, suggesting that an alternative pathway dependent on drug metabolism and antigen processing may exist (Bell et al., 2013). It remains to be determined whether similar mechanisms are involved in the development of NVP IHR.

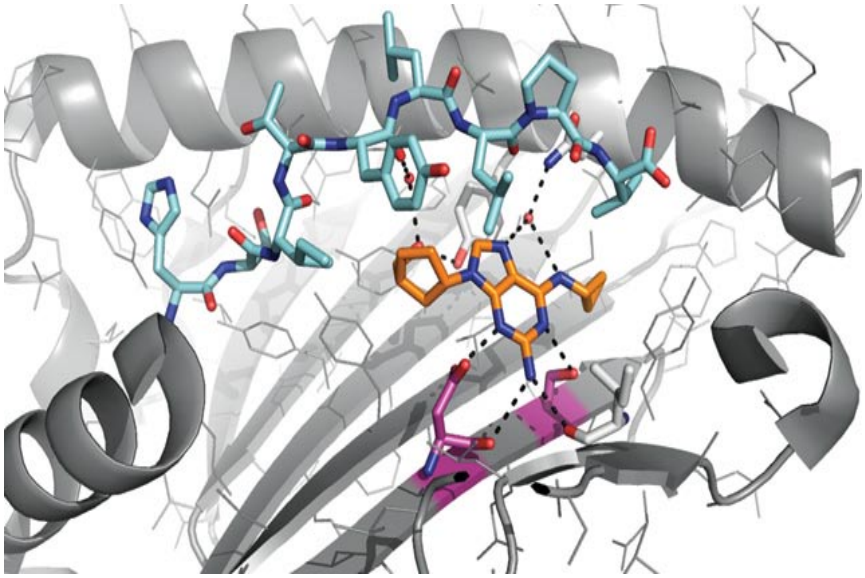

Figure 1.20. Interaction of abacavir (orange and blue) with amino acid residues 114D and 116S (purple and red) and immunogenic peptide (light and dark blue with red) located in the peptide binding cleft of HLA-B*57:01 determined by X-ray crystallography. Taken from Illing et al., 2012 .

1.5. Problem statement

Differences in host genetics can play modulatory roles in disease outcome in individuals and populations. HLA class I molecules are central in modulating both innate and adaptive immune responses to HIV-1exposure, via respective NK cell and CTL activities. In turn, the effectiveness of these responses influences both HIV-1 susceptibility to infection as well as disease progression and associated transmission risk. Due to the extensive polymorphism in *HLA class I* genes at a population and individual level, the efficacy of the immune response is dependent on inherited *HLA class I* repertoires, since particular alleles can be classified as protective or deleterious. Immunogenic peptides derived from viral proteins are restricted by particular HLA class I molecules that mediate immune detection of infected cells by CTLs and also influence NK cell surveillance via stimulation of KIRs. The KIR family of receptors are also highly polymorphic at population and individual levels and notably KIRs and HLAs assort to different chromosomes and are not necessarily inherited in complementary pairs. The effectiveness of NK cell responses is further complicated by variability in KIR gene repertoires, allelic variation and gene copy number variation. Both NK cell and CTL activities are also affected by the propensity of HIV-1 to induce mutations in immunogenic peptides that are restricted by HLA class I molecules that in turn affects both CTL and NK cell responses. Lastly, particular *HLA class I* alleles are also implicated in the

development of adverse events during treatment of HIV-1 infection with antiretroviral drugs via the development of immune-mediated hypersensitivity reactions.

HIV-1 infection has a low prevalence in South African (SA) Caucasian (SAC), Indian (SAI) and Mixed ancestry (SAM) populations, but is high in SA Black (SAB) populations. This raises the possibility that, in part, genetic factors may contribute to HIV-1 susceptibility or transmission risk. SAB and SAC populations are well characterised, but little data on *KIR* and *HLA class I* genotypes are available for the SAI and SAM populations. A subset of HLA class I alleles are also associated with immune hypersensitivity reactions to antiretroviral drugs such as Abacavir and Nevirapine that are in use in South Africa. This PhD study aims to generate *HLA class I* and *KIR* genotypic data for healthy controls in the understudied populations for subsequent analysis and comparison to data available from SAB and SAC populations. Viral factors such as immune escape from virus-specific HLA class I restricted CTL detection and NK cell surveillance are also of importance, however, due to the low prevalence of HIV-1 in SAI and SAM populations, an investigation of evolution in HIV-1 Gag epitopes from infected SAB individuals will be included and analysed in the context of *KIR* and *HLA class I* allele repertoires.

The three main hypotheses that this study will address are as follows:

1. That the differential HIV-1 prevalence rates in SAI, SAM, SAC and SAB populations may be affected, in part, by differences in *HLA class I* and *KIR* genotypes associated with HIV-1 susceptibility and transmission risk that can be extrapolated from genotypic data obtained from studies of healthy individuals from these populations.

2. That the frequencies of *HLA class I* alleles associated with immune hypersensitivity reaction to Abacavir and Nevirapine determined for healthy SAI, SAM, SAC and SAB populations are significant and justifies the need for pre-treatment screening.

3. That viral evolution in HIV-1 Gag contributes to disease progression in infected individuals carrying different *HLA class I* alleles due to viral immune escape from CTL and NK cell responses.

1.6. Study objectives

To test these hypotheses the following study objectives will be met.

1. Sequence based typing methods will be used to determine *HLA class I A, -B* and *-C* allelic variants present in healthy SAI and SAM populations for comparison to available data for SAB and SAC populations (**Chapter 3**).

2. To develop and/or optimise qualitative rapid real-time AS-PCR based methods for the determination of *HLA* ligand and *KIR* genotype profiles in healthy controls from SAI and SAM populations for comparison to available data for SAB and SAC populations (**Chapter 3**).

3. To optimise quantitative real-time AS-PCR based methods for the determination of gene copy number variation for *KIR3DL1/S1* and *KIR2DS4f/v* in SAI and SAM population (**Chapter 3**).

4. To develop rapid real-time AS-PCR based screening assays to identify *HLA class I* alleles associated with Abacavir and Nevirapine immune hypersensitivity reaction, apply the method to healthy controls in SAI and SAM populations, and with SAC and SAC data included, determine the potential number of individuals who are at risk (**Chapter 4**).

5. To compare and analyse the population data generated for SAI and SAM, with SAB and SAC data included, with respect to genotypes associated with HIV-1 acquisition and transmission to generate a hypothesis that host genetic factors might influence the HIV-1 prevalence in South Africa (**Chapter 5**).

6. To design RT-PCR and sequencing methods to characterise the evolution of Gag over time in HIV-1-positive SAB individuals and to assess the contribution of viral immune escape mutations from CTL and NK cell responses in the context of *KIR* and protective/deleterious *HLA class I* alleles (**Chapter 6**).

Chapter 2:
Materials and Methods

2.1. Study cohorts

2.1.1. ESKOM cohort

Buffy coats stored at -80°C were available for healthy individuals of SAM (n=40) and SAI (n=22) ethnicity from the ESKOM cohort recruited in a previous study (Paximadis et al., 2012). Additional volunteers were recruited prospectively from staff/students at the National Institute for Communicable Diseases (NICD), the University of the Witwatersrand (WITS) Medical School and academic hospital in Johannesburg, to make up totals of 50 individuals for the SAI and SAM populations. Signed consent was obtained from each participant and ethics approval for the study was granted by the Human Research Ethics Committee (Medical) of the WITS medical faculty. Clearance certificate M130873 appears in **Appendix A**.

2.1.2. Progressor cohort

HLA class I A, -B and *-C* genotypes, $CD4^+$ T cell counts and viral load data were available for individuals (n=113) classified as progressors, based on $CD4^+$ T cell counts declining to <350 cells/µl prior to initiation of treatment. These participants were part of a larger cohort of untreated HIV-1-positive SAB individuals recruited from Soweto, Johannesburg, for a previous study (Martinson et al., 2014). Based on *HLA class I* genotypic data, 9/113 (8.0%) progressors were *HLA-B*58:01* carriers and 22/113 (19.5%) were *HLA-B*58:02* carriers. Baseline and follow up plasma samples stored at -80°C, taken prior to initiation of treatment, were available for *HLA-B*58:01* carriers (n=5) and for *-B*58:02* carriers (n=13). Ethics approval for the study was granted by the Human Research Ethics Committee (Medical) of the WITS medical faculty. Clearance certificate M130873 appears in **Appendix A**.

2.2. Methods

2.2.1. DNA extraction

Genomic DNA was extracted from buffy coats or whole blood using the QiaAmp DNA blood mini kit (Qiagen, Hilden, Germany) as per the manufacturer's instructions and quantified using a Nanodrop spectrophotometer instrument (Thermo Scientific, Waltham, USA).

2.2.2. *KIR* genotyping assay

A real-time AS-PCR assay developed previously (Hong et al., 2011) was used in combination with additional PCR primers to generate *KIR* genotype profiles. The assay was enhanced by inclusion of a second set of primers designed to target *KIR2DL3, -2DS1* and *-3DP1*, since the original assay used a single set of primers for detection of these *KIRs*, but dual primer sets for all other *KIRs*. The use of dual primer sets for detection of all the *KIRs* improves coverage of allelic variants. The new *KIR3DP1* primer sets were designed to include detection of allelic variants *KIR3DP1*00902* and *-3DP1*004* that were missed by the original *KIR3DP1* primer set. All PCR primers were designed using *KIR* alignments available at the Immuno Polymorphism Database (IPD) (Robinson et al., 2013a). The internal control gene, *GALC* (*Galactosylceramidase*), was found to be compatible in multiplex with the new *KIR* primer sets. (A complete list of the dual sets of PCR primers used in the assay as well as *KIR* allelic variants missed or detectable by some primers appears in **Table 2.1**.

Real-time AS-PCR amplification was performed in 96-microwell PCR plates on an ABI7500 real-time PCR instrument (Life Technologies, Carlsbad, USA). Reaction volumes were 5 µl containing 5 ng of genomic DNA, 2X Universal SYBRgreen Mastermix (Roche, Mannheim, Germany), 0.5 µM *KIR* and 0.5 µM *GALC* forward/reverse primers (Inqaba Biotec, Pretoria, South Africa). Cycling conditions were an initial incubation of 95°C for 10 minutes followed by 30 cycles of 95°C for 15s and 60°C for 60s. Melting curve analysis was performed to determine the presence/absence of the gene of interest (**Figure 2.1**). True negatives showed the presence of control amplicon only due to a lower melting temperature. If no melting curves were present the sample was repeated.

Table 2.1. PCR primer sets used for the *KIR* real-time AS-PCR genotyping assay

Well	Primer name	Primer sequence (5'-3')	*KIR* alleles not detected
1	*2DL1* f1[1]	GTTGGTCAGATGTCATGTTTGAA	*013N
	2DL1 r1[1]	GGTCCCTGCCAGGTCTTGCG	
2	*2DL1* f2[1]	TGGACCAAGAGTCTGCAGGA	*005
	2DL1 r2[1]	TGTTGTCTCCCTAGAAGACG	*0040102
3	*2DL2* f1[1]	CTGGCCCACCCAGGTCG	*004,*00601,*00602
	2DL2 r1[1]	GGACCGATGGAGAAGTTGGCT	*00303,*004
4	*2DL2* f2[2]	AAACCTTCTCTCTCAGCCCA	
	2DL2 r2[2]	GCCCTGCAGAGAACCTACA	*009
5	*2DL3* f1[2]	AGACCCTCAGGAGGTGA	
	2DL3 r1[2]	CAGGAGACAACTTTGGATCA	*010,*017
6	*2DL3* f2[4]	GGCCCAGAAGTGCCCTCT	
	2DL3 r2[4]	ATTGGAGCTGGCAACCCA	
7	*2DL4* f1[1]	CAGGACAAGCCCTTCTGC	
	2DL4 r1[1]	CTGGGTGCCGACCACT	
8	*2DL4* f2[1]	ACCTTCGCTTACAGCCCG	
	2DL4 r2[1]	CCTCACCTGTGACAGAAACAG	
9	*2DL5* f1[1]	GCGCTGTGGTGCCTCG	
	2DL5 r1[1]	GACCACTCAATGGGGGAGC	
10	*2DL5* f2[1]	TGCAGCTCCAGGAGCTCA	
	2DL5 r2[1]	GGGTCTGACCACTCATAGGGT	
11	*2DS1* f1a[2]	TCTCCATCAGTCGCATGAG	
	2DS1 f1b[2]	TCTCCATCAGTCGCATGAA	
	2DS1 r1[2]	GGTCACTGGGAGCTGAC	
12	*2DS1* f2[4]	CCGCTCTTGAGCGAGCAA	
	2DS1 r2[4]	CAACTCCACCTCCAGGCCTATATA	
13	*2DS2* f1[1]	TTCTGCACAGAGAGGGGAAGTA	
	2DS2 r1[1]	GGGTCACTGGGAGCTGACAA	
14	*2DS2* f2[1]	CGGGCCCCACGGTTT	*00104
	2DS2 r2[1]	GGTCACTCGAGTTTGACCACTCA	
15	*2DS3* f1[3]	AAACCTTCTCTCTCAGCCCA	
	2DS3 r1[3]	GCATCTGTAGGTTCCTCCT	
16	*2DS3* f2[1]	CTATGACATGTACCATCTATCCAC	
	2DS3 r2[1]	AAGCAGTGGGTCACTTGAC	
17	*2DS4* f1[1]	CTGGCCCTCCCAGGTCA	
	2DS4 r1[1]	TCTGTAGGTTCCTGCAAGGACAG	
18	*2DS4* f2[1]	GTTCAGGCAGGAGAGAAT	
	2DS4 r2[1]	GTTTGACCACTCGTAGGGAGC	
19	*2DS5* f1[1]	TGATGGGGTCTCCAAGGG	
	2DS5 r1[1]	TCCAGAGGGTCACTGGGC	*003
20	*2DS5* f2[1]	ACAGAGAGGGGACGTTTAACC	
	2DS5 r2[1]	ATGTCCAGAGGGTCACTGGG	
21	*2DP1* f1[1]	GTCTGCCTGGCCCAGCT	
	2DP1 r1[1]	GTGTGAACCCCGACATCTGTAC	
22	*2DP1* f2[1]	CCATCGGTCCCATGATGG	
	2DP1 r2[1]	CACTGGGAGCTGACAACTGATG	
23	*3DL1* f1[1]	CGCTGTGGTGCCTCGA	*009;*042;*057
	3DL1 r1[1]	GGTGTGAACCCCGACATG	
24	*3DL1* f1a[2]	CCATCGGTCCCATGATGCT	*054
	3DL1 f1b[2]	CCATTGGTCCCATGATGCT	*054
	3DL1 f1c[2]	TCCATCGGTCCCATGATGTT	*054
	3DL1 r2[2]	CCACGATGTCCAGGGGA	
25	*3DL2* f1[1]	CAAACCCTTCCTGTCTGCCC	*01301;*01302;*014; *024;*025;*037;*040; *046;*052;*054
	3DL2 r1[1]	GTGCCGACCACCCAGTGA	
26	*3DL2* f2[1]	CCCATGAACGTAGGCTCCG	
	3DL2 r2[1]	CACACGCAGGGCAGGG	*018
27	*3DL3* f1[1]	GTCAGGACAAGCCCTTCCTC	
	3DL3 r1[1]	GAGTGTGGGTGTGAACTGCA	
28	*3DL3* f2[1]	TTCTGCACAGAGAGGGGATCA	
	3DL3 r2[1]	GAGCCGACAACTCATAGGGTA	
29	*3DS1* f1[1]	AGCCTGCAGGGAACAGAAG	
	3DS1 r1[1]	GCCTGACTGTGGTGCTCG	

30	*3DS1f* 2a[2]	CATCGGTTCCATGATGCG	
	3DS1f 2b[2]	CATCAGTTCCATGATGCG	
	3DS1 r2[2]	CCACGATGTCCAGGGGA	
31	*3DP1* f1[2]	GTACGTCACCCTCCCATGATGTA	*00902;*004
	3DP1 r1[4]	GCACTCCCTCCCTCTATTCCT	
32	*3DP1* f2[4]	CATCCTCCTCTCTAAGGTGGC	
	3DP1 r2[2]	GAAAACGGTGTTTCGGAATAC	
All	*GALC* f[3]	TTACCCAGAGCCCTATCGTTCT	
	GALC r[3]	GTCTGCCCATCACCACCTATT	

GALC, galactosylceramidase; *KIR*, killer-cell immunoglobulin-like receptor; PCR, polymerase chain reaction; AS-PCR, allele-specific PCR. [1]Primers taken from Martin and Carrington, 2008 ; [2]Primers taken from Vilches et al., 2007 ; [3]Primers taken from Alves et al., 2009 ; [4]Primers designed for this study.

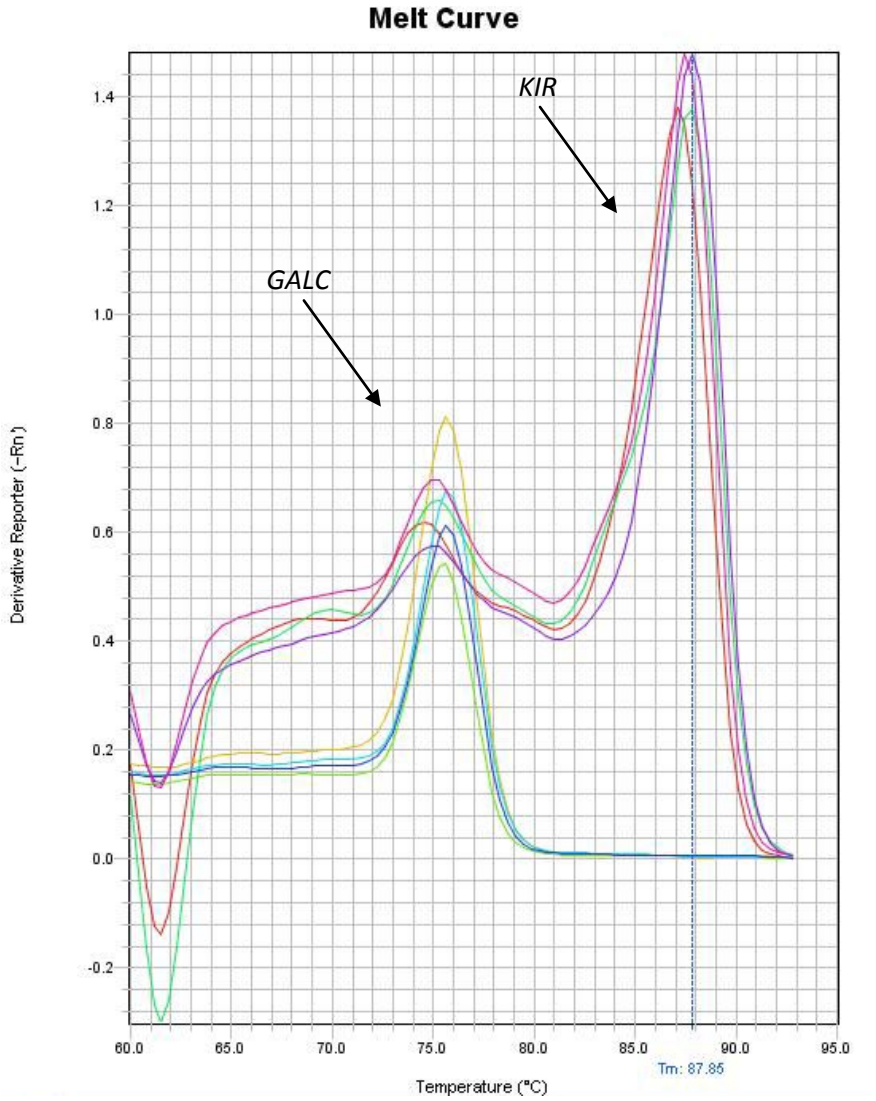

Figure 2.1. Melting curve analysis of *KIR* gene of interest and *GALC* control gene.

2.2.3. *HLA-Bw4^{80I}/Bw4^{80T}/C1/C2* ligand genotyping assay

A real-time AS-PCR assay developed previously (Hong et al., 2011) was used with additional primer modifications to generate *HLA-Bw4^{80I}/Bw4^{80T}/C1/C2* ligand profiles based on genetic variation at codon-80. A complete list of the PCR primers used in the assay appears in **Table 2.2**. Further optimisation and enhancement of the *HLA* ligand assay was done by modifying the PCR primers as follows: (i) the *HLA-A non-Bw4^{80T}* reverse primer was extended to 20 bases, (ii) the *HLA-B* specific forward primers were replaced with a single re-designed common forward primer, (iii) the *HLA-B Bw6^{80N}* reverse primer was extended to 19 bases, (iv) the *HLA-B Bw4^{80TA}* reverse primer was extended to 20 bases and (v) optional *HLA-B C1^{80N}* reverse primers were included to pair with the common *HLA-B* specific forward primer to detect the rare subset of *HLA-B* alleles that encode HLA-C1 epitopes. The assay makes use of 3'locked nucleic acid (LNA) modification to the primers (indicated by [ ] in **Table 2.2**) to reduce false-positive amplification due to mispriming. The redesigned common *HLA-B* forward primer when paired with the *HLA-Bw4^{80TA}* reverse primer showed improvement in the correct identification of two *HLA-B*38:02*-positive DNA samples obtained from the International Histocompatibility Working Group (IHWG) (www.ihwg.org) that had previously failed to type correctly in the original assay (Hong et al., 2011). Detection of *HLA-B* allelic variants encoding C1 epitopes were successfully tested using three *HLA-B*46* positive samples obtained from the IHWG, however, due to a nucleotide polymorphism in *HLA-B*73:01,* we included an additional reverse primer to detect this variant, but no control sample was available for testing.

Since *HLA class I* genotyping was done at a later time (discussed in **Section 2.2.7.1.)**, all real-time AS-PCR results were confirmed to be in agreement with *HLA* typing results. However, we did notice a potential false positive result for two samples that were homozygous for *HLA-B Bw4^{80TA}*, where a small melting peak was visible when *HLA-B Bw4^{80IA}* primers were used. The *HLA-B Bw4^{80IA}* and *HLA-B Bw4^{80TA}* primers differ by a single base at the 3'end, and even though 3'LNA modified primers are used, a small amount of misprimed product may be generated in the presence of two copies of *HLA-B Bw4^{80TA}*. This was not seen in samples with a single copy of *HLA-B Bw4^{80TA}* and the false-positive peak height is considerably smaller than a true *HLA-B Bw4^{80IA}* result. Since the melting curve peak height is different for each primer set, we recommend that analysis is also done comparing all samples positive for the same primer set so that a false-positive call can be avoided.

Table 2.2. PCR primer sets used for the determination of genetic variability at codon-80 of *HLA-ABC.*

Well	Primer name	Primer sequence (5'-3')
1	*HLA-A Bw4*^{80I} f1[1]	CCATTGGGTGTCGGGTTTC[C]
	HLA-A/B Bw4^{80I} r1[1]	CTCTGGTTGTAGTAGCGGAGCGCG[A]
2	*HLA-A non-Bw4*^{80T} f1[1]	AATCAGTGTCGTCGCGGTC[G]
	HLA-A non-Bw4^{80T} r1[2]	GTTGTAGTAGCCGCGCAGG[G]
3	*HLA-B* f1[2]	GCGAGGGGACCGCAGGC[G]
	HLA-B Bw6^{80N} r1[2]	TTGTAGTAGCCGCGCAGG[T]
4	*HLA-B* f1[2]	GCGAGGGGACCGCAGGC[G]
	HLA-A/B Bw4^{80I} r1[1]	CTCTGGTTGTAGTAGCGGAGCGCG[A]
5	*HLA-B* f1[2]	GCGAGGGGACCGCAGGC[G]
	HLA-B Bw4^{80TA} r1[2]	GTTGTAGTAGCGGAGCGCG[G]
6	*HLA-B* f1[2]	GCGAGGGGACCGCAGGC[G]
	HLA-B Bw4^{80TL} r1[1]	CTCTGGTTGTAGTAGCGGAGCAGG[G]
7	*HLA-C C1*^{80N} f1[1]	AGCCAATCAGCGTCTCCGC[A]
	HLA-C C1^{80N} r1[1]	GCTCTGGTTGTAGTAGCCGCGCAG[G]
8	*HLA-C C2*^{80K} f1[1]	CCATTGGGTGTCGGGTTCT[A]
	HLA-C C2^{80K} r1[1]	GCTCTGGTTGTAGTAGCCGCGCAG[T]
9	*HLA-B* f1[2]	GCGAGGGGACCGCAGGC[G]
	HLA-B C1^{80N} r1[2]	GCAGGTTCCGCAGGCTC[A]
	HLA-B C1^{80N} r2[2]	GCAGGTTCCGCAGGCCC[A]
All	*ALB* f1[1]	TCGATGAGAAAACGCCAGTAA
	ALB r1[1]	ATGGTCGCCTGTTCACCAA

ALB, albumin; *HLA*, human leukocyte antigen; PCR, polymerase chain reaction; AS-PCR, allele-specific PCR; [ ] denotes 3' locked nucleic acid (LNA) base modification. [1]Primers taken from Hong et al., 2011 ; [2]Primers modified or designed for this study.

Real-time AS-PCR amplification was performed in 96-microwell PCR plates on an ABI7500 real-time PCR instrument (Life Technologies, Carlsbad, USA). Reaction volumes were 5 µl containing 5 ng of genomic DNA, 2X Universal SYBRgreen Mastermix (Roche, Mannheim, Germany), 0.5 µM *HLA* ligand and 0.5 µM *ALB* control gene forward/reverse primers (Inqaba Biotec, Pretoria, South Africa). Cycling conditions were an initial incubation of 95°C for 10 minutes followed by 35 cycles of 95°C for 15s, 62°C for 15s and 72°C for 30s. Melting curve analysis was performed after generation of *HLA* ligand amplicons to determine the presence/absence of the gene of interest. True negatives showed the presence of control amplicon only due to a lower melting temperature compared to the gene of interest.

2.2.4. *CCR5/CCR5Δ32* genotyping assay.

CCR5/CCR5Δ32 genotypes were determined by a real-time AS-PCR threshold cycle (Ct) shift assay. Allele-specific forward primers *CCR5WTf* and *CCR5Δ32f* and a common reverse primer *CCR5r* were designed (**Table 2.3**) to amplify the full-length wild-type and 32bp-deletion variant of *CCR5* using genomic sequences available in Genbank (www.ncbi.nlm.nih.gov/genbank). Genomic DNA samples representing homozygous *CCR5*, homozygous *CCR5Δ32* and heterozygous *CCR5/CCR5Δ32* genotypes, previously characterised by DNA sequencing (Picton et al., 2010), were available as controls and for evaluation of the assay. Additional confirmation of results was done using an agarose gel-

based amplicon size discrimination method. Briefly, genomic DNA was amplified with PCR primers spanning the 32bp-deletion, *CCR5F* and *CCR5R* (**Table 2.3**), followed by 2% agarose gel electrophoresis, ethidium bromide staining and visualisation of amplicons under UV light.

Table 2.3. PCR primers used to determine *CCR5/CCR5Δ32* genotypes

Primer name	Primer sequence (5'-3')
CCR5WTf	TGCAGCTCTCATTTTCCATACAGTC
CCR5Δ32f	TGCAGCTCTCATTTTCCATACATTA
CCR5r	GATTCCCGAGTAGCAGATGACC
CCR5F	GTGGTGGCTGTGTTTGCGT
CCR5R	CATTTCGACACCGAAGCAGAG

Real-time AS-PCR amplification was performed in 96-microwell PCR plates using an ABI7500 real-time PCR instrument (Life Technologies, Carlsbad, USA). Reaction volumes were 5 µl containing 5 ng of genomic DNA, 2X Universal SYBRgreen Mastermix (Roche, Mannheim, Germany) and 0.5 µM *CCR5* forward/reverse primers (Inqaba Biotec, Pretoria, South Africa). Cycling conditions were an initial incubation of 95°C for 10 minutes followed by 45 cycles of 95°C for 15s, 62°C for 5s and 72°C for 30s and results were determined by analysis of relative shifts of threshold cycle (Ct) values (**Figure 2.2**).

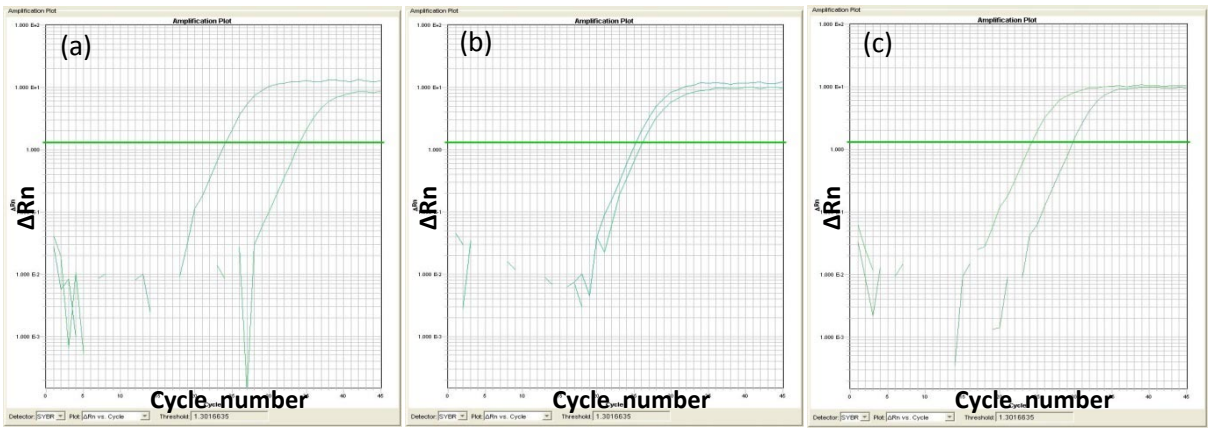

Figure 2.2. Schematic representation of real-time PCR amplification curves of the *CCR5* gene and *CCR5Δ32* allelic variant. The SYBRgreen amplification plots of the threshold cycle (Ct) shift assay were designed to discriminate wild-type *CCR5* from the *CCR5Δ32* deletion mutant. (a) Homozygous CCR5 wild-type, primer mispriming results in amplifying 10 cycles later; (b) *CCR5/CCR5Δ32* heterozygous, no primer mispriming and similar Ct values; (c) Homozygous *CCR5Δ32* mutant, primer mispriming results in amplification 6 cycles later. ΔRn; derivative of reporter.

2.2.5. Copy number variation assay for *KIR2DS4f/v* and *KIR3DL1/S1*

Published primers and probe specific for *KIR2DS4f/v* (Jiang et al., 2012) and *KIR3DL1/S1* (Pelak et al., 2011) were used in combination with the primers and probe specific for the human *beta-globin* (*BGB*) reference gene (Shostakovich-Koretskaya et al.,

2009) in a probe hydrolysis-based relative quantification real-time AS-PCR assay to determine gene copy number variation. A complete list of the PCR primers and probes used in the copy number variation assay are listed in **Table 2.4**. To facilitate target and reference gene multiplexing, *KIR*-specific probes were labelled at the 5'-end with the fluorochrome VIC, while *BGB* probes were labelled with the fluorochrome FAM. Standard curves were generated in two-fold dilution steps over a range of 20 - 0.625 ng of genomic DNA. The control DNA for the *KIR2DS4f/v* assay was derived from a haplotype AA1 donor carrying a single copy each of *KIR2DS4f* and *KIR2DS4v*, while DNA extracted from the CEPH cell lines GM11840 and GM12752, containing single copies of *KIR3DL1* and *-3DS1* was used for the *KIR3DL1/S1* assay (Pelak et al., 2011).

Table 2.4. List of PCR primers and probe sequences used for real-time quantitative AS-PCR copy number determination of *KIR3DL1/S1* and *KIR2DS4f/v.*

Well	Primer name	Primer sequence (5'-3')
1	*KIR3DL1* f1[1]	GCCTCGTTGGACAGATCCA[T]
	KIR3DL1 r1[1]	TAGGTCCCTGCAAGGGCA[A]
	KIR3DL1/S1 probe[1]	VIC-GGGTCTCCAAGGCCAATTTCTCCAT-MGB
2	*KIR3DS1* f1[1]	CTCGTTGGACAGATCCATG[A]
	KIR3DS1 r1[1]	GTCCCTGCAAGGGCA[C]
	KIR3DL1/S1 probe[1]	VIC-GGGTCTCCAAGGCCAATTTCTCCAT-MGB
3	*KIR2DS4f* f1[2]	CCGGAGCTCCTATGACATG
	KIR2DS4f/v r1[2]	TGACGGAAACAAGCAGTGG[A]
	KIR2DS4f/v probe[2]	VIC-AACATTCCAGGCCGACTTTCCTCTG-MGB
4	*KIR2DS4v* f1[2]	CCTTGTCCTGCAGCTCCAT
	KIR2DS4f/v r1[2]	TGACGGAAACAAGCAGTGG[A]
	KIR2DS4f/v probe[2]	VIC-AACATTCCAGGCCGACTTTCCTCTG-MGB
All	*BGB* f1[3]	TCGCTTTCTTGCTGTCCAATTTCTA
	BGB r1[3]	ATGCTCAAGGCCCTTCATAATATCC
	BGB probe[3]	FAM-CCTAAGTCCAACTACTAAACTG-MGB

VIC and FAM, probe reporter fluorescent dyes; MGB, minor groove binding protein (non-fluorescent quencher); AS-PCR, allele-specific PCR; *BGB*, human beta-globin, [ ] denotes 3' locked nucleic acid (LNA) base modification. [1]Primers/probes taken from Pelak et al., 2011 ; [2]Primers taken from Jiang et al., 2012 ; [3]Primers/probes taken from Shostakovich-Koretskaya et al., 2009 .

Real-time AS-PCR amplification was performed in 96-microwell PCR plates using an ABI7500 real-time PCR instrument (Life Technologies, Carlsbad, USA). Reaction volumes were 5 µl containing 5 ng of genomic DNA, 2X Lightcycler 480 probes Mastermix (Roche, Mannheim, Germany), 0.5 µM *KIR3DL1/S1* or *KIR2DS4f/v* forward/reverse primers (Inqaba Biotec, Pretoria, RSA), 0.5 µM *BGB* forward/reverse primers (Inqaba Biotec, Pretoria, RSA), 0.1 µM VIC-labelled *KIR3DL1/S1* or *KIR2DS4f/v* probe and 0.1 µM FAM-labelled *BGB* probe (Life Technologies, Carlsbad, USA). Cycling conditions were an initial incubation of 95°C for 10 minutes followed by 40 cycles of 95°C for 15s and 60°C for 60s. Gene copy numbers were estimated from the ratio of the relative amounts of target and reference genes. All samples were run in duplicate and repeated if the copy numbers were not in agreement.

2.2.6. Determination of *HLA-C 3'UTR* +263 insertion/deletion variants

The *HLA-C 3'UTR* was amplified by conventional PCR using published primers (Kulkarni et al., 2011). PCR cycling conditions were 94°C for 2 minutes followed by 30 cycles of 94°C for 15s, 65°C for 30s and 72°C for 90s and a final extension of 72°C for 7 minutes. PCR products were purified using Agencourt Ampure XP magnetic bead separation (Becton-Dickenson, Franklin lakes, USA). Sequence products were generated using a published reverse primer (Gentle et al., 2013) and the BigDye terminator v3.1 cycle sequencing kit (Life Technologies, Carlsbad, USA). Sequence products were purified using ethanol-sodium acetate precipitation and resolved on an ABI3100 PRISM Genetic Analyser instrument (Life Technologies, Carlsbad, USA). DNA sequences were analysed using Sequencher v4.5 software (Gene Codes Corporation, Ann Arbor, USA).

2.2.7. High resolution *HLA class I ABC* genotyping

2.2.7.1. ESKOM cohort

The SBT resolver kit (Conexio Genomics, Fremantle, Australia) was used to generate *HLA class I* A, B and C amplicons for DNA sequencing as described by the manufacturers. PCR products were purified using Agencourt Ampure XP magnetic bead separation (Becton-Dickenson, Franklin lakes, USA). Sequence products were generated using the BigDye terminator v3.1 cycle sequencing kit (Life Technologies, Carlsbad, USA), purified using ethanol-sodium acetate precipitation and resolved on an ABI3100 PRISM Genetic Analyser instrument (Life Technologies, Carlsbad, USA). Alleles were assigned using ASSIGN-SBT v4.7 software (Conexio Genomics, Fremantle, Australia). When encountered, ambiguous *HLA class I* genotypes were resolved using a "common" allele predictive approach based on allele frequency data available from the same population or ancestral populations. Ambiguities for the SAI population were resolved using data available from a previous *HLA* class I study of SAI individuals (n=51) as well as data available from studies in mainland India (Gonzalez-Galarza et al., 2011). Ambiguities found in the SAM population were resolved by referring to published *HLA class I* alleles detected in this population (Salie et al., 2014) as well as referring to *HLA class I* allele frequency data available for two of the ancestral populations, SAC and SAB (Paximadis et al., 2012). In addition, since *HLA-A, -B*

and -*C* alleles were determined for each sample, common *HLA class I* haplotypes could be used to support the linkage of specific alleles (Gonzalez-Galarza et al., 2011).

2.2.7.2. Progressor cohort

High resolution *HLA class I* genotypes were determined for the study participants using the Roche 454 GS FLX next-generation sequencing platform (Roche Diagnostics, Indianapolis, USA). This work was outsourced to the HLA typing service available at the Mallal Laboratory, Institute for Immunology and Infectious Diseases, Murdoch University, Western Australia. Briefly, genomic DNA was used to generate DNA sequences spanning exons 2 and 3 of *HLA-A*, *-B* and *-C* alleles by PCR simultaneously incorporating DNA adaptor molecules. Amplicons were purified, quantified by fluorimetry, diluted and subjected to emulsion PCR. Sequence data was then generated on a Roche 454 GS FLX instrument as per instructions by the manufacturers. Alleles were assigned by matching exon consensus sequences to a reference library.

2.2.8. Determination of HLA-B signal peptide p2-M/T variants

The presence of methionine or threonine at position-2 of the HLA-B signal peptide was determined by sequence analysis of exon 1 that is available through amplicon sequencing results from the *HLA-B* SBT typing method

2.2.9. Statistical analysis

Gene frequencies were determined by direct counting. Significance values and odds ratios were calculated using the two-tailed Fisher's exact test or the Mann-Whitney test available in GraphPad Prism v4.02 software (GraphPad software, San Diego, USA). Due to the hypothesis-generating nature of this study, correction for multiple testing was not applied, since, given the complex and multifactorial nature of HIV-1 susceptibility/transmission, we considered it more important to identify potential factors that may play a role in HIV-1 acquisition and donor infectivity rather than simply dismissing these leads as due to chance variations brought about by multiple comparisons.

2.2.10. Phylogenetic analysis

Neighbour-joining phylogenetic trees were constructed using the Kimura 2-parameter model and n=1000 bootstrap settings available in the software *MEGA* v5 (Tamura, Dudley, Nei and Kumar., 2007).

2.2.11. *HLA class I* IHR real-time AS-PCR assay development

Due to the extensive genetic polymorphism in *HLA-B*57* allelic variants (n=85), the design of AS-PCR primers capable of direct discrimination of alleles of *HLA-B*57:01* (n=16) from other *HLA-B*57* alleles necessitated the use of a novel strategy. In order to detect the presence of any *HLA-B*57* allele, two group-specific primer sets were designed. Two primer sets were included to detect the presence of *HLA-B*57* alleles encoding amino-acid residues 114D and 116S. It has been shown that these two key amino-acid residues located in the peptide-binding pocket of HLA-B*57:01 determine the *HLA class I* allele specificity in ABC IHR, since closely related HLA molecules, such as HLA-B*57:02 and -B*57:03, that contain 114N and 116Y, are insensitive to ABC (Chessman et al., 2008). A final primer set to identify the guanine (G) variant of the *HCP5* rs2395029 single nucleotide polymorphism (SNP), also termed "*HCP5* 568G", was also included, since *HCP5* 568G is known to be in strong linkage disequilibrium with *HLA-B*57:01* (Colombo et al., 2008, Rodriguez-Novoa et al., 2010). A positive PCR amplification for all five primer sets is highly indicative of the presence of *HLA-B*57:01*. However, since linkage between *HCP5* 568G and *HLA-B*57:01* is not absolute (Badulli et al., 2012, Colombo et al., 2008), it is useful to confirm the presence of *HLA-B*57* as well as codons 114D and 116S in cases where *HCP5* 568G occurs in *HLA-B*57:01*-negative individuals or is absent in *HLA-B*57:01*-positive individuals.

The primers used to detect alleles of *HLA-B*57* (n=85) and *-B*35:05* (n=3) were designed using sequence alignments of *HLA-B* alleles (n=3,089) available in the International Immunogenetics project (IMGT)/HLA database (Robinson et al., 2013b), last updated in October 2013 (**Table 2.6**). To detect the presence of codons 114D and 116S in *HLA-B*57:01*, a common forward primer B57E was designed to pair with two allele-specific reverse primers 114D and 116Sa (**Table 2.6**). Due to a sequence polymorphism in codon 116S in *HLA-B*57:01:02*, an additional reverse primer 116Sb was also included. Notably, other *HLA* alleles also encode 114D and 116S, hence, the B57E primer was designed to

selectively target *HLA-B*57* specifically, however *HLA-B*57:11* is missed. This does not affect detection of *HLA-B*57:01*, however, codons 114D and 116S are also present in other *HLA-B*57* alleles and it is not certain whether these alleles also associate with ABC IHR, although, these residues are present in HLA-B*58:01 molecules, which is insensitive to ABC. As a precaution, individuals testing positive for the presence of other *HLA-B*57* alleles that encode 114D and 116S should either be excluded from ABC therapy or be closely monitored for early signs of clinical development of IHR. The primers used to detect alleles of *HLA-C*04* (n=221) and *-C*08* (n=121) were designed using sequence alignments of *HLA-C* alleles (n=2,032) available in the IMGT/HLA database (Robinson et al., 2013b), last updated in October 2013 (**Table 2.6**). Although attempts were made to avoid false-positive PCR amplifications, due to the large number of *HLA-B* and *-C* alleles documented, it was inevitable that a small number of alleles would be co-amplified by the primer sets used in the assay (**Table 2.7**). The number of alleles co-amplified was minimized to 21/3089 (0.7%), 12/3089 (0.4%), 18/2032 (0.9%) and 9/2032 (0.4%) for *HLA-B*57*, *-B*35:05*, *-C*04* and *-C*08*, respectively.

The real-time AS-PCR assay was performed in 96-microwell PCR plates using an ABI7900HT real-time PCR instrument (Life Technologies, Carlsbad, USA). Reaction volumes of 5 µl containing 5 ng of genomic DNA, 2X universal SYBRgreen mastermix with ROX passive reference dye (Roche, Mannheim, Germany), 0.5 µM *HLA* gene-specific forward/reverse primers and 0.5 µM *ALB* forward/reverse primers (Inqaba Biotec, Pretoria, RSA). Cycling conditions were an initial denaturation of 95°C for 10 minutes followed by 35 cycles of 95°C for 15s, 65°C for 5s and 72°C for 60s. Melting curve analysis was performed to determine the presence/absence of the gene of interest. The presence of a melting peak corresponding to the positive control gene (*ALB*) and absence of a melting peak corresponding to the gene of interest was considered a true negative (**Figure 2.2**). The sample was repeated if no melting peaks were visible at all.

Table 2.5. PCR primers used to identify *HLA class I* alleles associated with ABC and NVP IHR

Well	Primer name	Primer sequence (5-3')
1	B57A (sense) B57B (antisense)	CGCGAGTCCGAGGATGG[C] TCTCGGTAAGTCTGCGCGG[A]
2	B57C (sense) B57D (antisense)	GAGACACGGAACATGAAGGCCT[C] CCACGTCGCAGCCATACATCA[C]
3	B57E (sense) B57114D (antisense)	GGGTCTCACATCATCCAGG[T] CCGTCGTAGGCGGACTGGT[C]
4	B57E (sense) 116Sa (antisense) 116Sb (antisense)	GGGTCTCACATCATCCAGG[T] TAATCCTTGCCGTCGTAGGCG[G] TAATCCTTGCCGTCGTAGGCA[G]
5	HCP5GF (sense) HCP5GR (antisense)	GGCCTTGGGAAATTACCTGG TACACATTACAGCTGCC[C]
6	B3505F (sense) B3505R (antisense)	CGGGTCTCACACCCTCCAGAG[C] CGGAGGAGGCGCCCGTC[G]
7	C04A (sense) C04B (antisense)	AGCGACGCCGCGAGTCC[A] GCGGTCCAGGAGCGCAG[A]
8	C04C (sense) C04D (antisense)	TGAGGTATTTCTCCACATCCGTGTCC[T] GCCGTCGTAGGCGAACTGGT[T]
9	C08A (sense) C08B (antisense)	ACGACACGCAGTTCGTG[C] CGCAGGTTCCGCAG[G]
10	C08C (sense) C08D (antisense)	AGGCACAGACTGACCGAGTG[A] GGTCCAGGAGCGCAGGTCCT[C]
All	ALBF (sense) ALBR (antisense)	TCGATGAGAAAACGCCAGTAA ATGGTCGCCTGTTCACCAA

[ ] denotes 3'Locked Nucleic Acid (LNA) modification; AS-PCR, allele-specific PCR; HLA, human leukocyte antigen; ABC, abacavir; NVP, nevirapine; IHR, immune hypersensitivity reaction.

Table 2.6. *HLA class I* alleles predicted to be missed or co-amplified by the AS-PCR primers used

Primer set	***HLA* alleles missed**	***HLA* alleles co-amplified**
B57A/B	*B*57:16*; *-57:45*; *57:51*	*B*15:16:01-03; 15:17:01:01; 15:17:01:02; 15:17:02; 15:67; 15:95; 15:162; 15:168; 15:177; 15:196; 15:208; 15:216; 15:222; 15:230; 15:254; 15:268; 15:273; 58:36*
B57C/D	*B*57:05; 57:11*	*B*58:14*
B57E/114D	*B*57:11*	*B*35:208; 55:14; 58:14*
B57E/116S	*B*57:11*	*B*35:208; 55:14; 58:14*
B3505F/R	none	*B*35:58; 35:66; 35:72; 35:89; 35:97; 35:114; 35:232; 40:02:08; 40:24; 41:04; 53:14; 58:18*
C04A/B	*C*04:01:05; 04:01:07; 04:01:09; 04:01:23; 04:04:02; 04:160; 04:94:02*	*C*03:04:25; 08:01:07; 08:02:07; 08:33:02; 12:03:20; 14:02:03; 14:03; 14:08; 14:10; 14:22; 14:35N; 14:38; 14:41; 14:53; 14:54; 15:77; 16:01:06*
C04C/D	*C*04:03:01; 04:03:02; 04:06; 04:16; 04:23; 04:42; 04:54; 04:80; 04:100; 04:103; 04:107; 04:108; 04:140; 04:147; 04:160*	*C*05:78*
C08A/B	*C*08:10; 08:13; 08:16:01; 08:16:02; 08:25; 08:94*	*C*01:43; 07:101; 07:148; 07:161*
C08C/D	*C*08:10; 08:01:06; 08:01:07; 08:02:04; 08:02:07; 08:33:01; 08:33:02*	*C*01:34; 03:87; 07:02:35; 07:41; 12:02:07*

HLA, human leukocyte antigen; AS-PCR, allele-specific PCR

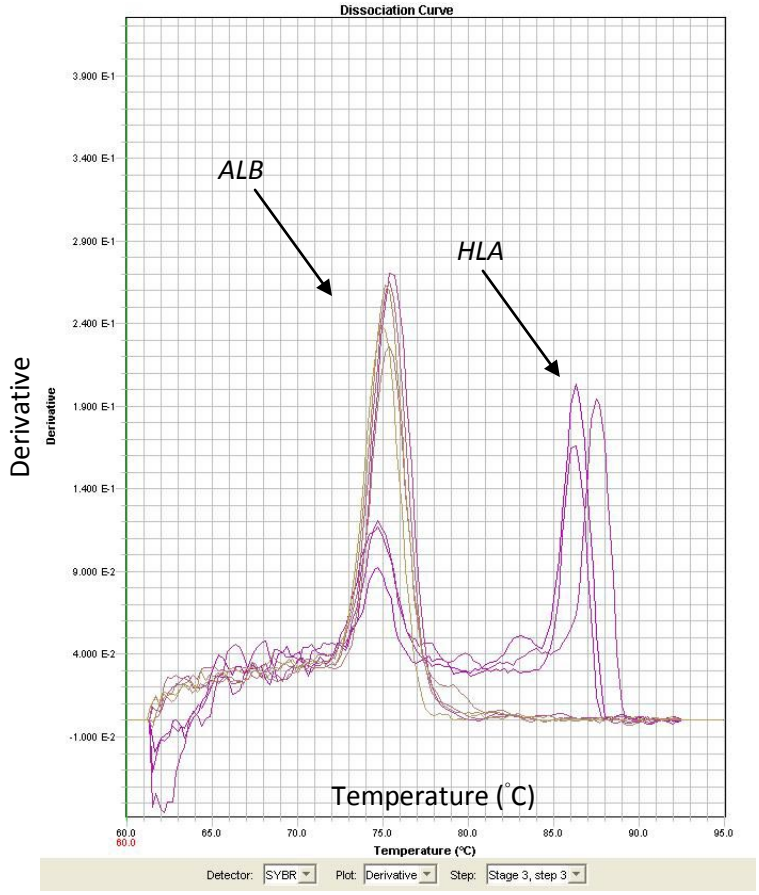

Figure 2.3. Melting curve analysis of *HLA* gene of interest and *ALB* control gene.

2.2.12. Cumulative analysis of protective/deleterious factors

2.2.12.1. HIV-1 acquisition

A cumulative analysis of factors associated with HIV-1 acquisition was done for each individual by assigning equal weighting to protective factors known to be associated with resistance including *KIR2DL2/3+HLA-C1/C2*, *KIR3DL1/S1+HLA-Bw4*, specific *HLA-ABC* alleles and HLA-B p2-T/T phenotypes; while deleterious factors known to be associated with susceptibility included *KIR2DL2/3+HLA-C1/C2*, *KIR3DL1/S1+HLA-Bw4*, HLA-B p2-M/M phenotypes and specific *HLA-ABC* alleles. *KIR-HLA* ligand combinations associated with HIV-1 acquisition have been reported (Guerini et al., 2011, Habegger de Sorrentino et al., 2013, Jennes et al., 2006) and include *KIR2DL2*(+)/*KIR2DL3*(+)/*HLA-C2/C2*(+), *KIR3DL1*(+)/*KIR3DS1*(-)/*HLA-B Bw6/Bw6*(+), *KIR3DL1*(-)/*KIR3DS1*(+)/*HLA-B* $Bw4^{80I}$(+), *KIR3DL1*(+)/*KIR3DS1*(+)/*HLA-B* $Bw4^{80I}$(+)*/Bw6*(+) and *KIR3DL1*(+/-)/*KIR3DS1*(+)/*HLA-A/B* $Bw4^{80I}$(+) that associate with resistance and *KIR2DL3*(+)/*KIR2DL2*(-)/*HLA-C1/C2*(+)

and *KIR3DL1*(+)/*KIR3DS1*(-)/*HLA-A/B* $Bw4^{80I/80T}$(+) that associate with susceptibility. Due to some *KIR3DL1/S1* and *HLA-Bw4* ligand combinations having overlapping profiles, a single protective *KIR-HLA* ligand genotype was assigned to an individual when more than one protective combination was present. HLA-B signal peptide p2-T/T phenotypes were counted as resistance factors, while the HLA-B signal peptide p2-M/M phenotype was counted as a susceptibility factor (Merino et al., 2013, Merino et al., 2012). Although heterozygous HLA-B signal peptide p2-M/T phenotypes have been shown to confer an intermediate degree of protection from HIV-1 acquisition (Merino et al., 2012), we did not include heterozygotes in our analysis. A number of studies have reported associations of individual *HLA class I A, -B* and *-C* alleles with heterosexual HIV-1 acquisition and where possible, taking all 6 alleles into account, these data were used to determine whether an individual had an overall resistance or susceptible *HLA class I* profile scoring the total number of protective vs. deleterious alleles. The relevant *HLA class I* alleles found in this study are listed alongside references in **Appendix B**. In this analysis, individuals with equal numbers of protective and deleterious *HLA* alleles or with alleles that were not associated with acquisition were considered neutral. In some cases contradictory reports on associations of *HLA class I* alleles and HIV-1 acquisition were observed, most likely due to population or viral subtype differences. In these cases, we opted for reports in populations closest in match to our populations under study or if there was an excess of studies in favour of an association.

2.2.12.2. HIV-1 viral load setpoint

A similar cumulative analysis of factors associated with HIV-1 viral load setpoint (VLS) was done for each individual by assigning equal weighting to known protective factors associated with low VLS including *CCR5/CCR5Δ32* heterozygosity (van Manen et al., 2011), *KIR3DL1/S1+HLA-Bw4, HLA-C 3'UTR* +263-del/del genotypes and specific *HLA-ABC* alleles; while deleterious factors associated with high VLS included *KIR3DL1/S1+HLA-Bw4, KIR2DS4f* (Merino et al., 2011, Merino et al., 2014), *HLA-C 3'UTR* +263 ins/ins genotypes and specific *HLA-ABC* alleles. Reported *KIR-HLA* ligand combinations associated with VLS (Pelak et al., 2011, Jiang et al., 2013) were used in the cumulative analysis and included *KIR3DL1*(+/-)/*KIR3DS1*(+)/*HLA-B* $Bw4^{80I}$(+), *KIR3DL1*(+)/*KIR3DS1*(+)/*HLA-B* $Bw4^{80I}$(+) and *KIR3DL1*(+)/*KIR3DS1*(-)/*HLA-B* $Bw4^{80I}$(+) associated with low VLS and *KIR3DL1*(+)/*KIR3DS1*(+/-)/*HLA-B* $Bw4^{80I}$(-) and *KIR3DL1*(+)/*KIR3DS1*(-)/*HLA-B* $Bw4^{80I}$(-) associated with high VLS. Since the genotypes involving *KIR3DL1/S1* and *HLA-Bw4* had

overlapping profiles, an overall analysis per individual was done by only counting a protective genotype once, even if there was more than one protective genotype present. *HLA-C 3'UTR* +263-ins/ins homozygosity was counted as a protective factor associated with low VLS, while *HLA-C 3'UTR* +263 del/del homozygosity was counted as a deleterious factor associated with high VLS (Kulkarni et al., 2011). Although *HLA-C 3'UTR* +263 ins/del heterozygosity has been shown to exert an intermediate influence in reducing VLS, we only considered homozygotes in our analysis. A number of studies have reported associations of individual *HLA class I A, -B* and *-C* alleles with VLS following heterosexual HIV-1 acquisition and these data were used to determine whether an individual had an overall low or high VLS *HLA class I* profile by adding up the total number of protective vs. deleterious alleles. The relevant *HLA class I* alleles found in this study and references are listed in **Appendix C**. Individuals with equal numbers of protective and deleterious alleles or with alleles that have not been shown to associate with VLS were considered neutral. Again, in some cases contradictory reports on associations of *HLA class I* alleles and HIV-1 VLS were observed, most likely due to population or viral subtype differences. In these cases, we opted for reports in populations closest in match to our populations under study or if there was an excess of studies in favour of an association**.**

2.2.13. HIV-1 RNA isolation, reverse transcription and *gag* sequencing

A set of nested PCR primers and internal sequencing primers were designed to amplify and sequence near full-length gag genes encompassing partial $p17^{Gag}$ coding regions to the Gag stop codon (**Table 2.5**). PCR and sequencing primers were optimally designed using alignments of sub-Saharan subtype C *gag* sequences available at the Los Alamos HIV Database (http://www.hiv.lanl.gov). Viral RNA was extracted from 500 µl of plasma using the high performance MagNA Pure LC total nucleic acid isolation kit and a MagNA Pure LC instrument (Roche Diagnostics, Indiapolis, USA). The elution volume was set to 50 µl and viral RNA was stabilised by addition of MS2 carrier RNA (Roche Diagnostics, Indiapolis, USA) to a final concentration of 10 ng/µl. Complementary DNA (cDNA) was synthesised using the outer reverse primer, gagoutR. An initial primer annealing step was performed by adding 50 pmol of gagoutR in 0.5 µl to 12.5 µl of viral RNA and heating to 65°C for 10 minutes before cooling on ice for 2 minutes. The annealed RNA mixture was made up to a 20 µl reaction volume by addition of 0.5 µl of reverse transcriptase enzyme and 4 µl of 5X buffer from the Transcriptor cDNA kit (Roche Diagnostics, Indiapolis, USA), 2 µl of 10 mM

dNTP mix (Roche Diagnostics, Indiapolis, USA) and 0.5 µl of Protector RNAse inhibitor (Roche Diagnostics, Indiapolis, USA). The cDNA reaction was incubated at 50°C for 1 hour and inactivated at 80°C for 5 minutes.

A 25 µl first-round PCR amplification was carried out using the FastStart High-Fidelity PCR system (Roche Diagnostics, Indiapolis, USA) by addition of 3 µl of cDNA to 22 µl of PCR mastermix containing 2.5 µl 10X buffer, 0.5 µl of 10 mM dNTP mix, 0.25 µl DNA polymerase, 1 µl containing 30 pmol of outer forward primer (gagoutF), 1 µl containing 30 pmol of outer reverse primer (gagoutR) and 16.75 µl of PCR grade water. Cycling conditions were 95°C for 2 minutes followed by 40 cycles of 95°C for 30sec, 55°C for 30s and 72°C for 3 minutes. A 50 µl second-round PCR amplification was carried out using the FastStart High-Fidelity PCR system (Roche Diagnostics, Indiapolis, USA) by addition of 5 µl of PCR product to 45 µl of PCR mastermix containing 5 µl 10X buffer, 1 µl of 10 mM dNTP mix, 0.5 µl DNA polymerase, 2 µl containing 30 pmol of inner forward primer (gaginF), 2 µl containing 30 pmol of inner reverse primer (gaginR) and 34.5 µl of PCR grade water. Cycling conditions were 95°C for 2 minutes followed by 40 cycles of 95°C for 30s, 58°C for 30s and 72°C for 3 minutes.

Second-round PCR products were purified using Agencourt Ampure XP magnetic bead separation (Becton-Dickenson, Franklin lakes, USA) and eluted in 50 µl water. Sequence products were generated using 2 µl of purified amplicon in a 10 µl reaction volume containing 5 pmol of sequencing primer in 1 µl, 0.5 µl of BigDye terminator v3.1 cycle sequencing mastermix (Life Technologies, Carlsbad, USA), 1.75 µl of 5X sequence reaction buffer (Life Technologies, Carlsbad, USA) and 4.75 µl of PCR grade water. The two inner PCR primers and six overlapping internal sequencing primers were used to generate *gag* sequences. Cycling conditions were 96°C for 60s followed by 25 cycles of 96°C for 10s, 52°C for 10s and 60°C for 4 minutes. Sequence products were purified using ethanol-sodium acetate precipitation and resolved on an ABI3100 PRISM Genetic Analyser instrument (Life Technologies, Carlsbad, USA). DNA sequences were analysed using Sequencher v4.5 software (Gene Codes Corporation, Ann Arbor, USA).

Table 2.7. PCR primers used for RT-PCR amplification and sequencing of HIV-1 *gag.*

Primer name	Primer sequence (5'-3')
gagoutF	TGTTAAAACACTTAGTATGGGCAAG
gagoutR	TTACTTTGATAAAACCTCCAATTCC
gaginF	TTGCACTTAACCCTGGCCTTTTAGA
gaginR	ATTTATTTCTTCTAATACTGTATCATCTGC
gagsegF1	GCTCTTCAGACAGGAACAGA
gagsegF2	GGACATCAAGCAGCCATGCA
gagsegF3	GAAGAAATGATGACAGCATG
gagsegR1	CCTGCTATGTCACTTCCCCT
gagsegR2	TTTCCACATTTCCAACAGCC
gagsegR3	GACAAGGGGTCGCTGCCAAA

2.2.14. *Gag* sequence analysis

A total of 18 participants were included in the analysis of the *gag* sequences generated which comprised of 16 sequences at baseline, 16 sequences at the 6 month time-point and 10 sequences at the 12 month time-point. Sequences were aligned to a consensus South African subtype C reference sequence using BioEdit software v7.0.9.0 (Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41:95-98.). Deduced amino acid sequences were used to determine the presence of relevant polymorphisms cited in the literature. CTL escape mutations in known HLA class I-restricted epitopes for HLA-B*58:01 as well as in other epitopes restricted by other HLA molecules in the participant's HLA repertoire were also determined.

Chapter 3:
Immunogenetic characterisation of healthy individuals of South African Indian, Mixed ancestry, Caucasian and Black ethnicity

3.1. Introduction

The South African (SA) resident population is extensive in size (n=52,982,000) and ethnically diverse; consisting of individuals of SA Black (SAB), SA mixed ancestry (SAM), SA Caucasian (SAC) and SA Indian/Asian (SAI) ethnicities that respectively represent 79.8%, 9.0%, 8.7% and 2.5% of the total population (www.statssa.gov.za/publications/P0302/P03022013.pdf). The SAC population has its origins in the Western Cape region ~360 years ago when Dutch sailors established a refreshment station in Cape Town in 1652, while the SAM population has its origins around the same time period due to admixture of multiple ethnic groups incorporating Caucasian, Black African, South Asian and Indonesian ancestry (de Wit et al., 2010, Patterson et al., 2010). The SAI population was established ~150 years ago in the KwaZulu-Natal region by a founder population of immigrants from mainland India (www.sahistory.org.za).

There is a high HIV-1 burden in South Africa with a national HIV-1 prevalence of 12.2% estimated in 2012, however prevalence rates are not proportionate among the different SA populations. In 2008, the prevalence of HIV-1 was estimated at 0.3%, 0.3%, 1.7% and 13.6% in SAI, SAM, SAC and SAB populations, respectively (www.mrc.ac.za/pressreleases/2009/sanat.pdf). In 2012, these estimates changed to 0.8%, 3.1%, 0.3% and 15.0% in SAI, SAM, SAC and SAB populations, respectively (www.hsrc.ac.za/en/research-outputs/view/6871), indicating differential influences on HIV-1 spread within each population, which could reflect potential host genetic differences.

Immunogenetic studies in different local populations are important in terms of immunity in HIV-1 infection (Martin and Carrington, 2013) as well as the effect of immune genes on prevention/treatment options, such as immune hypersensitivity reactions to antiretroviral drug therapy (Bharadwaj et al., 2012). Of particular importance is the role played by human leukocyte antigen (HLA) class I molecules found on the surface of most nucleated cells, that regulate both adaptive and innate immune responses following exposure to intracellular pathogens, via activation of $CD8^+$ cytotoxic T cell (CTLs) and natural killer (NK) cells. The three *HLA class I* genes, (*HLA-A, -B* and *-C*) are encoded by a highly polymorphic gene cluster mapping to the major histocompatibility complex (MHC) on chromosome 6 (Robinson et al., 2013b). NK cell activity is dependent on interaction with

HLA class I molecules via killer-cell immunoglobulin-like receptors (KIRs); a specific family of cell-surface activating/inhibitory molecules mapping to the leukocyte receptor complex (LRC) on chromosome 19 (Middleton and Gonzelez, 2010). Extensive genetic variation exists within both the *HLA class I* and *KIR* genes due to unequal distribution of *HLA class I* allelic variants and *KIR* haplotypes within different populations. It is therefore necessary to identify the presence of genetic variants associated with disease susceptibility and outcome, particularly in the context of HIV-1 infection, in different South Africa populations where the HIV-1 prevalence is high.

HLA class I genotyping has been done in SAB and SAC populations (Paximadis et al., 2012) but limited information on *KIR* genotypes is available in these populations, with one *KIR* genotyping study published for the SAB population (Hong et al., 2011). Furthermore, in terms of *HLA class I* and *KIR* genotypes, the SAM and SAI populations are relatively understudied in comparison to SAB and SAC populations. The purpose of the current study was to characterise several immunogenetic factors in healthy individuals of SAI (n=50) and SAM (n=50) ethnicity and to include data generated by others for the SAC (n=94) and SAB (n=159) populations. Real-time allele-specific PCR (AS-PCR) assays were used to determine *CCR5/CCR5Δ32*, *KIR* and *HLA-Bw4*^{80I}*/Bw4*^{80T}*/C1/C2* genotypes, while relative quantification real-time AS-PCR assays were used to determine copy number variation of *KIR3DL1/S1* and *KIR2DS4f/v*. *HLA-ABC* genotyping was done using commercial sequence-based typing (SBT) methods and HLA-B signal peptide p2-M/T variants were determined from these sequences. *HLA-C 3'UTR* +263-ins/del variants were determined by direct DNA sequencing. For detailed methodologies, see **Chapter 2**.

3.2. Results

3.2.1. *KIR* genotyping

3.2.1.1. Analysis of KIR frequencies

The presence/absence of each *KIR* gene was determined by real-time AS-PCR assay for SAM and SAI populations. Results for SAC and SAB populations, generated using the same assay, were provided by Ms. Nikki Gentle. A comparison of *KIR* gene frequencies between the four populations as well as significant differences are shown in **Figure 3.1**. The

frequencies of *KIR2DL1*, *-2DL4*, *-3DL2*, *-3DL3*, *-DS5* and *-3DP1* were similar between the four populations. *KIR2DL2* was found at the lowest frequency in the SAC population (58.5%). *KIR2DL3*, *-3DL1*, *-2DS4* and *-2DP1* frequencies were lowest in the SAI population (64.0%, 88.0%, 88.0% and 94.0%, respectively) and the difference was significant for *KIR2DL3* in comparison to the other populations. *KIR2DL5*, *-2DS1*, *-2DS2* and *-2DS3* were found at the highest frequencies in the SAI population (84.0%, 66.0%, 78.0% and 56.0%) respectively) and the difference was significant for *KIR2DS1* and *-2DS3* in comparison to the other populations. Low frequencies of *KIR2DS1* and *-3DS1* (12.6% and 6.9%, respectively) and the difference was significant for *KIR3DS1* in comparison to the other populations. The framework genes (*KIR2DL4, -3DL2, -3DL3* and *-3DP1*) were present in all individuals in all populations except one SAI individual shown to lack *KIR2DL4*.

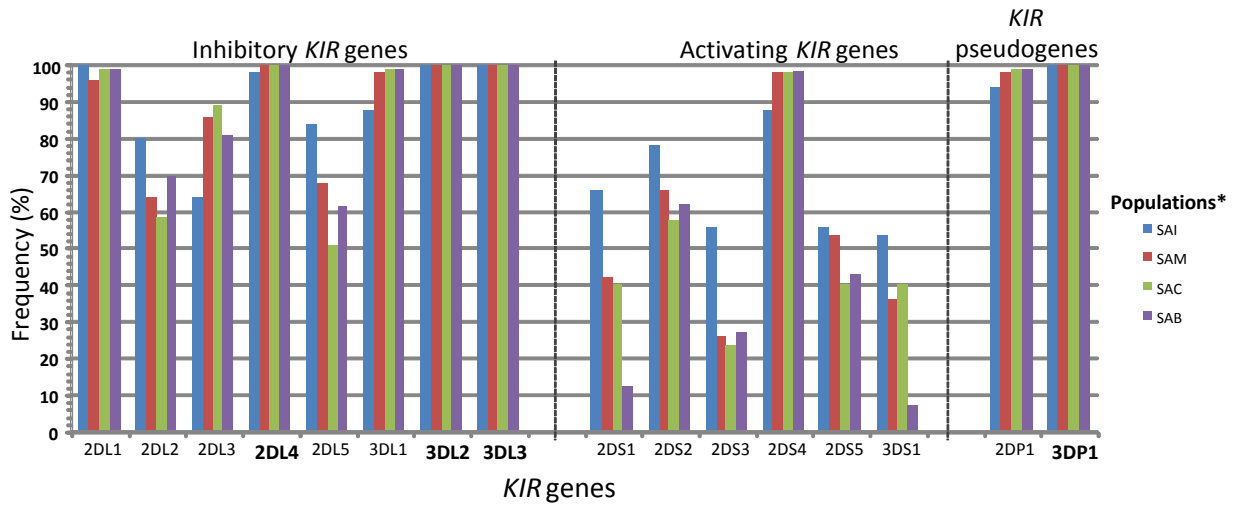

Figure 3.1. Frequency (%) of *KIR* (killer-cell immunoglobulin-like receptor) genes determined for SAI, South African Indian (n=50); SAM, South African Mixed ancestry (n=50); SAC, South African Caucasian (n=94) and SAB, South African Black (n=159) populations; Bold *KIRs* are framework genes; *KIR2DL4* is listed as an inhibitory gene although this may not be true at a functional level.

***Inter-population gene frequency comparisons showing significance ($p<0.05$):**

KIR2DL2	SAI vs. SAC	p=0.0101	***KIR3DL1***	SAC vs. SAI	p=0.0211	***KIR2DS3***	SAI vs. SAM	p=0.0042
KIR2DL3	SAM vs. SAI	p=0.0198		SAB vs. SAI	p=0.0008		SAI vs. SAC	p=0.0002
	SAC vs. SAI	p=0.0007	***KIR2DS1***	SAI vs. SAM	p=0.0268		SAI vs. SAB	p=0.0003
	SAB vs. SAI	p=0.0196		SAI vs. SAC	p=0.0049	***KIR2DS4***	SAC vs. SAI	p=0.0211
KIR2DL5	SAI vs. SAC	p=0.0001		SAI vs. SAB	p<0.0001		SAB vs. SAI	p=0.0027
	SAI vs. SAB	p=0.0033		SAM vs. SAB	p<0.0001	***KIR3DS1***	SAI vs. SAB	p<0.0001
				SAC vs. SAB	p<0.0001		SAM vs. SAB	p<0.0001
			KIR2DS2	SAI vs. SAC	p=0.0173		SAC vs. SAB	p<0.0001
				SAI vs. SAB	p=0.0418	***KIR2DP1***	SAB vs. SAI	p=0.0433

3.2.1.2. Analysis of KIR genotype frequencies

KIR genotype designations were assigned for individuals genotyped for the presence/absence of the 16 *KIR*s in all four populations (**Table 3.1**) according to the classification system used by the allele frequency database (Gonzalez-Galarza et al., 2011). Overall, 55 different *KIR* genotypes were identified, including 5 undocumented *KIR* genotypes (three found in the SAI population, one in the SAM population and one in the SAC population). In SAI, SAM, SAC and SAB populations 21, 18, 24 and 32 different *KIR* genotypes, respectively, were observed. *KIR* genotypes unique to one population or shared between two, three or all of the populations are depicted in **Figure 3.2**. Predominant *KIR* genotypes detected at frequencies $\geq$10% were detected as follows: *Bx6* (12.0%), *Bx73* (12.0%) and *Bx71* (10.0%) in SAI individuals; *AA1* (22.0%), *Bx2* (10.0%) and *Bx21* (16%) in SAM individuals; *AA1* (26.6%), *Bx4* (16.0%), *Bx2* (11.7%) and *Bx3* (10.6%) in SAC individuals and *AA1* (25.8%), and *Bx21* (11.3%) in SAB individuals. Statistical differences were evaluated in populations where the same *KIR* genotype was present and are listed in **Table 3.2**.

Table 3.1. Frequency (%) of *KIR* genotypes determined for SAI (n=50), SAM (n=50), SAC (n=94) and SAB (n=159) populations

***KIR* geno-types**	**Inhibitory *KIR genes***								**Activating *KIR geness***						***KIR* pseudogenes**		**Populations***			
	2DL1	2DL2	2DL3	**2DL4**	2DL5	3DL1	**3DL2**	**3DL3**	2DS1	2DS2	2DS3	2DS4	2DS5	3DS1	2DP1	**3DP1**	SAI	SAM	SAC	SAB
AA1	■		■	■		■	■	■				■			■	■	8.0	22.0	26.6	25.8
Bx2	■		■	■	■	■	■	■	■			■	■	■	■	■	6.0	10.	11.7	0.6
Bx3	■	■	■	■	■	■	■	■	■	■		■	■	■	■	■	4.0	2.0	10.6	0.6
Bx4	■	■	■	■		■	■	■		■		■			■	■	6.0	8.0	16.0	9.4
Bx5	■	■	■	■	■	■	■	■		■	■	■			■	■	6.0	8.0	2.1	9.4
Bx6	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	12.0	6.0	6.4	3.1
Bx7	■	■	■	■	■	■	■	■	■	■	■	■		■	■	■	2.0	4.0	2.1	0.6
Bx9	■	■	■	■	■	■	■	■	■	■		■	■		■	■	6.0	4.0	0.0	3.8
Bx10	■		■	■		■	■	■		■		■			■	■	0.0	0.0	1.1	0.0
Bx11	■	■	■	■	■	■	■	■	■	■	■	■			■	■	0.0	0.0	0.0	0.6
Bx13	■	■	■	■	■	■	■	■		■	■	■		■	■	■	0.0	0.0	2.1	0.0
Bx15	■		■	■		■	■	■	■			■			■	■	0.0	0.0	1.1	0.0
Bx19	■	■	■	■		■	■	■				■			■	■	0.0	0.0	0.0	0.6
Bx20	■	■	■	■	■	■	■	■				■	■		■	■	0.0	0.0	0.0	5.0
Bx21	■	■	■	■	■	■	■	■		■		■	■		■	■	0.0	16.0	1.1	11.3
Bx24	■	■	■	■	■	■	■	■		■		■	■	■	■	■	0.0	0.0	1.1	0.0
Bx25	■	■	■	■	■	■	■	■		■	■	■	■		■	■	0.0	0.0	0.0	0.6
Bx30	■		■	■	■	■	■	■			■	■			■	■	0.0	0.0	0.0	0.6
Bx31	■	■	■	■	■	■	■	■		■		■			■	■	0.0	0.0	0.0	1.3
Bx32	■		■	■	■	■	■	■				■	■		■	■	0.0	0.0	0.0	1.9
Bx44	■	■	■	■		■	■	■		■		■	■		■	■	0.0	0.0	0.0	0.6
Bx48	■		■	■	■	■	■	■		■		■	■		■	■	0.0	0.0	1.1	0.6
Bx51	■	■	■	■	■	■	■	■			■	■			■	■	0.0	0.0	0.0	1.9
Bx58	■	■	■	■	■	■	■	■	■		■	■	■	■	■	■	0.0	0.0	1.1	0.0
Bx62	■	■	■	■		■	■	■		■	■	■			■	■	0.0	0.0	0.0	0.6
Bx64	■	■	■	■	■	■	■	■	■		■	■		■	■	■	2.0	0.0	1.1	0.0
Bx68	■	■	■	■	■		■	■	■	■			■	■	■	■	2.0	0.0	0.0	0.0
Bx69	■		■	■	■		■	■	■				■	■	■	■	4.0	0.0	0.0	0.0
Bx70	■	■	■	■	■		■	■	■	■	■		■	■	■	■	0.0	0.0	2.1	0.0
Bx71	■	■		■	■	■	■	■		■	■	■			■	■	10.0	0.0	3.2	4.4
Bx73	■	■		■	■	■	■	■	■	■	■	■	■	■	■	■	12.0	4.0	1.1	0.6
Bx76		■	■	■	■	■	■	■	■	■		■	■	■		■	0.0	2.0	0.0	0.0
Bx78		■		■	■	■	■	■	■	■		■	■	■	■	■	0.0	2.0	0.0	0.0
Bx81	■	■		■	■		■	■	■	■	■		■	■	■	■	4.0	0.0	0.0	0.6
Bx88	■		■	■	■		■	■	■	■			■	■	■	■	0.0	2.0	0.0	0.0
Bx89	■	■		■		■	■	■		■		■			■	■	0.0	0.0	3.2	0.6
Bx91	■	■		■	■	■	■	■	■	■	■	■	■		■	■	0.0	2.0	0.0	0.0
Bx92	■	■		■	■	■	■	■	■	■		■	■		■	■	0.0	2.0	1.1	0.6
Bx112	■	■		■	■	■	■	■		■	■	■	■		■	■	0.0	0.0	0.0	3.1
Bx113	■	■		■	■	■	■	■	■	■	■	■			■	■	2.0	0.0	0.0	0.0
Bx169		■	■	■	■	■	■	■	■	■			■	■	■	■	0.0	0.0	0.0	0.6
Bx172	■	■		■		■	■	■		■		■				■	2.0	0.0	1.1	0.6
Bx175	■	■		■	■	■	■	■				■	■		■	■	0.0	0.0	0.0	0.6
Bx188	■	■	■	■		■	■	■		■		■		■	■	■	0.0	2.0	0.0	0.0
Bx228	■	■		■	■	■	■	■		■		■	■		■	■	0.0	2.0	1.1	7.5
Bx270	■	■	■	■	■	■	■	■			■	■	■		■	■	0.0	0.0	0.0	0.6
Bx337	■	■	■	■	■	■	■	■			■	■		■	■	■	2.0	0.0	0.0	0.0
Bx382	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	0.0	0.0	1.1	0.0
Bx393	■		■	■	■	■	■	■	■	■		■	■		■	■	0.0	0.0	0.0	0.6
Bx429	■	■		■	■	■	■	■	■	■		■	■	■		■	2.0	0.0	0.0	0.0
BxN1	■	■			■		■	■	■	■	■	■	■			■	2.0	0.0	0.0	0.0
BxN2	■	■	■	■	■	■	■	■	■	■	■	■			■	■	4.0	0.0	0.0	0.0
BxN3	■			■	■	■	■	■	■	■	■	■		■	■	■	2.0	0.0	0.0	0.0
BxN4	■		■	■	■	■	■	■	■		■	■		■	■	■	0.0	2.0	0.0	0.0
BxN5		■	■	■	■	■	■	■		■	■	■	■	■	■	■	0.0	0.0	1.1	0.0

KIR, killer-cell immunoglobulin-like receptor, SAI, South African Indian, SAM, South African Mixed ancestry, SAC, South African Caucasian; SAB, South African Black. Grey tiles indicate the presence of a *KIR* gene; Bold *KIRs* are framework genes; *KIR2DL4* is listed as an inhibitory gene although this may not be true at a functional level.

***Inter-population KIR genotype comparisons showing significance ($p<0.05$):**

Bx2	SAI vs. SAB	p=0.0433	**Bx6**	SAI vs. SAB	p=0.0244
	SAM vs. SAB	p=0.0033	**Bx21**	SAM vs. SAC	p=0.0009
	SAC vs. SAB	p=0.0001		SAB vs. SAC	p=0.0022
Bx3	SAC vs. SAB	p=0.0003	**Bx73**	SAI vs. SAC	p=0.0073
Bx5	SAB vs. SAC	p=0.0348		SAI vs. SAB	p=0.0008
			Bx228	SAB vs. SAC	p= 0.0351

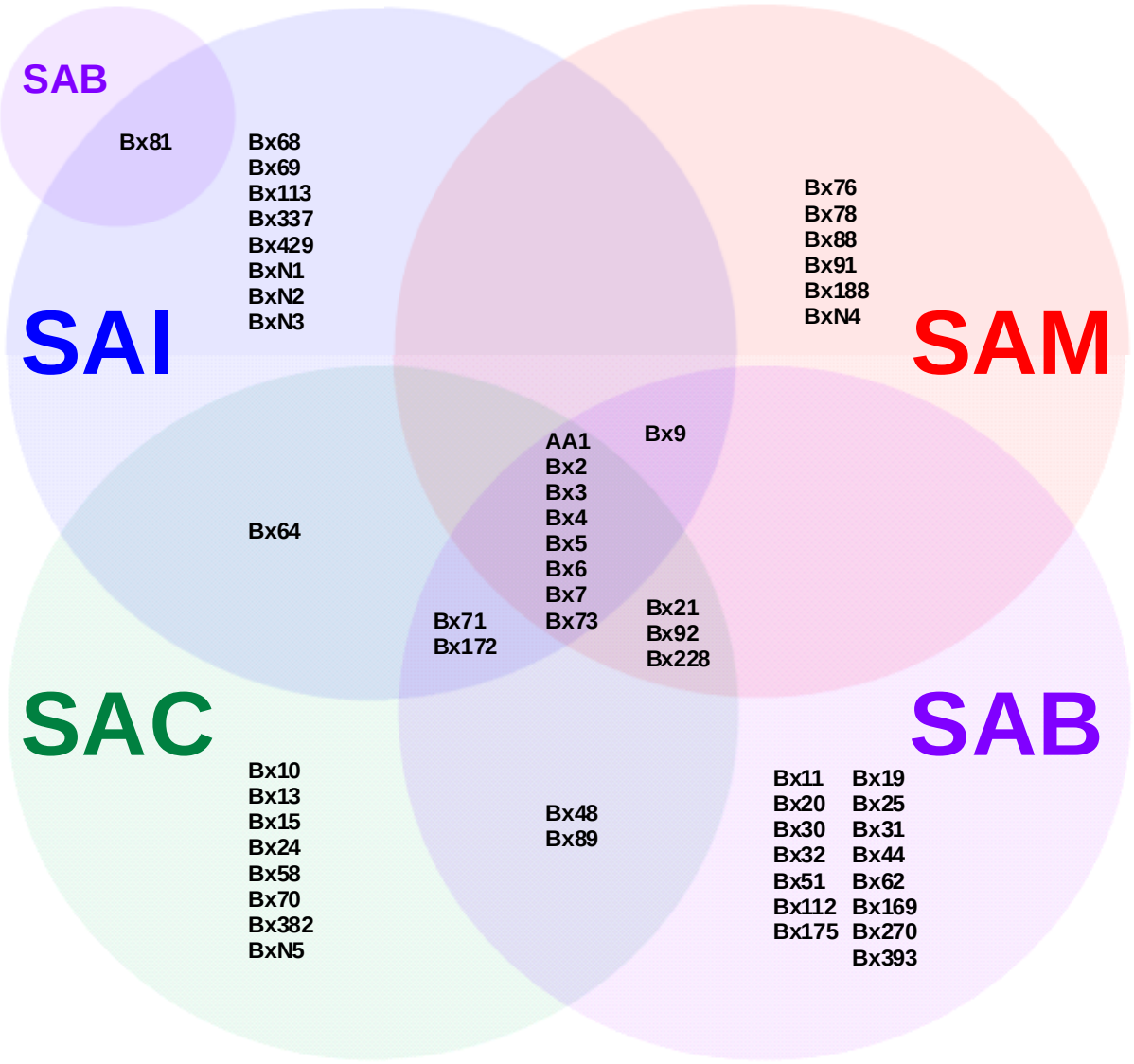

Figure 3.2. Distribution of shared and unique *KIR, killer-cell immunoglobulin-like receptor* genotypes determined for SAI, South African Indian (n=50), SAM, South African Mixed ancestry (n=50), SAC, South African Caucasian (n=94) and SAB, South African Black (n=159) populations.

3.2.2. *HLA-Bw4^{80I}/Bw4^{80T}/C1/C2* ligand genotyping

HLA-Bw4^{80I}/Bw4^{80T}/C1/C2 genotypes were initially determined by real-time AS-PCR assay for SAI and SAM populations and confirmed to be correctly assigned by analysis of *HLA class I* genotype data that was generated later. *HLA-Bw4^{80I}/Bw4^{80T}/C1/C2* genotypes for the SAC and SAB populations were determined from published *HLA-ABC* data (Paximadis et al., 2012) and additional *HLA-ABC* data was provided by Ms. Nikki Gentle. Identification of genetic variability at codon-80 of *HLA-A, -B* and *-C* loci was used to define combinations of *HLA* ligand genotypes present in all four populations. *HLA* ligand frequencies and significant differences are shown in **Figure 3.2**.

For *HLA-A*, *Bw4*^{80I} homozygotes were absent in the SAC population and found at low frequencies in the SAI, SAM and SAB populations (4.0%, 4.0% and 1.3%, respectively). The highest frequency of *non-Bw4*^{80T} homozygotes was found in SAB population (77.4%) and the frequency of *Bw4*^{80I}*/non-Bw4*^{80T} heterozygotes were highest in the SAI populations (38.0%).

At the *HLA-B* locus, the frequency of *Bw6*^{80N} homozygotes and *Bw4*^{80I}*/Bw6*^{80N} heterozygotes was highest the SAB population (38.4% and 40.9%, respectively. Low frequencies of *Bw6*^{80N}*/Bw4*^{80TL} heterozygotes were observed in SAI, SAM and SAC populations (10.0%, 4.0% and 6.4%, respectively), while this genotype was absent in the SAB population. The frequency of *Bw6*^{80N}*/Bw4*^{80TA} heterozygotes was highest in the SAC population (20.2%). Low frequencies of *Bw4*^{80I}*/Bw4*^{80TA} heterozygotes, *Bw4*^{80I} homozygotes and *Bw4*^{80TA}*/Bw4*^{80TL} heterozygotes were detected in all populations, but highest in the SAI population (8.0%, 6.0% and 8% respectively). The frequency of *Bw4*^{80TA} homozygotes was low in all populations (2.0%, 4.0%, 2.1% and 0.6%, respectively). *Bw4*^{80TA}*/Bw4*^{80TL} heterozygotes were absent in the SAB population, *Bw4*^{80I}*/Bw4*^{80TL} heterozygotes were absent in both the SAI and SAB populations and no *Bw4*^{80TL} homozygotes were deteced in any of the populations. Notably, *HLA-B* alleles that encode the *C1* epitope are prevalent in Asian populations (Abi-Rached et al., 2010) and could potentially be present in SAI and SAM populations, however, none were detected in this study.

For *HLA-C*, the frequency of *C1* homozygotes was highest in the SAC population (42.6%), while the frequency of *C2* homozygotes was highest in the SAB population (31.4%). All populations had similar frequencies of *C1/C2* heterozygotes ranging from 46.0% to 48.9%.

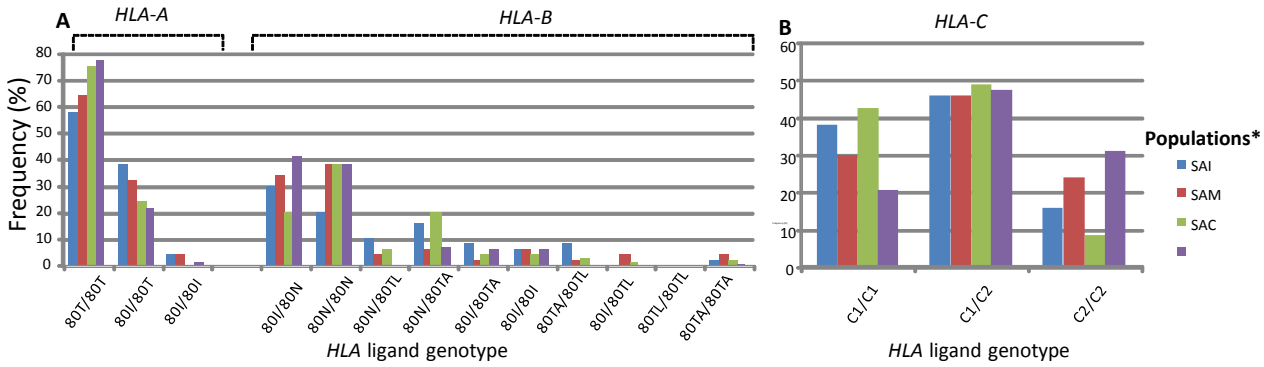

Figure 3.3. Frequency (%) of *HLA-A, -B* and *-C* ligand profiles determined for SAM, South African Mixed ancestry (n=50); SAC, South African Caucasian (n=94); SAI, South African Indian (n=50) and SAB, South African Black (n=159) populations; *HLA*, human leukocyte antigen; *HLA-A* locus 80I = Bw4, 80T = non-Bw4; At the *HLA-B* locus 80I, 80TA and 80TL = Bw4, 80N = Bw6.

***Inter-population *HLA-ABC* ligand frequency comparisons showing significance (p<0.05):**

HLA-A 80T/80T	SAC vs. SAI	p=0.0371	***HLA-B 80N/80TA***	SAC vs. SAM	p=0.0281
	SAB vs. SAI	p=0.0105		SAC vs. SAB	p=0.0048
HLA-A 80I/80T	SAI vs. SAB	p=0.0250	***HLA-C C1/C1***	SAI vs. SAB	p=0.0233
HLA-B 80I/80N	SAB vs. SAC	p=0.0009		SAC vs. SAB	p=0.0003
HLA-B 80N/80N	SAC vs. SAI	p=0.0260	***HLA-C C2/C2***	SAB vs. SAI	p=0.0454
	SAB vs. SAI	p=0.0171		SAM vs. SAC	p=0.0205
				SAB vs. SAC	p<0.0001

3.2.3. Analysis of known *KIR-HLA* ligand pairs

Since *KIRs* and *HLAs* are encoded by different chromosomes (19 and 6, respectively), *KIRs* and cognate *HLA* ligands are not necessarily inherited together. The frequencies of known KIR-HLA ligand pairs were calculated for the SAI, SAM, SAC and SAB populations using the available *KIR* and *HLA-Bw4*^{80I}*/Bw4*^{80T}*/C1/C2* ligand genotypes as well as specific *HLA class I* alleles using *HLA class I* high-resolution typings. *KIR-HLA* ligand pair frequencies and significant differences are shown in **Table 3.3.**

The frequency of *KIR3DL1+HLA-B Bw4*^{80T}, *KIR2DL2+HLA-C1* and *KIR2DS1+HLA-C2* was highest in the SAI population (40.0%, 66.0% and 48.0%, respectively) with a significant difference in the frequency of *KIR2DS1+HLA-C2* in comparison to the other three populations. The SAM population showed the highest frequency of *KIR3DL1+HLA-A/B Bw4*^{80I} (68.0%), while in the SAC population the frequency of *KIR2DL3+HLA-C1* was highest (80.9%) and significantly different in

comparison to the other three populations. *KIR3DL1+HLA B Bw4*^{80I} and *KIR2DL2+HLA-C2* were highest in frequency in the SAB population (52.8% and 78.6%, respectively).

The real-time AS-PCR *HLA* ligand assay does not determine the presence of specific *HLA class I* alleles, however, since high-resolution *HLA class I* genotypes were also determined by SBT, it was possible to analyse *KIR* and *HLA* allele-specific ligand pairs for the four populations (**Table 3.2**). *KIR2DS4+HLA-A*11* and *KIR3DL2+HLA-A*11* was found at the highest frequency in the SAI population (28.0% and 32.0%, respectively) and significantly different in comparison to the other three populations. *KIR3DL2+HLA-A*03* was found at the highest frequency in the SAC population (24.5%) and significantly different in comparison to the other three populations.

KIR2DS4 is also known to bind subsets of HLA-C1 and -C2 allotypes with varying affinities (Graef et al., 2009) and the corresponding *KIR-HLA* ligand genotypes were analysed in the four populations. *KIR2DS4+HLA-C*01:02, KIR2DS4+HLA-C*03:02* and *KIR2DS4+HLA-C*14:02* were found at the highest frequencies in the SAI population (8.0%, 10.0% and 10.0%, respectively). The SAC population showed the highest frequencies of *KIR2DS4+HLA-C*02:02* and *KIR2DS4+HLA-C*05:01* (11.7% and 10.6%, respectively), while *KIR2DS4+HLA-C*04:01* and *KIR2DS4+HLA-C*16:01* was found at the highest frequency in the SAB population (24.5% and 13.8%, respectively).

Table 3.2. Frequency (%) of known *KIR-HLA* ligand pairs determined for SAI (n=50); SAM (n=50); SAC (n=94) and SAB (n=159) populations

***KIR-HLA* ligand pairs**	**Populations***			
	SAI	**SAM**	**SAC**	**SAB**
KIR3DL1+HLA-A Bw4^{80I}	36.0	36.0	14.9	22.6
KIR3DL1+HLA-B Bw4^{80I}	38.0	46.0	29.8	52.8
KIR3DL1+HLA-A/B Bw4^{80I}	58.0	68.0	44.7	65.4
KIR3DL1+HLA-B Bw4^{80T}	40.0	20.0	37.2	13.8
KIR2DL1+HLA-C2	62.0	68.0	56.4	78.6
KIR2DL2+HLA-C1	66.0	50.0	51.1	49.1
KIR2DL3+HLA-C1	54.0	64.0	80.9	56.6
KIR2DS1+HLA-C2	48.0	24.0	21.3	9.4
*KIR2DS4+HLA-A*11*	28.0	8.0	11.7	1.9
*KIR3DL2+HLA-A*03*	8.0	10.0	24.5	10.7
*KIR3DL2+HLA-A*11*	32.0	10.0	11.7	1.9
*KIR2DS4+HLA-C*01:02*	8.0	4.0	4.3	0.0
*KIR2DS4+HLA-C*02:02*	0.0	8.0	11.7	1.9
*KIR2DS4+HLA-C*03:02*	10.0	2.0	0.0	3.1
*KIR2DS4+HLA-C*04:01*	10.0	18.0	17.0	24.5
*KIR2DS4+HLA-C*05:01*	0.0	4.0	10.6	0.6
*KIR2DS4+HLA-C*14:02*	10.0	2.0	2.1	1.3
*KIR2DS4+HLA-C*16:01*	2.0	8.0	12.8	13.8

KIR, killer-cell immunoglobulin-like receptor, *HLA*, human leukocyte antigen; SAI, South African Indian; SAM, South African Mixed ancestry; SAC, South African Caucasian; SAB, South African Black.

***Inter-population *KIR-HLA* ligand frequency comparisons showing significance (p<0.05):**

KIR3DL1+HLA-A Bw4^{80I}	SAI vs. SAC	p=0.0058	***KIR2DS4+HLA-A*11***	SAI vs. SAM	p=0.0174
	SAM vs. SAC	p=0.0058		SAI vs. SAC	p=0.0202
KIR3DL1+HLA-B Bw4^{80I}	SAB vs. SAC	p=0.0004		SAI vs. SAB	p<0.0001
KIR3DL1+HLA-A/B Bw4^{80I}	SAM vs. SAC	p= 0.0088		SAC vs. SAB	p=0.0027
	SAB vs. SAC	p=0.0016	***KIR3DL2+HLA-A*03***	SAC vs. SAI	p=0.0234
KIR3DL1+HLA-B Bw4^{80T}	SAI vs. SAM	p=0.0486		SAC vs. SAM	p=0.0462
	SAI vs. SAB	p=0.0002		SAC vs. SAB	p=0.0068
	SAC vs. SAM	p=0.0387	***KIR3DL2+HLA-A*11***	SAI vs. SAM	p=0.0128
	SAC vs. SAB	p<0.0001		SAI vs. SAC	p=0.0062
KIR2DL1+HLA-C2	SAB vs. SAI	p=0.0250		SAI vs. SAB	p<0.0001
	SAB vs. SAC	p=0.0003		SAM vs. SAB	p=0.0204
KIR2DL3+HLA-C1	SAC vs. SAI	p=0.0010		SAC vs. SAB	p=0.0027
	SAC vs. SAM	p=0.0421	***KIR2DS4+HLA-C*02:02***	SAC vs. SAB	p=0.0027
	SAC vs. SAB	p<0.0001	***KIR2DS4+HLA-C*04:01***	SAB vs. SAI	p=0.0291
KIR2DS1+HLA-C2	SAI vs. SAM	p=0.0213	***KIR2DS4+HLA-C*05:01***	SAC vs. SAB	p=0.0003
	SAI vs. SAC	p=0.0013	***KIR2DS4+HLA-C*14:02***	SAI vs. SAC	p=0.0494
	SAI vs. SAB	p<0.0001		SAI vs. SAB	p=0.0094
	SAM vs. SAB	p=0.0137	***KIR2DS4+HLA-C*16:01***	SAC vs. SAI	p=0.0343
	SAC vs. SAB	p=0.0134		SAB vs. SAI	p=0.0185

3.2.4. Copy number variation of *KIR3DL1/S1* and *KIR2DS4f/v*

A probe hydrolysis-based relative quantification real-time AS-PCR assay was used to estimate copy numbers of *KIR2DS4f/v* and *KIR3DL1/S1* relative to the *BGB* gene in the four populations. The frequency of copy number variants and significant differences detected within the four populations is shown in **Figure 3.4**.

3.2.4.1. Copy number variation of KIR3DL1/S1

The absence of *KIR3DL1/S1* was only found in the SAI population (2.0%). The frequency of *KIR3DL1/S1* copy variants 0/1, 1/0, and 1/1 was highest in SAI population (8.0%, 14.0% and 48%, respectively), while the frequency of *KIR3DL1/S1* copy variants 0/2, 1/2 and 2/1 was highest in the SAC population (2.1%, 4.3% and 3.2%, respectively). A low frequency of *KIR3DL1/S1* copy variant 1/1 was found in the SAB population (6.3%) and the difference was significant in comparison to the other three populations. The *KIR3DL1/S1* copy variant 2/0 was highest in frequency in the SAB population (83.6%) and the difference was significant in comparison to the other three populations. The *KIR3DL1/S1* copy variants 0/1 were absent in the SAM and SAC populations, while *KIR3DL1/S1* copy variant 1/2 was not detected in the SAM and SAB populations. The *KIR3DL1/S1* copy variants 0/2 and 2/1 were not present in the SAB population. The detection of *KIR3DL1/S1* copy variants 0/0, 0/1 and 1/0 is suggestive of deletions at the *KIR3DL1/S1* locus. Similarly, the detection of *KIR3DL1/S1* copy variants 1/2 and 2/1 is suggestive of a duplication at the *KIR3DL1/S1* locus.

3.2.4.2. Copy number variation KIR2DS4f/v

For *KIR2DS4f/v* copy number variation, the absence of *KIR2DS4* was detected in all populations, but the highest frequency was found in the SAI population (12.0%) and the difference was significant in comparison to the other three populations. The *KIR2DS4f/v* copy variant 0/1 was found at the highest frequency in the SAB population (61.7%) and the difference was significant in comparison to the other three populations. The *KIR2DS4f/v* copy variant 2/0 was absent in the SAC and SAB populations and detected at the highest frequency in the SAI population (22.0%). The *KIR2DS4f/v* copy variants 1/0 and 1/1 was present at the highest frequency in the SAB population (50.3% and 34.6%, respectively) and a significant difference was observed in frequency of *KIR2DS4f/v* copy variants 1/0 in comparison to the other three populations. The *KIR2DS4f/v* copy variant 0/2 was absent in the SAC and SAB populations and present at the highest frequency in the SAM population (12.0%).

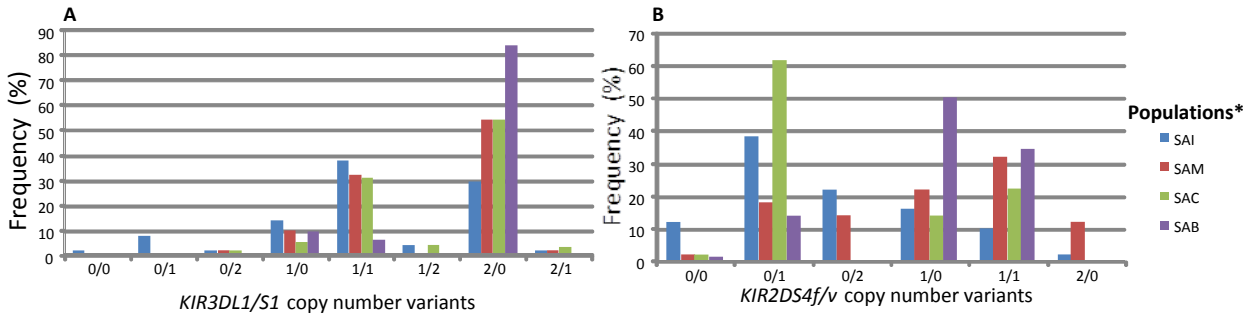

Figure 3.4. Frequency (%) of copy number variants for (A) *KIR3DL1/S1* and (B) *KIR2DS4f/v* genotypes determined for SAI, South African Indian (n=50), SAM, South African Mixed ancestry (n=50); SAC, South African Caucasian (n=94) and SAB, South African Black (n=159) populations.

***Inter-population *KIR3DL1/S1* and *KIR2DS4f/v* copy number frequency comparisons showing significance (p<0.05)**:

***KIR3DL1/S*:**	*0/1*	SAI vs. SAC	p=0.0134	***KIR2DL4f/v*:**	*0/0*	SAI vs. SAC	p=0.0211
		SAI vs. SAB	p=0.0122			SAI vs. SAB	p=0.0027
	1/1	SAI vs. SAB	p<0.0001		*0/1*	SAI vs. SAM	p=0.0440
		SAM vs. SAB	p<0.0001			SAI vs. SAB	p=0.0004
		SAC vs. SAB	p<0.0001			SAC vs. SAI	p=0.0085
	1/2	SAC vs. SAB	p=0.0183			SAC vs. SAM	p<0.0001
	2/0	SAM vs. SAI	p=0.0253			SAC vs. SAB	p<0.0001
		SAC vs. SAI	p=0.0081		*0/2*	SAI vs. SAC	p<0.0001
		SAB vs. SAI	p<0.0001			SAI vs. SAB	p<0.0001
		SAB vs. SAM	p<0.0001			SAM vs. SAC	p=0.0005
		SAB vs. SAC	p<0.0001			SAM vs. SAB	p<0.0001
					1/0	SAB vs. SAI	p<0.0001
						SAB vs. SAM	p=0.0005
						SAB vs. SAC	p<0.0001
					1/1	SAM vs. SAI	p=0.0128
						SAB vs. SAI	p=0.0006
						SAB vs. SAC	p=0.0470
					2/0	SAM vs. SAC	p=0.0014
						SAM vs. SAB	p=0.0001

3.2.5. *CCR5/CCR5Δ32* genotyping

The presence/absence of the *CCR5Δ32* variant (*Mt*) was determined in SAI, SAM and SAC individuals by real-time AS-PCR Ct-shift assay, however, this was not done for the SAB population due to the known absence of the *CCR5Δ32* variant in South African African individuals (Picton et al., 2010). The frequencies of *CCR5* wild-type (*WT*) and *Mt* genotypes and significant differences detected within the populations are shown in **Figure 3.5**. Low frequencies of *WT/Mt* heterozygotes were detected in the SAI and SAM populations (4.0% and 6.0%, respectively), while the highest frequency was present in the SAC population (17.0%). A low frequency of *Mt* homozygotes was found in the SAC population (1.1%), while this genotype was absent in the SAI and SAM populations. The frequency of *WT/Mt* heterozygotes and *Mt* homozygotes was assumed to be 0.0% in the SAB population.

3.2.6. HLA-B signal peptide p2-M/T variants

The presence of a methionine (M) or threonine (T) residue at position-2 (p2) of the HLA-B signal peptide was determined by analysis of exon 1 coding sequences generated by commercial *HLA-B* genotyping or from *HLA class I* sequences available in the International Histocompatibility Working Group (IHWG) (www.ihwg.org). The frequencies of *HLA-B p2-M/T* genotypes and significant differences detected within the populations are shown in **Figure 3.5.** *HLA-B p2-M/M* homozygotes were absent in the SAI population, present at low frequencies in the other three populations and highest in the SAB population (8.8%). *HLA-B p2-M/T* heterozygotes were present at the highest frequency in the SAC population (47.9%). The highest frequency of *HLA-B p2-T/T* homozygotes was found in the SAI population (78.0%) and was significantly different in comparison to the other three populations.

3.2.7. *HLA-C 3'UTR* +263 Ins/Del variants

The insertion (Ins) or deletion (Del) of a single guanine (G) residue at position +263 in the *HLA-C 3'untranslated region (3'UTR)* was determined for the SAI and SAM populations by direct sequencing and genotypes were inferred from these results. Genotypic data for the SAC and SAB populations was provided by Ms. Nikki Gentle. The frequencies of *HLA-C 3'UTR +263 Ins/Del* genotypes and significant differences detected within the populations are shown in **Figure 3.5.** *HLA-C 3'UTR +263 Del/Del* homozygotes were the least common genotypes detected in all populations with the highest frequency found in the SAM population (20%). *HLA-C 3'UTR +263 Ins/Ins* homozygotes were found at the highest frequency in the SAC population (52.1.0%), while *HLA-C 3'UTR +263 Ins/Del* heterozygotes were the most common in the SAI population (60.0%).

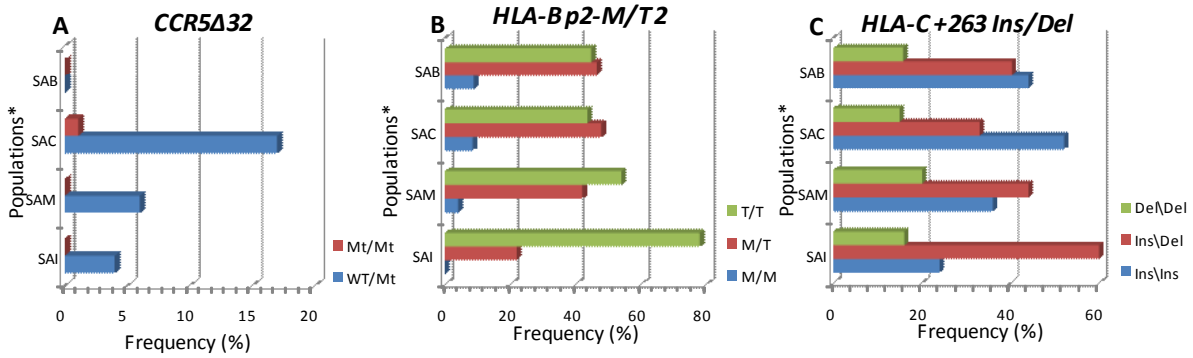

Figure 3.5. Frequency (%) of genotypes for (A) *CCR5Δ32 WT/Mt,* (B) *HLA-B p2-M/T* and (C) *HLA-C +263 Ins/Del* determined for SAI, South African Indian (n=50), SAM, South African Mixed ancestry (n=50); SAC, South African Caucasian (n=94) and SAB, South African Black (n=159) populations.

***Inter-population frequency comparisons for *CCR5Δ32*, *HLA-Bp2-M/T2* and *HLA-C +263 Ins/Del* showing significance (p<0.05):**

CCR5Δ32	*WT/Mt*	SAC vs. SAI	p=0.0323	***HLA-C +263 Ins/Del***	*Ins/Ins*	SAC vs. SAI	p=0.0014
		SAM vs. SAB	p=0.0131			SAB vs. SAI	p=0.0127
		SAC vs. SAB	p<0.0001		*Ins/Del*	SAI vs. SAC	p=0.0025
HLA-B p2-M/T	*M/M*	SAB vs. SAI	p=0.0248			SAI vs. SAB	p=0.0219
	M/T	SAC vs. SAI	p=0.0024				
		SAB vs.SAI	p=0.0027				
	T/T	SAI vs. SAM	p=0.0196				
		SAI vs. SAC	p<0.0001				
		SAI vs. SAB	p<0.0001				

3.2.8. *HLA class I A, -B* and *-C* allelic variation

HLA-ABC alleles were determined for SAI and SAM individuals by commercial SBT methods. *HLA-ABC* genotypes for the SAC and SAB populations were derived from published data (Paximadis et al., 2012) and additional typings were done by Ms Nikki Gentle.

3.2.8.1. Alleles at the HLA-A locus

The allele frequencies (f) of *HLA-A* allelic variants detected and significant differences found within the populations are shown in **Table 3.3**. Overall, a total of 42 different alleles were detected in the four populations with 19, 24, 25 and 31 different alleles identified in SAI, SAM, SAC and SAC populations, respectively. *HLA-A* alleles unique to one population or shared between two, three or all of the populations are depicted in **Figure 3.6**. The most common *HLA-A* allele detected in each population was *HLA-A*11:01* (f=0.17), *-A*02:01* (f=0.21), *-A*02:01* (f=0.255) and *-A*30:01* (f=0.11) in SAI, SAM, SAC and SAC populations, respectively. The frequency of individuals who were homozygous at the *HLA-A* locus was 16.0%, 8.0%, 12.8% and 9.4% in the SAI, SAM, SAC and SAB populations, respectively.

3.2.8.2. Alleles at the HLA-B locus

The allele frequencies of *HLA-B* allelic variants detected and significant differences found within the populations are shown in **Table 3.4.** Overall, a total of 61 different alleles were detected in the four populations with 30, 41, 38 and 29 different alleles identified in SAI, SAM, SAC and SAC populations, respectively. *HLA-B* alleles unique to one population or shared between two, three or all of the populations are depicted in **Figure 3.7**. The most common *HLA-B* allele detected in each population was *HLA-B*40:06* (f=0.14), *-B*07:02* (f=0.08), *-B*07:02* (f=0.154) and *-B*58:02* (f=0.094) in SAI, SAM, SAC and SAC populations, respectively. The frequency of individuals who were homozygous at the *HLA-B* locus was 4.0%, 6.0%, 5.3% and 5.0% in the SAI, SAM, SAC and SAB populations, respectively.

3.2.8.3. Alleles at the HLA-C locus

The allele frequencies of *HLA-C* allelic variants detected and significant differences found within the populations are shown in **Table 3.5.** Overall, a total of 38 different alleles were detected in the four populations with 20, 24, 26 and 27 different alleles identified in SAI, SAM, SAC and SAC populations, respectively. *HLA-C* alleles unique to one population or shared between two, three or all of the populations are depicted in **Figure 3.8**. The most common *HLA-C* allele detected in each population was *HLA-C*06:02* (f=0.14), *-C*06:02* (f=0.12), *-C*07:01* (f=0.176) and *-C*06:02* (f=0.157) identified in SAI, SAM, SAC and SAC populations, respectively. The frequency of individuals who were homozygous at the *HLA-C* locus was 6.0%, 8.0%, 6.4% and 9.4% in the SAI, SAM, SAC and SAB populations, respectively.

Table 3.3. Allele frequencies of *HLA-A* alleles determined for SAI (n=50), SAM (n=50), SAC (n=94) and SAB (n=159) populations.

	Populations*			
HLA-A	**SAI**	**SAM**	**SAC**	**SAB**
A*01:01	0.130	0.130	0.207	0.019
A*01:03				0.003
A*02:01	0.040	0.210	0.255	0.075
A*02:02			0.011	0.025
A*02:03		0.010		
A*02:05	0.010		0.005	0.072
A*02:06	0.030			
A*02:11	0.040	0.020		
A*02:14				0.003
A*02:17			0.005	
A*03:01	0.030	0.060	0.122	0.053
A*03:02	0.010			
A*11:01	0.170	0.060	0.053	0.006
A*11:03	0.020			
A*11:12			0.005	0.003
A*23:01	0.010	0.040	0.016	0.085
A*24:02	0.160	0.090	0.069	0.019
A*24:07	0.030	0.040	0.005	
A*25:01		0.020	0.005	
A*26:01	0.020	0.020	0.021	0.009
A*29:01				0.003
A*29:02		0.030	0.064	0.063
A*29:11				0.028
A*30:01		0.040	0.016	0.110
A*30:02	0.010	0.050	0.005	0.107
A*30:04		0.020	0.027	0.022
A*31:01	0.020	0.010	0.016	
A*32:01	0.030	0.010	0.027	0.016
A*33:01			0.011	0.003
A*33:03	0.150	0.010	0.005	0.022
A*34:01		0.010		
A*34:02		0.030	0.021	0.031
A*36:01				0.013
A*43:01		0.020		0.041
A*66:01		0.010		0.006
A*66:02				0.003
A*68:01	0.070	0.020	0.016	0.035
A*68:02	0.020	0.040	0.005	0.072
A*68:27				0.003
A*69:01			0.005	
A*74:01				0.041
A*80:01				0.006

HLA, human leukocyte antigen; SAI, South African Indian; SAM, South African Mixed ancestry; SAC, South African Caucasian; SAB, South African Black; Allele frequency defined as n/2 x number of individuals.

***Inter-population frequency comparisons for *HLA-A* alleles showing significance (p<0.05):**

HLA-A*01:01	SAI vs. SAB	p<0.0001	***HLA-A*23:01***	SAB vs. SAI	p=0.0054
	SAM vs. SAB	p<0.0001		SAB vs. SAC	p=0.0014
	SAC vs. SAB	p<0.0001	***HLA-A*24:02***	SAI vs. SAC	p=0.0223
HLA-A*02:01	SAM vs. SAI	p=0.0004		SAI vs. SAB	p<0.0001
	SAC vs. SAI	p<0.0001		SAM vs. SAB	p=0.0025
	SAM vs. SAB	p=0.0004		SAC vs. SAB	p=0.0064
	SAC vs. SAB	p<0.0001	***HLA-A*30:01***	SAB vs. SAM	p=0.0466
HLA-A*02:05	SAB vs. SAI	p=0.0233		SAB vs. SAC	p<0.0001
	SAB vs. SAC	p=0.0003	***HLA-A*30:02***	SAB vs. SAI	p=0.0008
HLA-A*03:01	SAC vs. SAI	p=0.0088		SAM vs. SAC	p=0.0204
	SAC vs. SAB	p=0.0097		SAB vs. SAC	p<0.0001
HLA-A*11:01	SAI vs. SAM	p=0.0249	***HLA-A*33:03***	SAI vs. SAM	p=0.0003
	SAI vs. SAC	p=0.0023		SAI vs. SAC	p<0.0001
	SAI vs. SAB	p<0.0001		SAI vs. SAB	p<0.0001
	SAM vs. SAB	p=0.0030	***HLA-A*68:02***	SAB vs. SAC	p=0.0003
	SAC vs. SAB	p=0.0013			

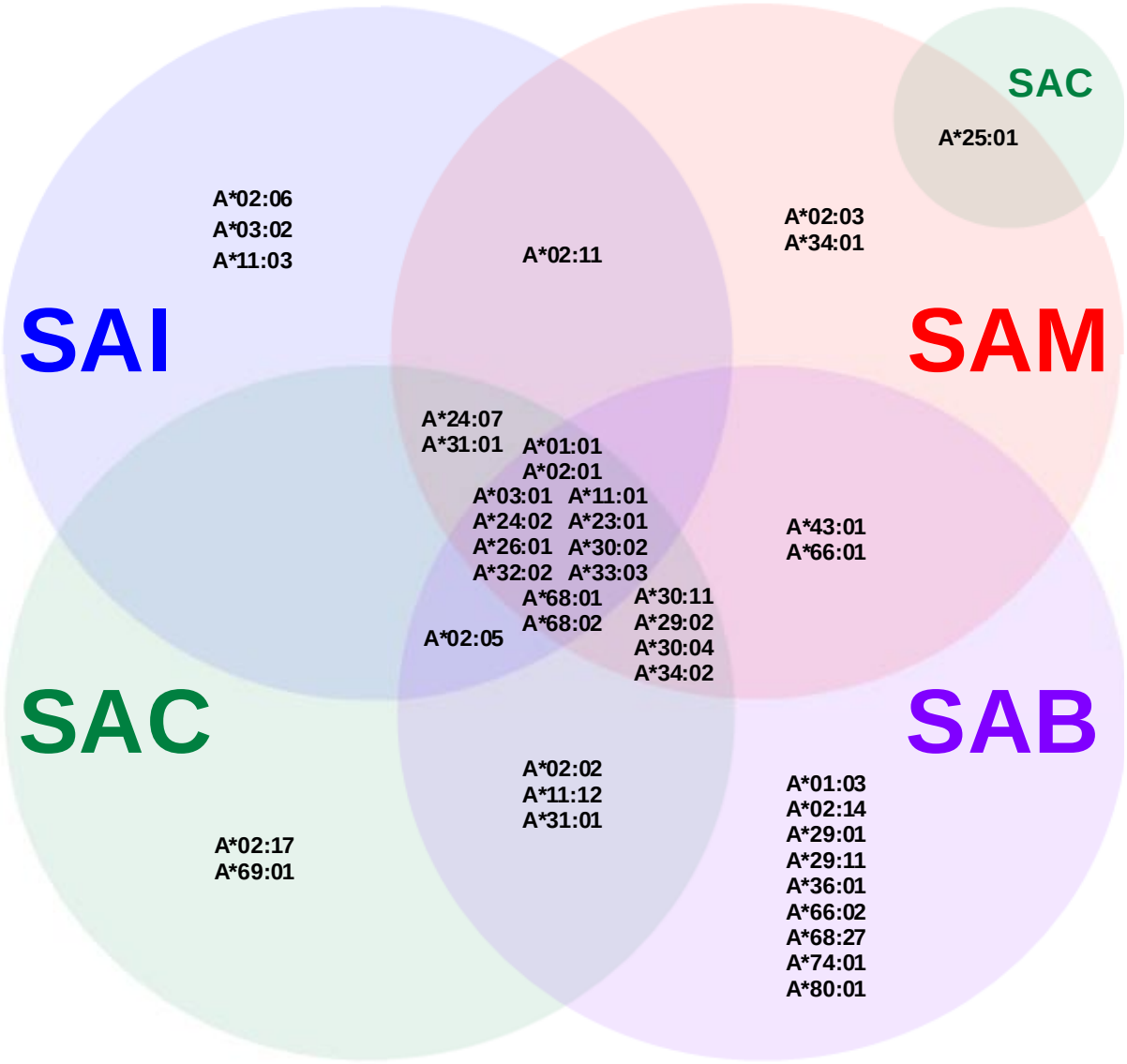

Figure 3.6. Distribution of shared and unique *HLA, human leukocyte antigen class I A* alleles determined for SAI, South African Indian (n=50), SAM, South African Mixed ancestry (n=50), SAC, South African Caucasian (n=94) and SAB, South African Black (n=159) populations.

Table 3.4. Allele frequency of *HLA-B* alleles determined for SAI (n=50), SAM (n=50), SAC (n=94) and SAB (n=159) populations**.**

	Populations*			
HLA-B	**SAI**	**SAM**	**SAC**	**SAB**
B*07:02	0.040	0.080	0.154	0.050
B*07:05	0.010			
B*07:06			0.005	0.016
B*08:01	0.020	0.020	0.133	0.057
B*13:01	0.040	0.010		
B*13:02			0.011	0.022
B*14:01	0.010	0.020	0.005	0.022
B*14:02		0.020	0.016	0.019
B*15:01		0.030	0.090	0.006
B*15:02	0.010	0.010	0.005	
B*15:03		0.020	0.005	0.085
B*15:08	0.020			
B*15:10	0.020	0.020	0.005	0.066
B*15:13		0.020	0.005	
B*15:16		0.010	0.005	0.016
B*15:17	0.010		0.011	
B*15:25	0.010			
B*18:01	0.010	0.070	0.053	0.028
B*27:02			0.005	
B*27:05		0.030	0.043	
B*27:07	0.010			
B*35:01	0.050	0.010	0.064	0.016
B*35:02		0.020		
B*35:03	0.050	0.030		
B*35:05		0.020	0.005	
B*35:08			0.011	
B*37:01	0.080	0.010	0.011	
B*38:01	0.010	0.010	0.011	
B*38:02		0.010		
B*39:06		0.010		
B*39:10		0.020		0.022
B*40:01	0.030	0.040	0.021	0.003
B*40:02		0.020	0.011	
B*40:06	0.140	0.040		
B*41:01	0.010	0.020	0.005	0.013
B*42:01		0.050		0.079
B*42:02				0.016
B*42:06				0.003
B*44:02	0.010		0.053	
B*44:03	0.130	0.040	0.074	0.053
B*44:04			0.005	
B*44:05		0.020	0.011	
B*44:27			0.005	
B*44:29		0.010		
B*45:01		0.020	0.005	0.079
B*45:07				0.006
B*47:01		0.010		
B*48:01	0.020			
B*49:01		0.030	0.005	0.006
B*50:01	0.010		0.005	
B*51:01	0.040	0.040	0.048	0.009
B*51:19	0.010			
B*52:01	0.060	0.020	0.011	
B*53:01		0.010		0.050
B*55:01	0.020		0.016	
B*57:01	0.060	0.040	0.043	
B*57:02		0.010		0.013
B*57:03				0.025
B*58:01	0.050	0.020	0.021	0.088
B*58:02	0.010	0.050	0.005	0.094
B*81:01		0.010		0.038

HLA, *human leukocyte antigen*; SAI, South African Indian; SAM, South African Mixed ancestry; SAC, South African Caucasian; SAB, South African Black; Allele frequency defined as n/2 x number of individuals.

*Inter-population frequency comparisons for *HLA-B* alleles showing significance (p<0.05):

HLA-B*07:02	SAC vs. SAI	p=0.0032	***HLA-B*40:01***	SAI vs. SAB	p=0.0441
	SAC vs. SAB	p=0.0001		SAM vs. SAB	p=0.0127
HLA-B*08:01	SAC vs. SAI	p=0.0011	***HLA-B*40:06***	SAI vs. SAM	p=0.0238
	SAC vs. SAM	p=0.0011	***HLA-B*44:03***	SAI vs. SAM	p=0.0398
	SAC vs. SAB	p=0.0045		SAI vs. SAB	p=0.0142
HLA-B*15:01	SAC vs. SAB	p<0.0001	***HLA-B*45:01***	SAB vs. SAM	p=0.0365
HLA-B*15:03	SAB vs. SAM	p=0.0238		SAB vs. SAC	p<0.0001
	SAB vs. SAC	p<0.0001	***HLA-B*51:01***	SAC vs. SAB	p=0.0117
HLA-B*15:10	SAB vs. SAC	p=0.0005	***HLA-B*52:01***	SAI vs. SAC	p=0.0227
HLA-B*35:01	SAC vs. SAM	p=0.0385	***HLA-B*58:01***	SAB vs. SAM	p=0.0241
	SAC vs. SAB	p=0.0083		SAB vs. SAC	p=0.0022
HLA-B*37:01	SAI vs. SAM	p=0.0349	***HLA-B*58:02***	SAB vs. SAI	p=0.0034
	SAI vs. SAC	p=0.0039		SAM vs. SAC	p=0.0204
				SAB vs. SAC	p<0.0001

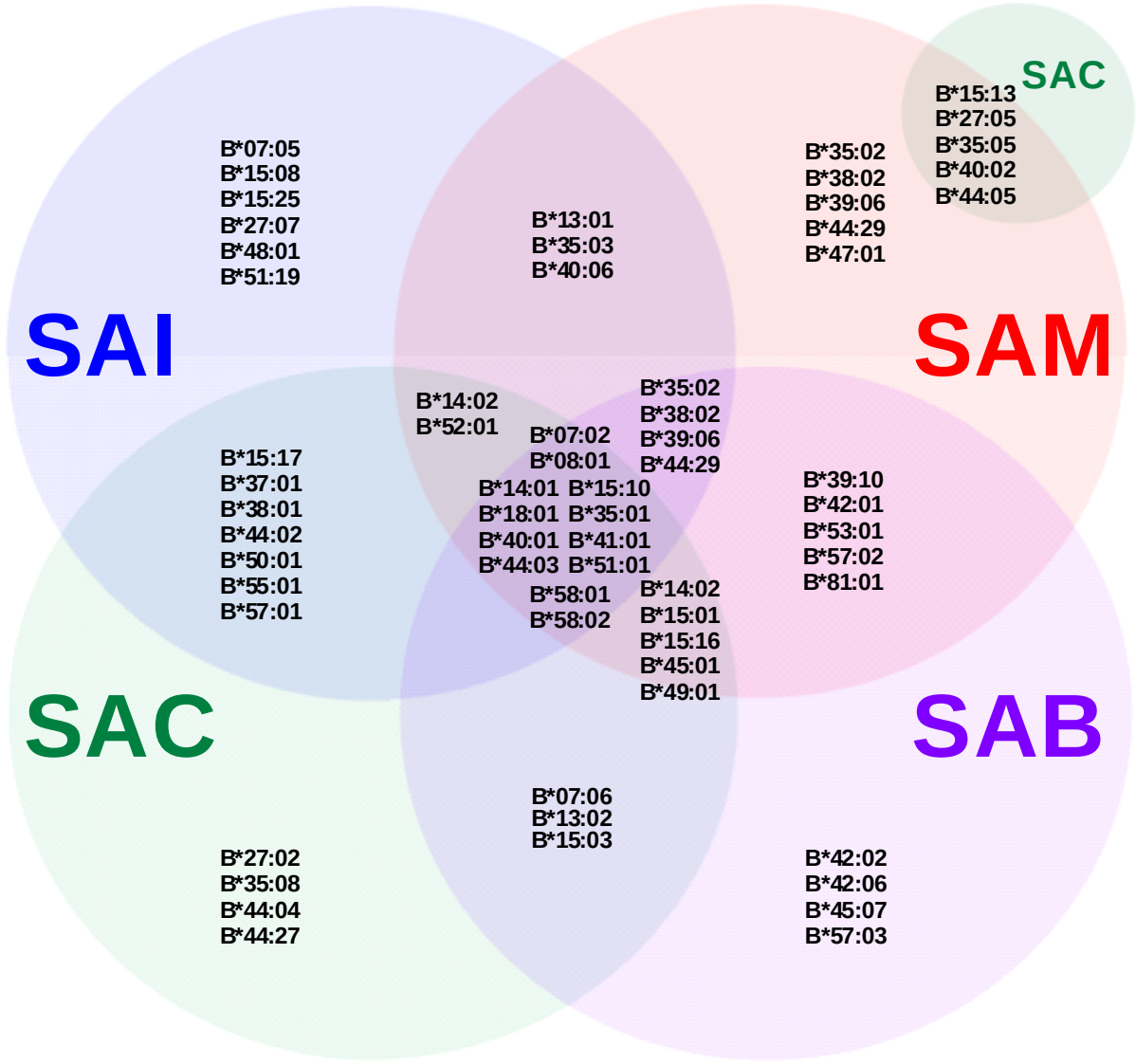

Figure 3.7. Distribution of shared and unique *HLA, human leukocyte antigen class I B* alleles determined for SAI, South African Indian (n=50), SAM, South African Mixed ancestry (n=50), SAC, South African Caucasian (n=94) and SAB, South African Black (n=159) populations.

Table 3.5. Allele frequencies *HLA-C* alleles determined for SAI (n=50), SAM (n=50), SAC (n=94) and SAB (n=159) populations.

	Populations*			
HLA-C	**SAI**	**SAM**	**SAC**	**SAB**
C*01:02	0.060	0.020	0.021	
C*02:02		0.050	0.059	0.009
C*02:05			0.005	0.003
C*02:10		0.010		0.085
C*03:02	0.050	0.010		0.016
C*03:03		0.010	0.059	
C*03:04	0.030	0.070	0.059	0.041
C*03:16			0.005	
C*04:01	0.090	0.090	0.090	0.129
C*04:03	0.040	0.010		
C*04:04				0.003
C*04:08			0.005	
C*05:01		0.020	0.053	0.003
C*06:02	0.140	0.120	0.069	0.157
C*06:06				0.003
C*06:11			0.011	
C*07:01	0.120	0.080	0.176	0.069
C*07:02	0.060	0.090	0.149	0.066
C*07:04	0.010	0.040	0.011	0.013
C*07:06	0.020		0.021	0.041
C*07:11				0.003
C*07:18			0.016	0.044
C*07:26	0.010			
C*08:01	0.020	0.040	0.005	0.003
C*08:02	0.010	0.040	0.021	0.019
C*08:03	0.010			
C*08:04			0.011	0.031
C*12:02	0.070	0.020	0.011	
C*12:03	0.070	0.060	0.021	0.016
C*14:02	0.050	0.010	0.011	0.006
C*15:02		0.030	0.027	0.003
C*15:05		0.010		0.006
C*15:07	0.010			
C*16:01	0.010	0.040	0.069	0.075
C*16:02	0.010		0.011	
C*17:01		0.070	0.005	0.110
C*18:01		0.030		0.016
C*18:02		0.030		0.028

HLA, *human leukocyte antigen;* SAI, South African Indian; SAM, South African Mixed ancestry; SAC, South African Caucasian; SAB, South African Black; Allele frequency defined as n/2 x number of individuals.

***Inter-population frequency comparisons for *HLA-C* alleles showing significance (p<0.05):**

HLA-C*02:02	SAM vs. SAB	p=0.0215	***HLA-C*12:03***	SAI vs. SAB	p=0.0101
	SAC vs. SAB	p=0.0031		SAM vs. SAB	p=0.0262
HLA-C*02:10	SAB vs. SAM	p=0.0054	***HLA-C*14:02***	SAI vs. SAB	p=0.0100
HLA-C*05:01	SAC vs. SAB	p=0.0003	***HLA-C*15:02***	SAM vs. SAB	p=0.0441
HLA-C*06:02	SAB vs. SAC	p=0.0034		SAC vs. SAB	p=0.0286
HLA-C*07:01	SAM vs. SAC	p=0.0329	***HLA-C*16:01***	SAC vs. SAI	p=0.0396
	SAC vs. SAB	p=0.0003		SAB vs. SAI	p=0.0139
HLA-C*07:02	SAC vs. SAB	p=0.0030	***HLA-C*17:01***	SAM vs. SAC	p=0.0030
HLA-C*08:01	SAM vs. SAB	p=0.0127		SAB vs. SAC	p<0.0001
HLA-C*12:02	SAI vs. SAC	p=0.0096			

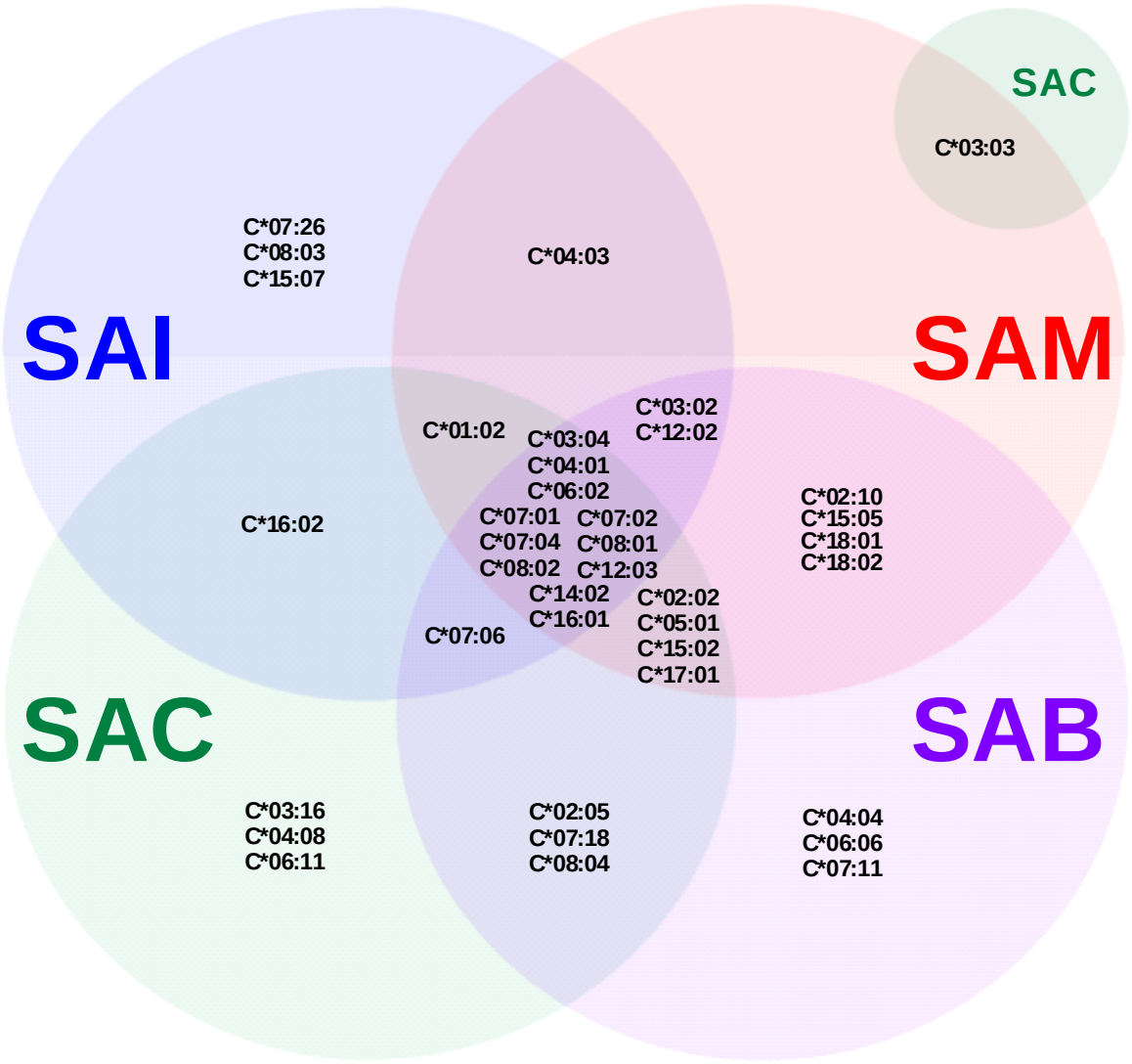

Figure 3.8. Distribution of shared and unique *HLA, human leukocyte antigen class I C* alleles determined for SAI, South African Indian (n=50), SAM, South African Mixed ancestry (n=50), SAC, South African Caucasian (n=94) and SAB, South African Black (n=159) populations.

3.3. Discussion

An extensive immunogenetic characterisation of healthy individuals of SAI, SAM, SAC and SAB ethnicity was achieved and the background frequencies of *HLA-ABC* alleles, *KIR* genotypes, *KIR-HLA* combinations, *KIR3DL1/S1* and *KIR2DS4f/v* copy number variation as well as the prevalence of *CCR5Δ32,* HLA-B signal peptide p2-M/T and *HLA-C 3'UTR +263*-ins/del variants were determined. The four way comparison of frequencies of genes important in innate and adaptive immunity between the SAI, SAM, SAC and SAB populations revealed, as expected, a large number of differences, many of which reached statistical significance which have been reported in the results section.

A few differences of interest were observed in the distribution of *KIR* genes that was most divergent between the SAI population and the other three populations. The highest frequencies of the activating *KIRs*, *KIR2DS1* (66.0%), *KIR2DS2* (78.0%) and *KIR2DS3* (56.0%), *KIR2DS5* (56.0%) and *KIR3DS1* (54.0%), although lower frequencies of *KIR2DS4* (88.0%) and the pseudogene *KIR2DP1* (94.0%) was observed in the SAI population comparison to the three other populations. Similarly, the inhibitory *KIRs, KIR2DL2* (80.0%) and *KIR2DL5* (84.0%) had the highest frequencies, except for *KIR2DL3* (64.0%) which had the lowest frequency in the SAI population. A rare individual from the SAI population was identified with the framework gene *KIR2DL4* absent. Low frequencies of *KIR2DL4* negative individuals have been reported previously in studies from Brazil, China (Gonzalez-Galarza et al., 2011), Equatorial Guinea (Gomez-Lozano et al., 2003, Gonzalez-Galarza et al., 2011) and Poland (Nowak et al., 2011). The SAB population had very low frequencies of the activating *KIRs*, *KIR2DS1* (12.6%) and *KIR3DS1* (6.9%) in comparison to the other three populations. *KIR3DS1* is known to be present at low frequencies in African populations (Norman et al., 2007).

The copy numbers of *KIR3DL1* and *KIR3DS1* were quantified and showed that the SAM and SAC populations were the most similar in the distribution of copy number variants, however the SAI and SAB populations had divergent profiles of copy number variants. The SAI population was the only population with one individual lacking both *KIR3DL1* and *KIR3DS1* entirely. Individuals with only a single copy of *KIR3DS1* were rare or absent in the other populations, but present in 8.0% of individuals in the SAI population. Approximately a third of SAI, SAM and SAC populations carried one copy each of *KIR3DL1* and *KIR3DS1*,

but a significantly low frequency of individuals in the SAB populations carried this copy number variant (6.3%). The majority of SAB individuals carried 2 copies of *KIR3DL1* in the absence of *KIR3DS1* (83.6%) which is expected considering the lack of *KIR3DS1* in this population. There was evidence of duplication of KIR3DS1 or KIR3DL1 in the SAI, SAM and SAC populations, but not in the SAB population. Evidence for deletion of KIR3DS1 or KIR3DL1 was found in the SAI and SAB populations, but not in the SAM and SAC populations. Duplication and deletion of *KIR3DL1/S1* has been reported in minority of Caucasian individuals (Jiang et al., 2012, Pelak et al., 2011).

The copy numbers of full-length *KIR2DS4f* and truncated *KIR2DS4v* were also measured. The absence of both *KIR2DS4f* and *KIR2DS4v* was at a low frequency and similar in the SAM, SAC and SAB populations, but significantly higher in the SAI population (12.0%). All of the individuals who did not carry *KIR2DS4* were in fact *KIR2DS1*-positive; hence the absence of *KIR2DS4* is unlikely to be due to a deletion. This is because *KIR2DS4* and *KIR2DS1,* although not present at the same locus, are inherited in a mutually exclusive pattern (Hsu et al., 2002). However, in some cases deletion of *KIR2DS4* can occur and has been described in a minority of Caucasian individuals (Jiang et al., 2012). In the SAC population 61.7% of SAC individuals had only a single copy of *KIR2DS4v*. Since *KIR2DS4v* encodes a truncated protein, this is large number of individuals without a functional activating receptor. The SAI population also showed high frequencies of individuals carrying a non-functional *KIR2DS4,* with 38% and 22% in individuals carrying single and double copies of *KIR2DS4v,* respectively, in the absence of *KIR2DS4f.*

In all populations, collectively, 55 different *KIR* genotypes could be identified. This included 5 new genotypes (BxN1-N5) that have not been reported before, three of which were identified in the SAI population (BxN1-N3), while one each was detected in the SAM and SAC population (BxN4 and BxN5, respectively). For each individual population, the number of *KIR* genotypes detected was 21, 18, 24 and 32 for SAI, SAM, SAC and SAB populations, respectively. The most common *KIR* genotypes detected in SAI populations were Bx6 and Bx76 both at frequencies of 12.0%, while the AA1 genotype was the most common in the other three populations at 22.0%, 26.6% and 25.8% in the SAM, SAC and SAB populations, respectively.

The frequencies of *HLA* ligands were different in the four populations. For *HLA-A*, *HLA-A Bw4*^{80I} homozygosity was rare in SAI, SAM and SAB populations and absent in the SAC population. *HLA-A non-Bw4* homozygosity was the most common genotype in all populations, but lowest in the SAI population (58.0%). For *HLA-B*, *HLA-B Bw4*^{80I}*/Bw6* heterozygotes and *HLA-B Bw6* homozygotes were the most common genotypes detected in all populations. All permutations of *HLA-B* ligands were detectable in the overall analysis, however, interestingly, no *HLA-B Bw4*^{TL} homozygotes were found in any of the populations studied. At the *HLA-C* locus, all populations had similar frequencies of *HLA-C1/C2* heterozygotes approaching 50% of the population. SAI and SAC populations had higher frequencies of *HLA-C1* homozygotes compared to SAM and SAB populations, while in SAM and SAB populations the frequencies of *HLA-C2* homozygotes were higher.

Since the genes that encode the *KIRs* and *HLA* ligands are present on separate chromosomes (6 and 19, respectively), *KIR-HLA* ligand pairs are not necessarily inherited together. The SAB and SAM populations had the highest frequencies of the *KIR3DL1/HLA-B Bw4*^{80I} combination (52.8% and 46.0%, respectively). The SAB and SAM populations also had the lowest frequencies of the *KIR3DL1/HLA-B Bw4*^{80T} combination (13.8% and 20.0%, respectively). The frequency of *KIR2DL1/HLA-C2* was significantly higher in the SAB population (78.6%) compared to the other three populations, while *KIR2DL3/HLA-C1* was significantly higher in the SAC population (80.9%) compared to the other populations. A very low frequency of the *KIR2DS1/HLA-C2* combination was detected in the SAB population (9.4%).

The prevalence of the *CCR5Δ32* variant was previously unknown in the SAI and SAM population and was shown to be present at low frequencies in both the SAI (4.0%) and SAM (6.0%) population. This result was not unexpected, however, since *CCR5Δ32* has been detected at low frequencies (1.5%) in mainland Indian populations (Verma et al., 2007) and due to SAC ancestry within the SAM population, it was predicted that *CCR5Δ32* would also be present in this population. Notably, none of the carriers of *CCR5Δ32* detected in the SAI and SAM populations were homozygous, however, a single individual was detected in the SAC population who was homozygous for *CCR5Δ32*. A high frequency of *CCR5Δ32* heterozygotes was determined in the SAC population (17.0%). This is higher than generally reported for Caucasian populations in Europe (~10%) (Martinson et al., 1997), however this

result is in agreement with a frequency of 14.3% reported for SAC individuals in another study (Picton et al., 2010).

HLA class I genotyping revealed extensive diversity at all three loci, with some alleles shared amongst all populations and many unique alleles only found in a single population. Overall, 141 different *HLA class I* alleles were detected consisting of 42 different *HLA-A* alleles, 61 different *HLA-B* alleles and 38 different *HLA-C* alleles. For *HLA-A,* 19, 24, 25 and 31 different alleles were identified in SAI, SAM, SAC and SAC populations, respectively. *HLA-A*02:01* was the most common in both the SAM and SAC populations (f=0.210 and f=0.255, respectively while *HLA-A*11:01* (f=0.170) and *-A*30:01* (f=0.110) were the most common alleles in the SAI and SAB populations, respectively. For *HLA-B,* 30, 41, 38 and 29 different alleles were identified in SAI, SAM, SAC and SAC populations, respectively. *HLA-B*07:02* was the most common allele in both the SAM and SAC populations (f=0.08 and f=0.154, respectively), while *HLA-B*40:06* (f=0.140) and *-B*58:02* (f=0.094) was the most common allele in the SAI and SAB populations, respectively. Overall, 38 different *HLA-C* alleles were detected in the four populations with 20, 24, 26 and 27 different alleles identified in SAI, SAM, SAC and SAC populations, respectively. For *HLA-C, HLA-C*06:02* was the most common allele found in three populations (SAI, SAM and SAB population) at frequencies of f=0.14, f=0.12 and f=0.157, respectively, while *HLA-C*07:01* was the most common allele in the SAC population (f=0.176).

In conclusion, **Chapter 3** has contributed important host genetic background information, previously unavailable for the SAI and SAM populations, with additional information added for the SAC and SAB populations for comparative purposes. This chapter thus serves as the basis for subsequent chapters that will address the HLA class I-mediated immune hypersensitivity reaction to antiretroviral drugs (**Chapter 4**) and the potential role of host genetic factors associated with HIV-1 acquisition and transmission (**Chapter 5**).

Chapter 4:
Frequencies of immune hypersensitivity reaction-associated *HLA class I* alleles in healthy South African Indian and Mixed ancestry populations determined by a novel real-time PCR assay

4.1. Introduction

Antiretroviral (ARV) therapy reduces morbidity and mortality associated with HIV-1infection; however, treatment success is often compromised by drug toxicity, viral resistance development and immune complications such as immune reconstitution inflammatory syndrome (IRIS) and immune hypersensitivity reaction (IHR). In susceptible individuals, IHR develops rapidly following the administration of the nucleoside reverse transcriptase inhibitor (NRTI), Abacavir (ABC) or the non-nucleoside reverse transcriptase inhibitor (NNRTI), Nevirapine (NVP). IHR is potentially life-threatening due to excessive activation of autoimmune CD8+ cytotoxic T cells (CTLs) and increased production of pro-inflammatory cytokines, such as interferon-γ (Pompeu et al., 2012). Symptoms include fever, rash, nausea, vomiting, diarrhoea, abdominal pain, dyspnoea, sore throat, cough and malaise (Yuen et al., 2008). In extreme circumstances, failure to withdraw the drug, or re-initiation of therapy after a period of discontinuation, can lead to multiple organ failure and death of the patient.

ABC IHR associates with the presence of *human leukocyte antigen (HLA)-B*57:01*, although, some carriers do not develop a response (Hetherington et al., 2002, Mallal et al., 2002, Martin et al., 2004b, Phillips et al., 2005). Similarly, NVP IHR associates with the presence of *HLA-B*35:05* documented in Thai populations (Chantarangsu et al., 2009, Yuan et al., 2011); *HLA-C*04* in Thai, Chinese and African populations (Likanonsakul et al., 2009, Gao et al., 2012, Yuan et al., 2011); *HLA-C*08* in Japanese populations (Gatanaga et al., 2007) and *HLA-B*14-C*08* haplotypes in Caucasian populations (Littera et al., 2006). Recently, the *HLA-C*04:01* allelic variant was specifically associated with NVP IHR in a Black African population (Carr et al., 2013a). An association of *HLA class II* alleles, such as *HLA-DRB1*01* (Vitezica et al., 2008) and *HLA-DRB1*0101* (Martin et al., 2005), with cutaneous rash/hepatotoxicity in Caucasians treated with NVP, has also been described and more recently, a study in South African populations has linked NVP hepatotoxicity with the presence of *HLA-DRB1*0102* as well as *HLA-B*58:01* (Phillips et al., 2013). The current study has focused on *HLA class I* alleles only, although detection of *HLA-B*58:01* has not been included.

*HLA-B*57:01* is uncommon in individuals of Black African descent, but is present at an allele frequency of 7.1% in Caucasians and 8.2% in individuals from the Indian subcontinent (Gonzalez-Galarza et al., 2011). The allele frequency of *HLA-B*35:05* is 10.2% in South America and 4.2% in Oceania, but is rarely detected in individuals of European descent and absent in South Asian and sub-Saharan African populations (www.ncbi.nlm.nih.gov/projects/gv/mhc/ihwg.cgi). The prevalence of *HLA-B*35:05* is high in Indonesia (16.5%) and the Philippines (12.0%) (Gonzalez-Galarza et al., 2011). The allele frequency of *HLA-C*04* ranges from 9.4% in North-East Asia to 39.9% in Australia; while *HLA-C*08* is present in approximately 5% of European and sub-Saharan African individuals and is common in Oceania at an allele frequency of 18.2% (www.ncbi.nlm.nih.gov/projects/gv/mhc/ihwg.cgi).

Diverse populations reside in South Africa and comprise of individuals of SA Black (SAB), SA Caucasian (SAC) and SA Indian (SAI) ethnicity as well as SA individuals of Mixed ancestry (SAM). *HLA class I* allele frequencies for the SAC and SAB population have been reported in a previous study; however, the SAI and SAM population was not included in this work (Paximadis et al., 2012). Individuals of the SAM population carry complex genetic admixtures derived from at least four lineages (European, Black African, South Asian and Indonesian) (de Wit et al., 2010, Patterson et al., 2010); hence, the presence of *HLA class I* alleles associated with ABC and NVP IHR is likely. A cohort of healthy SAI individuals was also included in this study, although a small study (n=51) reporting *HLA class I* allele frequencies in this population is available (Gonzalez-Galarza et al., 2011).

Prospective *HLA-B*57:01* screening has been successfully implemented in developed countries such as Australia and the USA (Rauch et al., 2006, Mallal et al., 2008) and recent *HLA-B*35:05* genotyping was shown to reduce the occurrence of NVP IHR in Thailand (Kiertiburanakul et al., 2013). Currently, this intervention is not in practice in resource-limited settings, such as countries in Africa, where NVP is widely used and ABC may have been introduced in some regions. ABC and NVP is prescribed in South Africa, where one of the largest national ARV treatment programs in the world is in operation, with two million HIV-positive individuals accessing treatment in October 2012 (www.doh.gov.za/docs/presentations/2013/Antenatal_Sentinel_survey_Report2012_final.pdf). There are limited reports, however, that describe the occurrence of IHR following exposure to ABC and NVP in South Africa. A 3.5% prevalence of NVP-associated rash in a local

national mother-to-child HIV-1 transmission prevention program has been reported (Black et al., 2008), while a sub-Saharan HIV-1 treatment study that included three local sites, reported the development of cutaneous rash of any grade in 54.0% of women receiving NVP twice-daily (Dong et al., 2012). A frequency of 6.0% of individuals developed hypersensitivity reactions to ABC within the first 28 days of treatment in a multisite study that included one local site (Wood et al., 2004), while in a multisite paediatric HIV-1 treatment study that included 4 sites in South Africa, 9/109 (8.3%) children developed suspected ABC-hypersensitivity reactions (Fortuny et al., 2014).

Several molecular and serological methods have been described previously for the detection of *HLA-B*57:01* (Stocchi et al., 2012). There was also the suggestion that detection of the guanine (G) variant of the *HCP5* (HLA complex protein 5) rs2395029 single nucleotide polymorphism (SNP), also termed "*HCP5* 568G", known to be in strong linkage disequilibrium with *HLA-B*57:01* (Colombo et al., 2008, Rodriguez-Novoa et al., 2010) could be used solely as a predictor of *HLA-B*57:01* carriage. The current study has used a novel strategy to determine the presence of *HLA-B*57:01* alleles using a real-time allele-specific PCR (AS-PCR) assay and also included the detection of *HLA-B*35:05, -C*04* and -*C*08.* Although a recent real-time AS-PCR assay for the detection of *HLA-B*57:01* has been published (Dello Russo et al., 2014), the detection strategy used in the current assay was different and more informative. To our knowledge, no real-time AS-PCR detection assays for *HLA-B*35:05*, *-C*04* and *-C*08* are available. The advantages of using real-time AS-PCR assays includes less template DNA required, lower cost of reagents and consumables and a faster turn-around time when compared to conventional sequence-based typing (SBT) methods. The assay was first developed and evaluated using DNA samples previously *HLA class I* genotyped by SBT, then subsequently applied to a cohort of healthy individuals of SAM and SAI ethnicity. Results were combined with the reported population frequencies of *HLA class I* alleles detected in the SAC and SAB population. In addition, South African demographic data was used to estimate the number of HIV-1 positive individuals potentially at risk of developing ABC and NVP IHR, who may benefit from the assay.

4.2. Results

4.2.1. Real-time AS-PCR assay development and validation

DNA samples of known *HLA class I* genotype (n=120) were available from SAB and SAC individuals from a previous study (Paximadis et al., 2012). Samples representing alleles of interest were selected for development and evaluation of the AS-PCR primers as follows: *HLA-B*57:01* (*n*=8), *-B*57:02* (*n*=6), *-B*57:03* (*n*=9), *-B*35:05* (*n*=1), *-C*04:01* (*n*=65), *-C*04:04* (*n*=2), *-C*04:08* (*n*=1), *-C*08:01* (*n*=2), *-C*08:02* (*n*=14) and *-C*08:04* (*n*=12). Each primer set was tested with all 120 samples to also assess false positive/negative amplification potential. The *HLA-B*57* group-specific primers were evaluated with positive control samples representing *HLA-B*57:01*, *-B*57:02* and *-B*57:03*. The presence of *HLA-B*57* was correctly detected in these samples and was also negative for all other *HLAs* tested. The primers specific for detection of coding sequences for residues 114D and 116S correctly identified the presence of these codons in *HLA-B*57:01* samples only and not in the samples that contained *HLA-B*57:02* and *-B*57:03*. *HCP5* 568G was detected in *HLA-B*57:01* samples only and not in samples containing *HLA-B*57:02* and *-B*57:03* or in samples containing other *HLA*s. **Figure 4.1** shows the difference in raw data plots between samples positive for *HLA-B*57:01*, *HLA-B*57:02* and *HLA-B*57:03.* The *HLA-B*35:05*-specific primers correctly amplified *HLA-B*35:05* with no observed false positive amplification of other *HLA*s tested. The *HLA-C*04* group-specific primers were tested successfully with samples representing *HLA-C*04:01*, *-C*04:04* and *-C*04:08*, while the *HLA-C*08* group-specific primers were tested successfully with samples representing *HLA-C*08:01*, *-C*08:02* and *-C*08:04*. No false positives were observed for other *HLAs* tested for *HLA-C*04* and *-C*08*.

For additional evaluation purposes, a panel of DNA samples (n=30) representative of more diverse global populations, available from the International Histocompatibility Working Group (IHWG), was included in the validation of the real-time AS-PCR assays and included the following relevant alleles: *HLA-B*35:01* (*n*=2), *-B*35:03* (*n*=1), *-B*35:05* (*n*=1), *-B*57:03* (*n*=2), *-C*04:01* (*n*=6), *-C*08:01* (*n*=2) and *-C*08:02* (*n*=2). The *HLA-B*57* primers detected *HLA-B*57:03* and did not show the presence of 114D/116S or *HCP5* 568G. *HCP5* 568G was not detected in any of the samples tested. *HLA-B*35:05* was correctly

identified and no amplification was observed for *HLA-B*35:01* or *-B*35:03*. The *HLA-C*04* primers correctly identified *HLA-C*04:01* and likewise, the *HLA-C*08* primers correctly identified *HLA-C*08:01* and *-C*08:02* with no false positive amplifications observed.

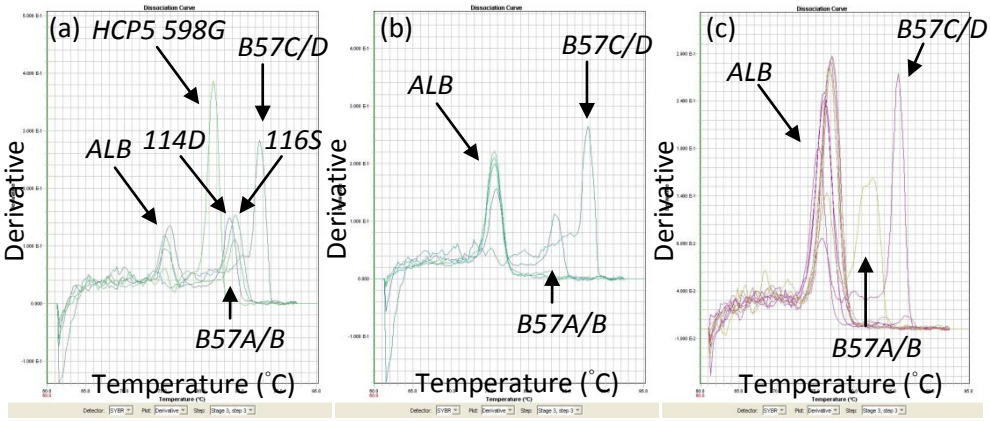

Figure 4.1. Melting curve analysis of (a) *HLA-B*57:01*-positive; (b) *HLA-B*57:02*-positive and (c) *HLA-B*57:03*-positive samples showing the presence of melting peak for ALB control gene and *HLA-B*57* group-specific primers in all samples, however, the presence of melting peaks corresponding to the *HCP5 598G* variant and residues 114D/116S are only present in *HLA-B*57:01*-positive samples.

4.2.2. Real-time AS-PCR detection of *HLA-B*57:01, -B*35:05, -C*04* and *-C*08*

*HLA-B*57:01* was detected in 8.0% and 12.0% of individuals in the SAM and SAI population, respectively (**Table 4.1**), which was similar to the frequency of 8.2% calculated for SAC populations (Paximadis et al., 2012). *HCP5* 568G was present in all *HLA-B*57:01* samples; however, *HCP5* 568G was also detected in one individual of SAM ethnicity, who was negative for *HLA-B*57:01* by the AS-PCR assay. *HLA class I* genotyping by SBT confirmed that this individual was not an *HLA-B*57:01* carrier and therefore *HCP5* 568G detection cannot be used as an absolute predictor of *HLA-B*57:01* positivity in the SAM population. *HLA-B*35:05* was not detected in SAI individuals, but was present in SAM individuals at a frequency of 6.0% (**Table 4.1**). *HLA-B*35:05* has not been reported in SAB or SAI populations and has a low population frequency of 1.0% in SAC individuals (Paximadis et al., 2012, Gonzalez-Galarza et al., 2011). *HLA-C*04* was detected at frequencies of 22.0% and 24.0% of individuals in the SAM and SAI population, respectively (**Table 4.1**), which was similar to the calculated population frequencies for SAC (18.6%) and SAB (23.7%) individuals (Paximadis et al., 2012). *HLA-C*08* was detected in 10.0% of SAM individuals and 8.0% of SAI individuals (**Table 4.1**) which was similar to the calculated frequencies for SAC (6.9%) and SAB (10.2%) populations (Paximadis et al., 2012). To confirm the validity of results generated by the AS-PCR assay, all SAI and SAM samples were *HLA class I* genotyped by SBT.

Table 4.1. Frequency of individuals carrying *HLA class I* alleles associated with ABC and NVP immune hypersensitivity reaction for healthy SAI, SAM, SAC and SAB populations.

***HLA class I* genotype**	**SAI**	**SAM**	**SAC**	**SAB**
*HLA-B*57:01*	6/50(12.0%)	4/50(8.0%)	8/97(8.2%)	0/196(0.0%)
*HLA-B*35:05*	0/50(0.0%)	3/50(6.0%)	1/96(1.0%)	0/196(0.0%)
*HLA-C*04*	12/50(24.0%)	11/50(22.0%)	19/102(18.6%)	47/198(23.7%)
*HLA-C*08*	4/50(8.0%)	5/50(10.0%)	7/102(6.9%)	20/197(10.2%)

HLA, human leukocyte antigen; ABC, Abacavir; NVP, Nevirapine; SAI, South African Indian; SAM, South African Mixed ancestry; SAC, South African Caucasian; SAB, South African Black

4.2.3. Hypothetical analysis of individuals potentially at risk of ABC and NVP IHR

The most recent South African national census (2013) provides estimates of the number of individuals per population (http://www.statssa.gov.za/publications/P0302/P03022013.pdf) as follows: SAB (n=42,284,100), SAC (n=4,602,400), SAM (n=4,766,200) and SAI (n=1,329,300). Updated HIV-1 prevalence rates stratified by population are not available, however, a national survey conducted in 2012 estimated rates of 0.8%, 3.1%, 0.3% and 15.0% in SAI, SAM, SAC and SAB populations, respectively (http://www.mrc.ac.za/pressreleases/2009/sanat.pdf). These figures were used to estimate that 14,235 HIV-positive individuals would potentially be at risk of developing IHR in response to ABC exposure and 2,212,700 HIV-positive individuals would potentially be at risk of developing IHR in response to NVP therapy (**Table 4.2**).

Table 4.2. Estimated numbers of South African individuals based on HIV-1 status and carriage of at least one *HLA class I* allele associated with ABC or NVP immune hypersensitivity reaction.

***HLA class I* genotype** **HIV-1 status**	**SAI** **n=1,329,300**	**SAM** **n=4,766,200**	**SAC** **n=4,602,400**	**SAB** **n=42,284,100**	**Total** **n=52,982,000**
HIV-1(+)	10,634	147,752	13,807	6,342,615	6,514,808
*HLA-B*57:01(+)*	159,516	381,296	379,579	0	920,391
*HLA-B*57:01(+) HIV-1(+)*	1,276	11,820	1,139	0	14,235
*HLA-B*35:05(+)*	0	285,972	47,447	0	333,419
*HLA-B*35:05(+) HIV-1(+)*	0	8,865	142	0	9,007
*HLA-C*04(+)*	319,032	1,048,564	857,310	10,037,135	12,262,041
*HLA-C*04(+) HIV-1(+)*	2,552	32,505	2,572	1,505,570	1,543,199
*HLA-C*08(+)*	106,344	476,620	315,851	4,292,802	5,191,617
*HLA-C*08(+) HIV-1(+)*	851	14,775	948	643,920	660,494

HLA, human leukocyte antigen; ABC, Abacavir; NVP, Nevirapine; HIV-1(+); HIV-1 positive; SAI, South African Indian; SAM, South African Mixed ancestry; SAC, South African Caucasian; SAB, South African Black

4.3. Discussion

*HLA-B*57:01* is absent in SAB individuals, but is present in 8.2% of SAC individuals (Paximadis et al., 2012) and has been reported in SAI individuals at an allele frequency of 10.2% (Gonzalez-Galarza et al., 2011). The presence and frequency of *HLA-B*57:01* in the SAM population is not known; however, using a novel real-time AS-PCR assay, 8.0% of individuals in this population were found to be carriers. Similarly, *HLA-B*57:01* was detected in 12.0% of SAI individuals. Furthermore, all samples identified as *HLA-B*57:01* positive by the AS-PCR assay also tested positive for *HCP5* 568G, in agreement with previous observations that these genetic loci are in strong linkage disequilibrium (Colombo et al., 2008, Rodriguez-Novoa et al., 2010). However, *HCP5* 568G was also detected in one *HLA-B*57:01*-negative individual of SAM ethnicity. Similar observations have been made in two Caucasian studies (Badulli et al., 2012, Rodriguez-Novoa et al., 2010) where 3/66 (4.5%) and 1/15 (6.7%) individuals with *HCP5* 568G were *HLA-B*57:01* negative. *HCP5* 568G can also be absent in *HLA-B*57:01* carriers as observed in 6/104 (5.8%) and 2/119 (1.7%) individuals from two Caucasian studies (Badulli et al., 2012, Colombo et al., 2008). In the current study, however, none of the *HLA-B*57:01* positive individuals were negative for *HCP5* 568G, suggesting that in the SAM and SAI populations, detection of *HCP5* 568G alone was positive predictor for exclusion from ABC treatment. However, as in one individual from the SAM population, the presence of *HCP5* 568G in the absence of *HLA-B*57:01* would deny eligibility for ABC treatment. The inclusion of *HLA-B*57* group-specific and 114D/116S primers can be used to resolve this situation.

*HLA-B*35:05* is absent in the SAI and SAB population and has a low frequency of 1.0% in the SAC population (Paximadis et al., 2012, Gonzalez-Galarza et al., 2011). Using the AS-PCR assay, *HLA-B*35:05* was not detected in SAI individuals, in agreement with the previous report (Gonzalez-Galarza et al., 2011), but was observed at a relatively high frequency of 6.0% in SAM individuals. This was an unexpected finding that questioned the origin of *HLA-B*35:05* in the SAM population considering the scarcity of this allele in the SAB and SAC populations, that are shared lineages. *HLA-B*35:05* is not found in Southern Asia, but is prevalent in South America and is also present in Oceania (http://www.ncbi.nlm.nih.gov/projects/gv/mhc/ihwg.cgi). Oceania includes Indonesia where a high population frequency (16.5%) of *HLA-B*35:05* has been reported (Gonzalez-Galarza et al., 2011). Since the SAM population is known to have Indonesian ancestry (de Wit et al.,

2010, Patterson et al., 2010), there is a likelihood that *HLA-B*35:05* may be derived from the Indonesian lineage, although a Caucasian influence cannot be excluded.

*HLA-C*04* is present in 18.6% of SAC and 23.7% of SAB individuals (Paximadis et al., 2012) and was detected at similar frequencies of 22.0% and 24.0% in SAM and SAI populations, respectively, by the AS-PCR assay. Notably, a recent African study identified the *HLA-C*04:01* allelic variant as a risk factor for the development of NVP IHR (Carr et al., 2013b). The real-time AS-PCR assay does not differentiate this allele from other *HLA-C*04* alleles, but does include its detection. *HLA-C*08* is present in 6.9% of SAB individuals and 10.2% of SAC individuals (Paximadis et al., 2012). An allele frequency of 4.2% has also been reported for the SAI population (Gonzalez-Galarza et al., 2011). Similar frequencies of *HLA-C*08* were detected by the AS-PCR assay in SAM and SAI populations (10.0% and 8.0%, respectively). A Caucasian study identified the *HLA-B*14-C*08* haplotype as a risk factor for NVP IHR (Littera et al., 2006), however, the role of *HLA-B*14* is uncertain due to the high linkage disequilibrium between *HLA-B*14* and *-C*08* known for this population (Sardinia). This strong linkage disequilibrium has also been reported in the SAC and SAB population (Paximadis et al., 2012). Since, *HLA-C*08* has been reported as an independent risk factor for NVP IHR, albeit in a Japanese population (Gatanaga et al., 2007) where *HLA-B*14* is uncommon (Gonzalez-Galarza et al., 2011), it was decided to omit detection of *HLA-B*14* in the current assay, even though *HLA-B*14* is prevalent in the SAB and SAC population at allele frequencies of 3.3% and 2.6%, respectively (Paximadis et al., 2012).

The current study has used South Africa an example to illustrate the potential benefits of a pre-screening assay by estimating the number of individuals potentially at risk of developing IHR. Two million individuals were accessing treatment in October 2012, making the South African national antiretroviral treatment program one of the largest in the world and currently, the South African Department of Health (DOH) recommends ABC for the treatment of adolescent and paediatric HIV-1 infection, while NVP is recommended for adult HIV-1 treatment when efavirenz is contra-indicated (www.doh.gov.za/docs/policy/2013/ART_Treatment_Guidelines_Final_25March2013.pdf). Using population statistics, HIV-1 prevalence data and the reported/measured frequencies of *HLA class I* alleles, the number of *HLA-B*57:01* carriers who are HIV-1 positive equates to 14,235 individuals. Notably, *HLA-B*57:01* has not been reported in the SAB population, although, low frequencies have been detected in individuals from other African countries,

including Zimbabwe, Kenya, Uganda, Sudan and Zambia (Gonzalez-Galarza et al., 2011). It can therefore be assumed that the risk of ABC IHR will be negligible in the SAB population. *HLA-B*35:05* was estimated to be present in 9,007 HIV-1 positive individuals. Notably, *HLA-B*35:05* is absent in the SAI and SAB populations and should not pose a risk for these individuals. It was estimated that *HLA-C*04* alleles would be present in 1,543,199 HIV-1 positive individuals and similarly, *HLA-C*08* alleles would be present in 660,494 HIV-1 positive individuals. The risk of NVP IHR is highly significant in the SAB population due to a large population size (n=42,284,100), high frequency of *HLA-C*04* (23.7%) and *-C*08* (10.2%) as well as a high HIV-1 prevalence (15.0%). Furthermore, the *HLA-C*04:01* allelic variant, associated with NVP IHR, is common in this population with a reported allele frequency of 11.9% and is also present at allele frequencies of 8.5% in SAC individuals (Paximadis et al., 2012) and 4.2% in SAI individuals (Gonzalez-Galarza et al., 2011). The allele frequency of *HLA-C*04:01* was determined at 9.0 % in the SAM population using *HLA class I* genotyping data. In total, an estimated 14,235 and 2,212,700 HIV-1 positive individuals are potentially at risk of developing IHR in response to ABC and NVP exposure, respectively. Based on these estimates, a substantial number of South African individuals would benefit from the availability of a pre-screening assay before starting treatment, as exemplified by prospective *HLA-B*57:01* and *-B*35:05* genotyping in Australia and Thailand.

Pre-screening of HIV-positive patients for the presence of *HLA class I* alleles associated with ABC and NVP IHR can markedly reduce the risk of morbidity and mortality, however, the implementation of cheaper and quicker molecular tests to substitute conventional *HLA class I* genotyping methods are needed to facilitate this intervention. This study presents the development of a real-time AS-PCR assay to achieve this aim and in addition the previously unknown percentages of *HLA class I* alleles associated with ABC and NVP IHR (*HLA-B*57:01*, *-B*35:05, -C*04* and *-C*08*) in healthy SAM and SAI individuals was determined. Limitations of the assay include the potential of detection of a small subset of false positive alleles that are inevitable due to the large number of *HLA class I* alleles currently described. It is recommended that if the assay was implemented in populations where *HLA class I* alleles that produce false positive results are prevalent, that the contra-indicated drug was not necessarily withheld based on the results, but used with close monitoring for early clinical signs of IHR. Using South Africa as an example, it was estimated that a substantial number of individuals accessing antiretroviral therapy may be at

risk of IHR and would therefore benefit from a genetic pre-screening assay. The assay may also be of use in populations, such as Australia, USA and Thailand, where these alleles are prevalent and ABC and NVP are in use for HIV-1 treatment.

Chapter 5:
Prevalence of protective/deleterious genotypes associated with HIV-1 acquisition/transmission in healthy individuals of South African Indian, Mixed ancestry, Caucasian and Black ethnicity

5.1. Introduction

Host genetic factors can modulate the rate of spread of HIV-1 within a population due to the influence of individual genes or gene combinations on susceptibility/resistance to HIV-1 infection. In addition, HIV-1 transmission, which predominantly occurs via a heterosexual route, can also be influenced by genetic determinants of infectivity, such as high viral load. A primary example of a gene variant consistently associated with both resistance to HIV-1 infection and a low viral load setpoint (VLS) is the 32bp-deletion variant of C-C chemokine receptor type 5, *CCR5Δ32* (Barmania et al., 2013). *CCR5Δ32* homozygotes (Liu et al., 1996), but not heterozygotes (Liu et al., 2012), are resistant to CCR5-tropic HIV-1 acquisition, while heterozygotes have a low VLS that mediates slow disease progression and lower infectivity (van Manen et al., 2011).

Other mediators of HIV-1 acquisition/transmission include cell-surface HLA class I molecules that regulate both adaptive and innate antiviral immune responses following HIV-1 exposure, via activation of CD8+ cytotoxic T cell (CTLs) and natural killer (NK) cells. Specific alleles of *HLA class I* genes are known to associate with VLS (Carlson et al., 2012, Leslie et al., 2010, Frahm et al., 2005, Geldmacher et al., 2007, Honeyborne et al., 2007, Frahm et al., 2006, Tang et al., 2002, Tang et al., 2010, Kiepiela et al., 2004). Susceptibility and resistance to HIV-1 acquisition has also been associated with particular *HLA class I* alleles (Raghavan et al., 2009, Umapathy et al., 2007, Koehler et al., 2010, Peterson et al., 2013, Song et al., 2011).

CTLs respond to immunogenic peptides derived from cytosolic proteins produced by intracellular pathogens, such as HIV-1. Peptide antigens bind to HLA class I molecules expressed on the cell surface. Three related subclasses of HLA class I molecules are expressed (HLA-A, -B and -C) that are encoded by three separate genes present in a highly polymorphic gene cluster mapping to the major histocompatibility complex (MHC) on chromosome 6 (Robinson et al., 2013b). NK cells are regulated in part, by killer-cell immunoglobulin-like receptors (KIRs); a specific family of cell-surface activating/inhibitory molecules mapping to the leukocyte receptor complex (LRC) on chromosome 19 (Middleton and Gonzelez, 2010). KIRs, in combination with cognate HLA class I ligands, are involved in detection and destruction of HIV-1-infected target cells by activated NK cells (Carrington

and Alter, 2012). NK cell responses precede the development of antiviral CTLs following HIV-1 exposure, underlining the importance of innate immune responses in preventing infection (Alter et al., 2007).

Studies in HIV-1 exposed seronegative (HESN) individuals, such as female sex workers, and uninfected partners in serodiscordant couples, have identified specific *KIR-HLA* combinations that associate with resistance/susceptibility to heterosexual HIV-1 acquisition. *KIR-HLA* combinations including *KIR2DL2(+)KIR3DL3(+)HLA-C1(-)HLA-C2(+); KIR3DL1(+)KIR3DS1(-)HLA-Bw4(-)HLA-Bw6(+);KIR3DL1(-)KIR3DS1(+)HLA-Bw4$^{80I/T}$(+); KIR3DL1(+)KIR3DS1(+)HLA-B Bw4$^{80I/T}$(+)HLA-Bw6(+)* and *KIR3DL1(+/-)KIR3DS1(+)HLA-A/B Bw4(+)* associate with resistance to HIV-1 infection, while *KIR2DL3(+)KIR2DL2(-)HLA-C1(+)HLA-C2(+)* and *KIR3DL1(+)KIR3DS1(-)HLA-A/B Bw4(+)* associate with susceptibility to HIV-1 infection (Jennes et al., 2006, Guerini et al., 2011, Habegger de Sorrentino et al., 2013).

Once infected by HIV-1, specific *KIR-HLA* combinations can subsequently influence HIV-1 disease progression via modulation of VLS or CD4+ helper T cell preservation (Gaudieri et al., 2005, Jiang et al., 2013, Martin et al., 2002, Pelak et al., 2011). The risk of heterosexual HIV-1 transmission also associates with viral load (Quinn et al., 2000, Merino et al., 2011) and *KIR-HLA* combinations such as *KIR3DL1(+/-)KIR3DS1(+)HLA-B Bw4^{80I}(+)* and *KIR3DL1(+)KIR3DS1(-)HLA-B Bw4^{80I}(+)* have been associated with a low VLS (Pelak et al., 2011, Jiang et al., 2013), while *KIR3DL1(+)KIR3DS1(+/-)HLA-B Bw4^{80I}(-)* and *KIR3DL1(+)KIR3DS1(-)HLA-B Bw4^{80I}(-)* genotypes have been associated with a high VLS (Jiang et al., 2013).

The cell-surface expression levels of HLA-C molecules have been shown to affect HIV-1 VLS via an insertion (Ins)/deletion (Del) polymorphism at position +263 in the 3'untranslated region (3'UTR) of *HLA-C* messenger RNA (mRNA) transcripts (Kulkarni et al., 2011). The *HLA-C* +263 Del variant is thought to disrupt the binding site of the Hsa-miR-148a microRNA that targets *HLA-C* mRNA transcripts for degradation.

Non-classical HLA class I molecules include HLA-E, -F and -G. NK cells express the inhibitory receptor CD94/NKG2A that interacts with HLA-E ligands on target cells (Adams and Luoma, 2013). HLA-E binds peptides derived from the HLA-B signal peptide sequence

that contain either a methionine (M) or threonine (T) residue at the anchoring position-2 (p2). The p2-T variant is associated with resistance to HIV-1 acquisition, but does not influence VLS (Yunis et al., 2007, Merino et al., 2013). It is thought that HLA-E is less stable in complex with the p2-T peptide variant which leads to lower cell-surface expression levels of HLA-E. Consequently NK cells are less inhibited and able detect HIV-1 infected target cells more efficiently.

In South Africa HIV-1 prevalence is 12.2% (2012) at a national level, however, prevalence rates are disproportionate among the different SA populations. In 2008, the prevalence of HIV-1 was estimated at 0.3%, 0.3%, 1.7% and 13.6% in SAI, SAM, SAC and SAB populations, respectively (www.mrc.ac.za/pressreleases/2009/sanat.pdf). In 2012, these estimates changed to 0.8%, 3.1%, 0.3% and 15.0% in SAI, SAM, SAC and SAB populations, respectively (www.hsrc.ac.za/en/research-outputs/view/6871). Although several reasons ranging from behavioural to biological differences may underlie the low prevalence of HIV-1 infection in the SAC, SAM and SAI populations compared to SAB populations, the potential contribution of host immunogenetic factors has not been investigated in these populations, particularly the SAI and SAM populations.

This study has characterised several immunogenetic factors in healthy individuals of SAI (n=50) and SAM (n=50) ethnicity and has included data generated by others for SAC (n=94) and SAB (n=159) populations in order to make comparisons between the populations with a focus on protective/deleterious genotypes that associate with HIV-1 acquisition and VLS. Real-time allele-specific PCR (AS-PCR) assays were used to determine *CCR5/CCR5Δ32*, *KIR* and *HLA-Bw4*^{80I}*/Bw4*^{80T}*/C1/C2* genotypes, while relative quantification real-time AS-PCR assays were used to determine copy number variation of *KIR3DL1/S1* and *KIR2DS4f/v*. *HLA class I A*, *-B* and *-C* genotyping was done using commercial sequence-based typing (SBT) methods and HLA-B signal peptide p2-M/T variants were determined from these sequences. *HLA-C 3'UTR* +263-ins/del variants were determined by direct DNA sequencing.

5.2. Results

In order to assess the prevalence of various genotypes associated with HIV-1 acquisition and VLS in the four populations, the frequency data from **Chapter 3** was used to perform the respective analyses. Most of the analyses required an examination of combinations of particular genes/alleles and thus provide new data in this chapter. Some of the frequency data have been used directly as already reported in **Chapter 3** and is consequently repetitive, however these data have been included where necessary to ensure this chapter can be read as a standalone study.

5.2.1. Genotypes associated with HIV-1 acquisition

5.2.1.1 Genotypes associated with resistance to HIV-1 infection

5.2.1.1.1. CCR5Δ32 homozygosity

Taken from **Chapter 3** and shown in **Figure 5.1**, no *CCR5Δ32* homozygotes were identified in the SAI and SAM populations. A low frequency of *CCR5Δ32* homozygotes was determined for the SAC population (1.1%), while the frequency of *CCR5Δ32* homozygotes in the SAB population was assumed to be 0%, due to the absence of this gene variant in this population (Picton et al., 2010).

5.2.1.1.2. HLA-B p2-T homozygosity

Taken from **Chapter 3** and shown in **Figure 5.1**, the highest frequency of HLA-B p2-T/T was observed in the SAI population (78.0%), followed by the SAM population (54.0%) and similar frequencies were observed in the SAC and SAB populations (43.6% and 44.7%, respectively). A significant difference was observed in the frequency of HLA-B p2-T homozygotes in the SAI population compared to the SAM, SAC and SAB populations (78.0% vs. 54.0%; $p=0.0196$, 78.0% vs. 43.6%; $p<0.0001$ and 78.0% vs. 44.7%; $p<0.0001$, respectively).

5.2.1.1.3. HLA-ABC

Based on published studies of *HLA class I* alleles associated with resistance or susceptibility to HIV-1 infection (**Appendix B**) and *HLA class I* genotypic data available for the SAI, SAM, SAC and SAB populations, an individual was considered a carrier of an overall *HLA-ABC* genotype profile associated with "resistance" if after assigning each *HLA-A, -B* and *-C* allele to a protective, deleterious or unknown status, the overall cumulative result was an excess of protective alleles (**Figure 5.1**). Individuals were defined as "neutral" if equal numbers of protective and deleterious alleles were present or if all six alleles were of unknown status. Based on this analysis, the frequency of resistance *HLA-ABC* genotypes was highest in the SAI population (44.0%), followed by the SAC population (39.4%), while similar frequencies were observed in the SAM and SAB populations (32.0% vs. 30.2%, respectively). No significant differences were observed between the populations ($p>0.05$).

5.2.1.1.4. KIR2DL2(+)KIR3DL3(+)HLA-C1(-)HLA-C2(+)

The frequency of individuals carrying the protective *KIR-HLA ligand* combination *2DL2(+)3DL3(+)C1(-)C2(+)*, defined by the presence of both *KIR2DL2* and *KIR3DL3* as well as an *HLA-C2* ligand, was determined by analysis of *KIR* genotypes and *HLA ligand* genotypes available for the SAI, SAM, SAC and SAB populations (**Figure 5.1**). The frequency of the *2DL2(+)3DL3(+)C1(-)C2(+)* genotype was highest in the SAB population (13.8%), followed by the SAM population (12.0%), and present at a similar frequency in the SAI and SAC populations (8.0% and 7.4%, respectively). No statistically significant differences were observed between the populations.

5.2.1.1.5. KIR3DL1(+)3DS1(-)HLA-Bw4(-)HLA-Bw6(+)

The frequency of individuals carrying the protective *KIR-HLA ligand* combination *3DL1(+)3DS1(-)Bw4(-)Bw6(+)*, defined by the presence of *KIR3DL1* and absent *KIR3DS1* as well as the presence of an *HLA-B Bw6* ligand and absence of an *HLA-Bw4* ligand, was determined by analysis of *KIR* genotypes and *HLA ligand* genotypes available for the SAI, SAM, SAC and SAB populations (**Figure 5.1**). The frequency of the *3DL1(+)3DS1(-)Bw4(-)Bw6(+)* genotype was the highest in the SAB population (37.1%), lowest in the SAI population (12.0%) and present at similar frequencies in the SAM and SAC populations

(22.0% and 22.3%, respectively). A significant difference was observed with this genotype when the SAI and SAC populations were compared to the SAB population (12.0 % vs. 37.1%; p=0.0007 and 22.3% vs. 37.1%; p=0.0173, respectively).

5.2.1.1.6. KIR3DL1(-)KIR3DS1(+)HLA-B Bw4$^{80I/T}$(+)

The frequency of individuals carrying the protective *KIR-HLA ligand* combination *3DL1(-)3DS1(+)HLA-B Bw4$^{80I/T}$(+),* defined as the presence of *KIR3DS1* and absent *KIR3DL1* as well as the presence of *HLA-B Bw4^{80I}* and/or *HLA-B Bw4^{80T}* ligands, was determined by analysis of *KIR* genotypes and *HLA ligand* genotypes available for the SAI, SAM, SAC and SAB populations (**Figure 5.1**). The frequency of *3DL1(-)3DS1(+)HLA-B Bw4$^{80I/T}$(+)* genotypes was highest in the SAI population (8%), lowest in the SAB population (0.6%) and present at a low frequency in SAM populations (2.0%). This genotype was absent in the SAC population. A statistically significant difference was observed in the frequency of *3DL1(-)3DS1(+)HLA-B Bw4$^{80I/T}$(+)* genotypes when the SAI population was compared to the SAB population (8.0% vs. 0.6%, respectively; p=0.0122).

5.2.1.1.7. KIR3DL1(+)KIR3DS1(+)HLA-B Bw4$^{80I/T}$(+)HLA-Bw6(+)

The frequency of individuals carrying the protective *KIR-HLA ligand* combination *3DL1(+)3DS1(+)HLA-B Bw4$^{80I/T}$(+)Bw6(+),* defined as the presence of both *KIR3DL1* and *KIR3DS1* as well as the presence of both an *HLA-Bw6* and an *HLA-B Bw4^{80I}* or *HLA-B Bw4^{80T}* ligand, was determined by analysis of *KIR* genotypes and *HLA ligand* genotypes available for the SAI, SAM, SAC and SAB populations (**Figure 5.1**). The frequency of *3DL1(+)3DS1(+)HLA-B Bw4$^{80I/T}$(+)Bw6(+)* genotypes was highest in the SAI population (26.0%), lowest in the SAB population (3.8%) and present at similar frequencies in the SAM and SAC populations (16.0% and 18.1%, respectively.). A significant difference in the frequency of *3DL1(+)3DS1(+)HLA-B Bw4$^{80I/T}$(+)Bw6(+)* genotypes was observed in the SAI, SAM and SAC populations compared to the SAB population (26.0% vs. 3.8%; p<0.0001, 16.0% vs. 3.8%; p=0.0059 and 18.1% vs. 3.8%; p=0.0002, respectively).

5.2.1.1.8. KIR3DL1(+/-)KIR3DS1(+)HLA-A/B Bw4(+)

The frequency of individuals carrying the protective *KIR-HLA ligand* combination *3DL1(+/-)3DS1(+)HLA-A/B Bw4(+),* defined as the presence/absence of *KIR3DL1* and the presence of *KIR3DS1* as well as the presence of *HLA-A* and/or *HLA-B Bw4* ligands, was determined by analysis of *KIR* genotypes and *HLA ligand* genotypes available for the SAI, SAM, SAC and SAB populations (**Figure 5.1**). The frequency of *3DL1(+/-)3DS1(+)HLA-A/B Bw4(+)* genotypes was highest in the SAI population (50.0%), lowest in the SAB population (5.7%) and present at similar frequencies in the SAM and SAC populations (26.0% and 27.7%, respectively). A statistically significant difference in the frequency of *3DL1(+/-)3DS1(+)HLA-A/B Bw4(+)* genotypes was observed in the SAI population compared to the SAM, SAC and SAB populations (50.0% vs. 26.0%; p=0.0228, 50.0% vs. 27.7%; p=0.0103 and 50.0% vs. 5.7%; p<0.0001, respectively) as well as in the SAM and SAC populations compared to the SAB population (26.0% vs. 5.7%; p=0.0002 and 27.7% vs. 5.7%; p<0.0001, respectively).

5.2.1.2. Genotypes associated with susceptibility to HIV-1 infection

5.2.1.2.1. HLA-B p2-M homozygosity

Taken from **Chapter 3** and shown in **Figure 5.1**, the lowest frequency of the HLA-B p2-M/M genotype was observed in the SAM population (4.0%) and was present at similar frequencies in the SAC and SAB populations (8.5% and 8.8%, respectively). This genotype was absent in the SAI population.

5.2.1.2.2. HLA-ABC

Based on published studies of *HLA class I* alleles associated with resistance or susceptibility to HIV-1 infection (**Appendix B**) and *HLA class I* genotypic data available for the SAI, SAM, SAC and SAB populations, an individual was considered a carrier of a "susceptibility" *HLA-ABC* genotype if after assigning each *HLA-A*, *-B* and *-C* allele to a protective, deleterious or unknown status, the overall cumulative result was an excess of deleterious alleles (**Figure 5.1**). Individuals were defined as neutral if equal numbers of protective and deleterious alleles were present or if all six alleles were of unknown status.

Based on this analysis, the frequency of susceptibility *HLA-ABC* genotypes was highest in the SAB population (57.9%), followed by the SAM population (46.0%) and the SAC population (40.4%). The lowest frequency was determined in the SAI population (32.0%). A significant difference was observed in the frequency of susceptibility *HLA-ABC* genotypes in the SAB population compared to the SAI and SAC populations (57.9% vs. 32.0%; p=0.0019 and 57.9% vs. 40.4%; p=0.0092, respectively).

5.2.1.2.3. KIR2DL3(+)KIR2DL2(-)HLA-C1(+)HLA-C2(+)

The frequency of individuals carrying the deleterious *KIR-HLA ligand* combination *2DL3(+)2DL2(-)C1(+)C2(+),* defined as the presence of *KIR2DL3* and absent *KIR2DL2* as well as the presence of both *HLA-C1* and *HLA-C2* ligands, was determined by analysis of *KIR* genotypes and *HLA ligand* genotypes available for the SAI, SAM, SAC and SAB populations (**Figure 5.1**). The frequency of the *2DL3(+)2DL2(-)C1(+)C2(+)* genotype was highest in the SAC population (22.3%), lowest in the SAI population (8.0%) and present at similar frequencies in the SAM and SAB populations (16.0% and 12.6%, respectively). A significant difference was observed in the frequency of the *2DL3(+)2DL2(-)C1(+)C2(+)* genotype in the SAC population compared to the SAI population (22.3% vs. 8.0%, respectively; p=0.0372).

5.2.1.2.4. KIR3DL1(+)KIR3DS1(-)HLA-A/B Bw4(+)

The frequency of individuals carrying the deleterious *KIR-HLA ligand* combination *3DL1(+)3DS1(-)HLA-A/B Bw4(+),* defined as the presence of *KIR3DL1,* absence of *KIR3DS1* and the presence of *HLA-A* and/or *HLA-B Bw4* ligands, was determined by analysis of *KIR* genotypes and *HLA ligand* genotypes available for the SAI, SAM, SAC and SAB populations (**Figure 5.1**). The frequency of the *3DL1(+)3DS1(-)HLA-A/B Bw4(+)* genotype was highest in the SAB population (67.3%), followed by the SAM population (52.0%) and the SAC population (44.7%). The frequency was lowest in the SAI population (40.0%). A significant difference in the frequency of the *3DL1(+)3DS1(-)HLA-A/B Bw4(+)* genotype was observed in the SAB population compared to the SAI and SAC populations (67.3% vs. 40.0%; p=0.0008 and 67.3% vs. 44.7%; p=0.0006, respectively).

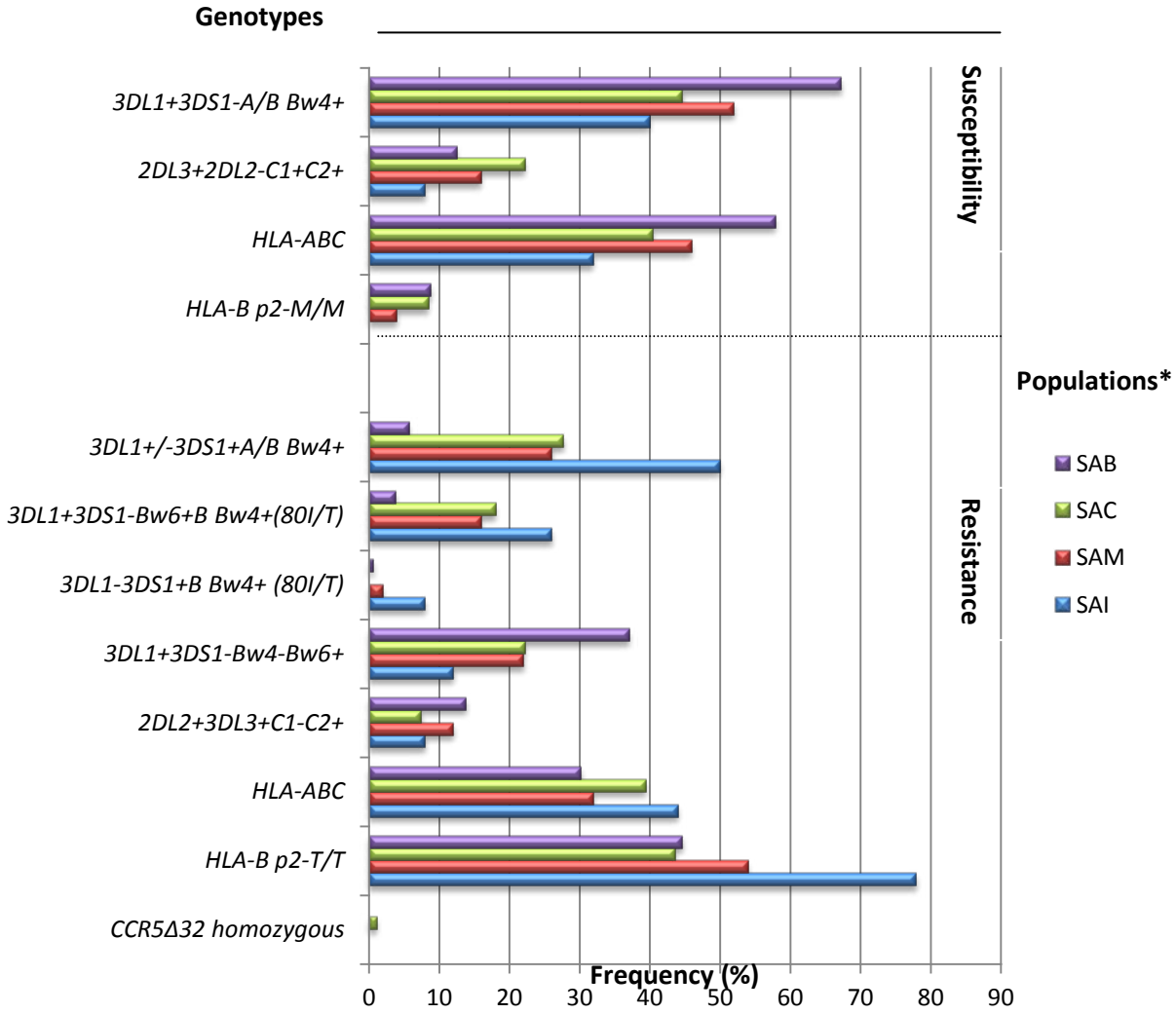

Figure 5.1. Bar graph showing frequency (%) of individuals of different populations carrying genotypes associated with HIV-1 resistance/susceptibility determined for SAI, South African Indian (n=50), SAM, South African Mixed ancestry (n=50), SAC, South African Caucasian (n=94) and SAB, South African Black (n=159) populations. *HLA*, human leukocyte antigen, M, methionine, T, threonine.

***Inter-population genotype frequency comparisons showing significance ($p<0.05$)**:

Resistance:			**Susceptibility:**		
HLA-B p2-T/T	SAI vs. SAM	p=0.0196	*HLA-B p2-M/M*	SAB vs. SAI	p=0.0248
	SAI vs. SAC	p<0.0001	*HLA-ABC*	SAB vs. SAI	p=0.0019
	SAI vs. SAB	p<0.0001		SAB vs. SAC	p=0.0092
3DL1+3DS1-Bw4-Bw6+	SAI vs. SAB	p=0.0007	*2DL3+2DL2-C1+C2+*	SAC vs. SAI	p=0.0372
	SAC vs. SAB	p=0.0173	*3DL1+3DS1-A/B Bw4+*	SAB vs. SAI	p=0.0008
3DL1-3DS1+B Bw4$^{80I/T}$+	SAI vs. SAB	p=0.0122		SAB vs. SAC	p=0.0006
3DL1+3DS1+B Bw4$^{80I/T}$+Bw6+	SAI vs. SAB	p<0.0001			
	SAM vs. SAB	p=0.0059			
	SAC vs. SAB	p=0.0002			
3DL1+/-3DS1+A/B Bw4+	SAI vs. SAM	p=0.0228			
	SAI vs. SAC	p=0.0103			
	SAI vs. SAB	p<0.0001			
	SAM vs. SAB	p=0.0002			
	SAC vs. SAB	p<0.0001			

5.2.1.3. Cumulative analysis of genotypes associated HIV-1 acquisition

The frequencies of each genotype associated with resistance or susceptibility to HIV-1 infection discussed in **Sections 5.2.1.1** and **5.2.1.2** were calculated at a population level. Since each of these genotypes can be present in a variable combination at an individual level, a cumulative analysis to determine an overall "resistant" or "susceptible" genotype profile was performed by scoring the number of genotypes associated with resistance and susceptibility present in each individual (**Figure 5.2**). An individual was considered "resistant" if an excess number of resistance genotypes were present or "susceptible" if an excess number of susceptibility genotypes were present. An individual was considered "neutral" if equal numbers of resistance and susceptibility genotypes were present or if no resistance or susceptibility genotypes were present at all. Of note, due to the overlap of some of the resistance genotypes involving *KIR3DL1/S1* and *HLA-Bw4* ligands (**Sections 5.2.1.1.5** to **5.2.1.1.8**), an individual was only scored once even if more than one of these genotypes was present.

Based on this analysis, the frequency of individuals carrying an overall resistant genotype profile was highest in the SAI population (68.0%), lowest in the SAB population (31.4%) and present at similar frequencies in the SAM and SAC populations (42.0% and 42.6%, respectively). A significant difference in the frequency of overall resistant genotype profiles was observed in the SAI population compared to the SAM, SAC and SAB populations (68.0% vs. 42.0%; $p=0.0154$, 68.0% vs. 42.6%; $p=0.0049$ and 68.0% vs. 31.4%; $p<0.0001$, respectively). Similarly, the frequency of individuals carrying an overall susceptible genotype profile was highest in the SAB population (42.8%), lowest in the SAI population (14.0%) and present at similar frequencies in the SAM and SAC populations (34.0% and 31.9%, respectively). A significant difference in the frequency of overall susceptible genotype profiles was observed in the SAM, SAC and SAB populations compared to the SAI population (34.0% vs. 14.0%; $p=0.0338$, 31.9% vs. $p=0.0267$ and 42.8% vs. 14.0%; $p=0.0002$, respectively). No significant differences were observed in the frequencies of individuals classified as neutral.

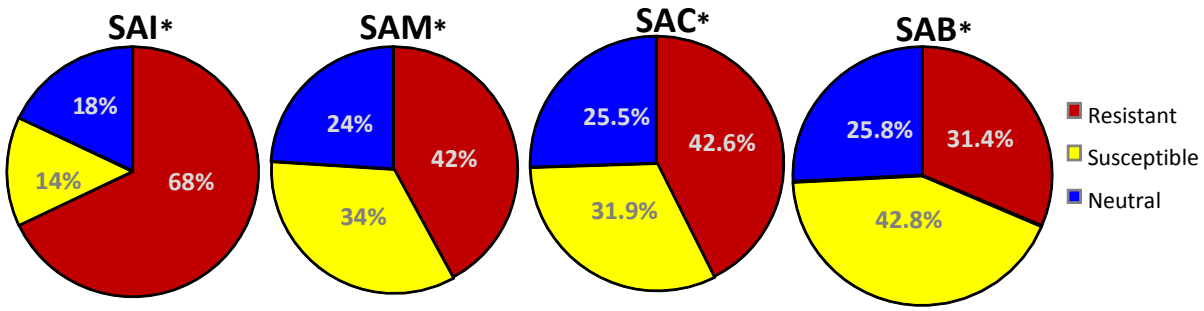

Figure 5.2. Pie charts showing frequency (%) of individuals carrying overall resistant, susceptible or neutral genotype profiles associated with HIV-1 acquisition determined for SAI, South African Indian (n=50), SAM, South African Mixed ancestry (n=50), SAC, South African Caucasian (n=94) and SAB, South African Black (n=159) populations.

***Inter-population overall genotype frequency comparisons showing significance ($p<0.05$)**:

Resistant:		**Susceptible:**	
SAI vs. SAM	p=0.0154	SAM vs. SAI	p=0.0338
SAI vs. SAC	p=0.0049	SAC vs. SAI	p=0.0267
SAI vs. SAB	p<0.0001	SAB vs. SAI	p=0.0002

Neutral: No significant differences

5.2.2. Genotypes associated with HIV-1 VLS

5.2.2.1. Genotypes associated with low HIV-1 VLS

5.2.2.1.1. HLA-ABC

Based on published studies of *HLA class I* alleles associated with VLS (**Appendix C**) and *HLA class I* genotypic data available for the SAI, SAM, SAC and SAB populations, an individual was considered a carrier of an *HLA-ABC* genotype profile associated with a "low" VLS if after assigning each *HLA-A*, *-B* and *-C* allele to a protective, deleterious or unknown status, the overall cumulative result was an excess of protective alleles (**Figure 5.3**). Individuals were defined as "neutral" if equal numbers of protective and deleterious *HLA-ABC* alleles were present or if all six *HLA class I* alleles were of unknown status. Based on this analysis, the frequency of *HLA-ABC* genotypes associated with low VLS was highest in the SAI population (66.0%), followed by the SAM population (44.0%) and present at similar frequencies in the SAC and SAB populations (39.4% vs. 40.3%, respectively). A significant difference in the frequency of *HLA-ABC* genotypes associated with low VLS was observed in the SAI population compared to the SAM, SAC and SAB populations (66.0% vs. 44.0%, p=0.0439; 66.0% vs. 39.4%, p=0.0029 and 66.0% vs. 40.3%, p=0.0019, respectively).

5.2.2.1.2. HLA-C +263 Del homozygosity

Taken from **Chapter 3** and shown in **Figure 5.3**, the highest frequency of *HLA-C +263 Del/Del* genotypes was present in the SAM population (20.0%) and similar frequencies were observed in the SAI, SAC and SAB populations (16.0%, 14.9% and 15.7%, respectively). No significant differences were observed between the populations.

5.2.2.1.3. CCR5Δ32(+)

Taken from **Chapter 3** and shown in **Figure 5.3**, the highest frequency of *CCR5Δ32* carriers was detected in the SAC population (18.1%) and low frequencies were present in the SAI and SAM populations (4.0% and 6.0%, respectively). The frequency of *CCR5Δ32* carriers in the SAB population was assumed to be 0%, due to the absence of this gene variant in this population (Picton et al., 2010). A significant difference in the frequency of *CCR5Δ32* carriers was observed in the SAC population compared to the SAI population (18.1% vs. 4.0%, respectively; p=0.0205).

5.2.2.1.4. KIR3DL1(+)KIR3DS1(-)HLA-B Bw4^{80I}(+)

The frequency of individuals carrying the protective *KIR-HLA ligand* combination *3DL1(+)3DS1(-)HLA-B Bw4^{80I}(+),* defined as the presence of *KIR3DL1,* the absence of *KIR3DS1* and the presence of *HLA-B Bw4^{80I}* ligands, was determined by analysis of *KIR* genotypes and *HLA ligand* genotypes available for the SAI, SAM, SAC and SAB populations (**Figure 5.3**). The frequency of the *3DL1(+)3DS1(-)HLA-B Bw4^{80I}(+)* genotype was highest in the SAB population (47.8%), followed by the SAM population (28.0%) and the SAI population (22.0%). The lowest frequency was present in the SAC population (17.0%). Significant differences in the frequency of the *3DL1(+)3DS1(-)HLA-B Bw4^{80I}(+)* genotype were observed in the SAB population compared to the SAI, SAM and SAC populations (47.8% vs. 22.0%, p=0.0016; 47.8% vs. 28.0%, p=0.0145 and 47.8% vs. 17.0%, p<0.0001, respectively).

5.2.2.2. Genotypes associated with high HIV-1 VLS

5.2.2.2.1. HLA-ABC

Based on published studies of *HLA class I* alleles associated with high or low VLS following HIV-1 infection (**Appendix C**) and *HLA class I* genotypic data available for the SAI, SAM, SAC and SAB populations, an individual was considered a carrier of an *HLA-ABC* genotype profile associated with a "high" VLS if after assigning each *HLA-A*, *-B* and *-C* allele to a protective, deleterious or unknown status, the overall cumulative result was an excess of protective alleles (**Figure 5.3**). Individuals were defined as neutral if equal numbers of protective and deleterious alleles were present or if all six alleles were of unknown status. Based on this analysis, the frequency of *HLA-ABC* genotypes associated with high VLS was the highest in the SAC population (33.0%), followed by the SAB population (29.6%) and the SAM population (24.0%). The lowest frequency was present in the SAI population (16.0%). A significant difference in the frequency of *HLA-ABC* genotypes associated with high VLS was observed in the SAC population compared to the SAI population (33.0% vs. 16.0%, respectively; p=0.0316).

5.2.2.2.2. HLA-C 3'UTR +263 Ins homozygosity

Taken from **Chapter 3** and shown in **Figure 5.3**, the highest frequency of *HLA-C 3'UTR +263 Ins/Ins* genotypes was detected in the SAC population (52.1%), followed by the SAB population (44.0%) and the SAM population (36.0%). The lowest frequency was present in the SAI population (24.0%). Significant differences in the frequency of *HLA-C 3'UTR +263 Ins/Ins* genotypes were observed in the SAC and SAB populations in comparison to the SAI population (52.1% vs. 24.0%; p=0.0014 and 44.0% vs. 24.0%; p=0.0127, respectively). Of note, *HLA-C 3'UTR* +263 Ins/Del heterozygotes, shown to have an intermediate effect on VLS, were present in 60.0%, 44.0%, 33.0% and 40.3% of SAI, SAM, SAC and SAB populations, respectively, but were not included in the cumulative analysis (**Section 5.2.2.3)**

5.2.2.2.3. KIR3DL1(+)KIR3DS1(+/-)HLA-B Bw4^{80I}(-)

The frequency of individuals carrying the deleterious *KIR-HLA ligand* combination *3DL1(+)3DS1(+/-)HLA-B Bw4^{80I}(-),* defined as the presence of *KIR3DL1*, the presence/absence of *KIR3DS1* as well as the absence of *HLA-B Bw4^{80I}* ligands, was determined by analysis of *KIR* genotypes and *HLA ligand* genotypes available for the SAI, SAM, SAC and SAB populations (**Figure 5.3**). The frequency of the *3DL1(+)3DS1(+/-)HLA-B Bw4^{80I}(-)* genotype was highest in the SAC population (68.1%) and similar frequencies were present in the SAI, SAM and SAB populations (50.0%, 54.0% and 46.5%, respectively). Significant differences in the frequency of the *3DL1(+)3DS1(+/-)HLA-B Bw4^{80I}(-)* genotype were observed in the SAC population compared to the SAI and SAB populations (68.1% vs. 50.0%; p=0.0471 and 68.1% vs. 46.5%; p=0.0011, respectively).

5.2.2.2.4. KIR3DL1(+)KIR3DS1(-)HLA-B Bw4^{80I}(-)

The frequency of individuals carrying the deleterious *KIR-HLA ligand* combination *3DL1(+)3DS1(-)HLA-B Bw4^{80I}(-),* defined as the presence of *KIR3DL1,* the absence of *KIR3DS1* and the absence of *HLA-B Bw4^{80I}* ligands, was determined by analysis of *KIR* genotypes and *HLA ligand* genotypes available for the SAI, SAM, SAC and SAB populations (**Figure 5.3**). The frequency of the *3DL1(+)3DS1(-)HLA-B Bw4^{80I}(-)* genotype was highest in the SAB population (45.3%), followed by the SAC population (42.6%) and the SAM population (36.0%). The lowest frequency was present in the SAI population (22.0%). Significant differences in the frequency of the *3DL1(+)3DS1(-)HLA-B Bw4^{80I}(-)* genotype were observed in both the SAC and SAB populations in comparison to the SAI population (42.6% vs. 22.0%; p= 0.0173 and 45.3% vs. 22.0%; p= 0.0045, respectively).

5.2.2.2.5. KIR2DS4f(+)

Taken from **Chapter 3** and shown in **Figure 5.3**, the frequency of *KIR2DS4f(+)* individuals was highest in the SAB population (84.9%), followed by the SAM population (66.0%) and the SAC population (36.2%). The lowest frequency was detected in the SAI population (28.0%). Significant differences in the frequency of *KIR2DS4f(+)* individuals were observed in both the SAB and SAM populations in comparison to the SAI population

(84.9% vs. 28.0%; p=0.0003 and 66.0% vs. 28.0%; p<0.0001, respectively). Significant differences were also observed in the SAM population compared to the SAC population (66.0% vs. 36.2%, respectively; p=0.0008) as well as in the SAB population in comparison to both the SAM and SAC populations (84.9% vs. 66.0%; p=0.0070 and 84.9% vs. 36.2%; p<0.0001; respectively).

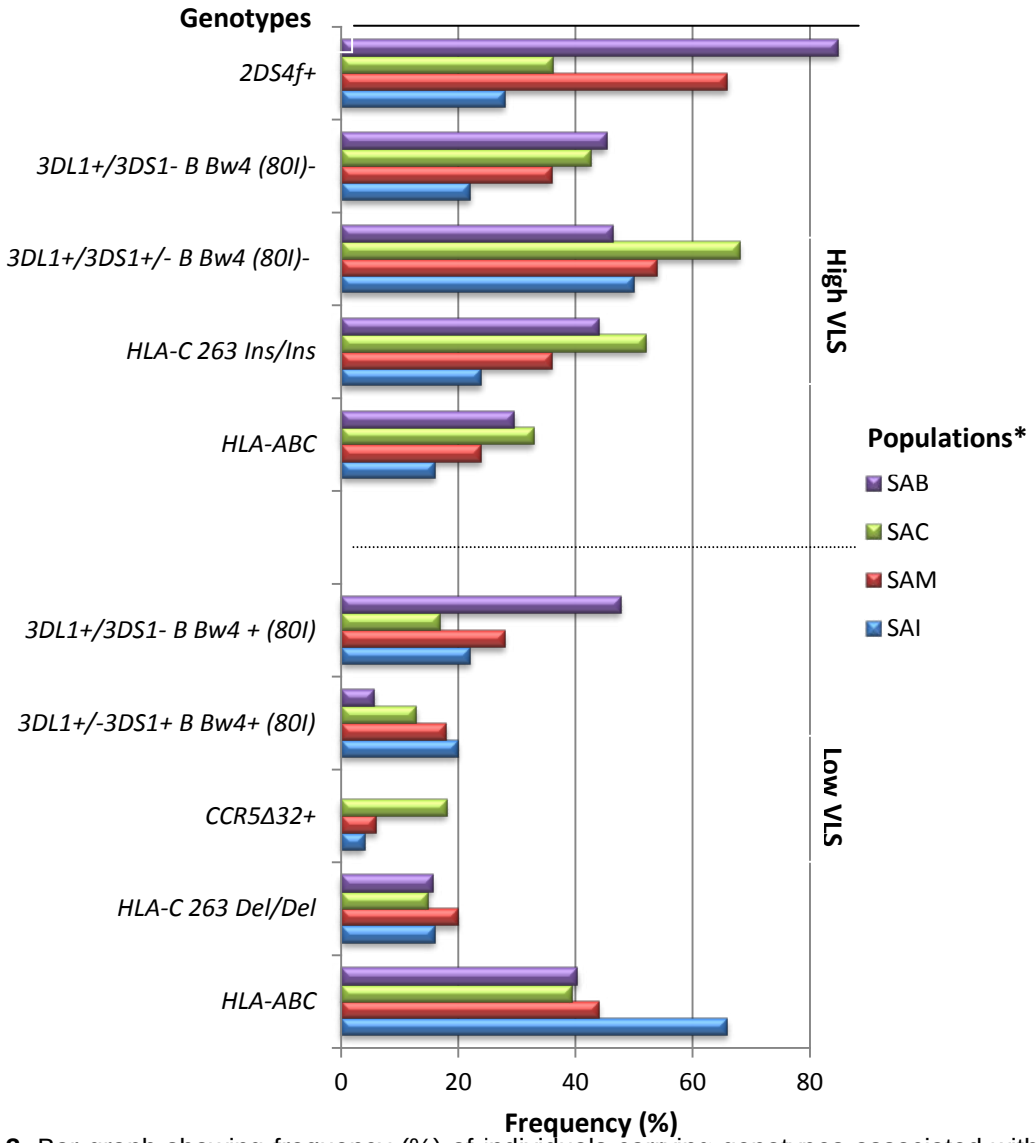

Figure 5.3. Bar graph showing frequency (%) of individuals carrying genotypes associated with high or low HIV-1 viral load setpoint (VLS) determined for SAI, South African Indian (n=50), SAM, South African Mixed ancestry (n=50), SAC, South African Caucasian (n=94) and SAB, South African Black (n=159) populations. *HLA*, human leukocyte antigen; *Del*, deletion; *Ins*, insertion.

***Inter-population genotype frequency comparisons showing significance ($p<0.05$):**

Low VLS:			**High VLS:**		
HLA-ABC	SAI vs. SAM	p=0.0439	*HLA-ABC*	SAC vs. SAI	p=0.0316
	SAI vs. SAC	p=0.0029	*HLA-C +263 Ins/Ins*	SAC vs. SAI	p=0.0014
	SAI vs. SAB	p=0.0019		SAB vs. SAI	p=0.0127
CCR5Δ32+	SAC vs. SAI	p=0.0205	*3DL1+3DS1+/-B Bw4^{80I}-*	SAC vs. SAI	p=0.0471
3DL1+/-3DS1+B Bw4^{80I}+	SAI vs. SAB	p=0.0192		SAC vs. SAB	p=0.0011
	SAM vs. SAB	p= 0.0166	*3DL1+3DS1-B Bw4^{80I}-*	SAC vs. SAI	p= 0.0173
3DL1+3DS1-B Bw4^{80I}+	SAB vs. SAI	p=0.0016		SAB vs. SAI	p= 0.0045
	SAB vs. SAM	p=0.0145	*2DS4f+*	SAM vs. SAI	p=0.0003
	SAB vs. SAC	p<0.0001		SAB vs. SAI	p<0.0001
				SAM vs. SAC	p=0.0008
				SAB vs. SAM	p=0.0070
				SAB vs. SAC	p<0.0001

5.2.2.3. Cumulative analysis of genotypes associated with HIV-1 VLS

The frequencies of each genotype associated with high or low HIV-1 VLS discussed in **Sections 5.2.2.1** and **5.2.2.2** were calculated at a population level. Since each of these genotypes can be present in a variable combination at an individual level, a cumulative analysis to determine an overall genotype profile was performed by scoring the number of genotypes associated with high and low VLS present in each individual (**Figure 5.4**). An individual was considered "low VLS" if an excess number of genotypes associated with low VLS were present, "high VLS" if an excess number of genotypes associated with high VLS were present and "neutral" if equal numbers of resistance and susceptibility genotypes were present or if no resistance or susceptibility genotypes were present at all.

Based on this analysis, the frequency of individuals carrying an overall low VLS genotype profile was highest in the SAI population (48.0%), lowest in the SAB population (13.2%) and present at similar frequencies in the SAM and SAC populations (22.0% and 22.3%, respectively). A significant difference in the frequency of overall resistant genotype profiles was observed in the SAI population compared to the SAM, SAC and SAB populations (48.0% vs. 22.0%; $p=0.0113$, 48.0% vs. 22.3%; $p=0.0024$ and 48.0% vs. 13.2%; $p<0.0001$, respectively). Similarly, the frequency of individuals carrying an overall high VLS genotype profile was lowest in the SAI population (30.0%) and present at similar frequencies in the SAM, SAC and SAB populations (56.0%, 60.6% and 59.7%, respectively). Significant differences in the frequency of overall high VLS genotype profiles were observed in the SAM, SAC and SAB populations compared to the SAI population (56.0% vs. 30.0%; $p=0.0149$, 60.6% vs. 30.0%; $p= 0.0008$ and 59.7% vs. 30.0%; $p= 0.0003$, respectively). No significant differences were observed in the frequencies of individuals classified as neutral.

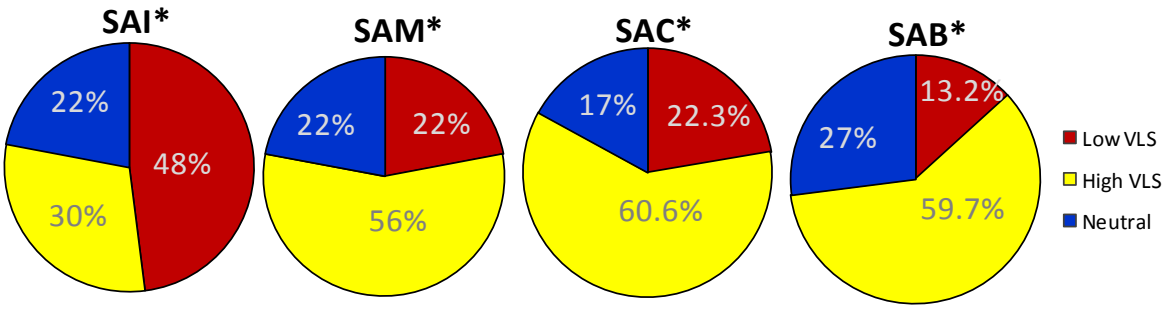

Figure 5.4. Pie charts showing frequency (%) of individuals carrying overall HIV-1 low VLS, high VLS or neutral genotype profiles determined for SAI, South African Indian (n=50), SAM, South African Mixed ancestry (n=50), SAC, South African Caucasian (n=94) and SAB, South African Black (n=159) populations.

***Inter-population overall genotype frequency comparisons showing significance ($p<0.05$)**:

Low VLS:	SAI vs. SAM	p=0.0113	**High VLS:**	SAM vs. SAI	p=0.0149
	SAI vs. SAC	p=0.0024		SAC vs. SAI	p= 0.0008
	SAI vs. SAB	p<0.0001		SAB vs. SAI	p= 0.0003
Neutral:	No significant difference				

5.3. Discussion

A multiplicity of factors, not necessarily mutually exclusive, contribute to the degree of spread of HIV-1 within a population and can include sexual behavioural practices, marital status, condom usage, circumcision status, antiretroviral drug therapy, host genetics, virus genetics and the presence of other sexually transmitted infections. In 2008, HIV-1 prevalence estimates were highest in the SAB population (13.6%), followed by the SAC population (1.7%) and lowest in the SAM and SAI populations (0.3% and 0.3%, respectively) respectively (www.mrc.ac.za/pressreleases/2009/sanat.pdf). In 2012, HIV-1 prevalence estimates increased and remained highest for the SAB population (15.0%), also increased the SAM population and SAI population (3.1% and 0.8%, respectively), while a decrease in the HIV-1 prevalence was recorded in the SAC population (0.3%) (www.hsrc.ac.za/en/research-outputs/view/6871). It is known that host genes involved in antiviral immunity can impact on HIV-1 prevalence since several immunogenetic factors associate with risk of HIV-1 acquisition. Influences on risk of HIV-1 transmission can also be effected via modulation of VLS. By extrapolation from the differences in the prevalence of protective/deleterious genetic factors determined for healthy individuals from the four populations, a possible influence of host genes in modulation of the observed differential HIV-1 prevalence rates can be examined.

A comparative immunogenetic characterisation study has not been done in healthy individuals from these populations. In this study, genomic DNA from healthy individuals, self-identified as belonging to the SAI, SAM, SAC or SAB populations, was used to first characterise the background frequencies of several host genes important in antiviral immunity (reported in **Chapter 3**). A subsequent analysis at a population level, was done regarding the prevalence of singular protective/deleterious genetic factors associated HIV-1 acquisition. Although some genotypes associated with resistance were significantly more prevalent in some populations compared to others, such as HLA-B p2-T/T in the SAI population (78%), HLA-ABC profiles in the SAB population (57.9%), *3DL1(+)3DS1(-)Bw6(+)HLA-B Bw4*$^{80I/T}$*(+)* and *3DL1(+/-)3DS1(+)Bw6(+)HLA-A/B Bw4(+)* genotypes in the SAI population (26.0% and 50.0%, respectively) and similarly for genotypes associated with susceptibility such as *2DL3(+)2DL2(+)C1(+)C2(+)* genotypes in the SAC population (22.3%) and *3DL1(+)3DS1(-)HLA-A/B Bw4(+)* genotypes in the SAB populations (67.3%), it was not possible to draw any conclusions regarding an overall host genetic influence on HIV-1 acquisition risk for the high and low prevalence populations.

Since individuals carry a combination of protective/deleterious genetic factors, a secondary analysis was undertaken an individual level by scoring each participant based on the overall number of protective/deleterious genetic factors present. Results from the cumulative analysis showed that a predominance of protective genotypes associated with resistance was present in 68.0%, 42.0%, 42.6% and 31.4% of SAI, SAM, SAC and SAB populations, respectively, while a predominance of deleterious genotypes associated with susceptibility was present in 14.0%, 34.0%, 31.9% and 42.8% of SAI, SAM, SAC and SAB populations, respectively. A significant difference was found in the frequency of individuals carrying predominantly resistant genotypes in the SAI population compared to the SAB population (68.0% vs. 31.4%, respectively; $p<0.0001$). Similarly, a significant difference in the frequency of individuals carrying predominantly susceptible genotypes was found in the SAB population compared to the SAI population (42.8% vs. 14.0%, respectively; $p=0.0002$). This result suggests that the prevalence of immunogenetic factors associated with resistance significantly higher and the frequencies of immunogenetic factors associated with susceptibility significantly lower in healthy SAI populations compared to healthy SAB populations and by extrapolation the low HIV-1 prevalence in the SAI population (0.8%) compared to the high HIV-1 prevalence in the SAB population (15.0%) may influenced, in part, by differences in the prevalence of host genetic factors associated with HIV-1

acquisition. Although statistical significance was not reached, the analysis also suggested that healthy SAM and SAC populations also have a higher prevalence of immunogenetic factors associated with resistance compared to the SAB population and a lower prevalence of immunogenetic factors associated with susceptibility.

It was also reasoned that the risk of transmission of HIV-1 within a population can be influenced by host genes that determine the infectivity of the donor. Studies in serodiscordant couples have demonstrated that HIV-1 viral load is the primary mediator of heterosexual transmission (Quinn et al., 2000, Merino et al., 2011). This was also confirmed by immunogenetic studies of index partners of serodiscordant couples, where specific *HLA class I* alleles such as *HLA-B*27, -B*57* and *-C*18* that associated with low risk of transmission, and *HLA-B*35* and *-A*36* that associated with high risk of transmission (Gao et al., 2010, Tang et al., 2008), were in fact the major determinants of VLS in the donors.

An analysis at the population level was done regarding the prevalence of singular protective/deleterious genetic factors associated with VLS. Although some genotypes associated with low VLS were significantly different in some populations compared to others, such as *HLA-ABC* genotypes in the SAI population (66.0%), *CCR5Δ32* carriage in the SAC population (18.1%), *3DL1(+/-)3DS1(+)HLA-B* $Bw4^{80I}$*(+)* genotypes in the SAI population (20.0%) and *3DL1(+)3DS1(-)HLA-B* $Bw4^{80I}$*(+)* genotypes in the SAB population (47.8%) and similarly for genotypes associated with high VLS such as *HLA-ABC* genotypes in the SAI population (16.0%), *HLA-C 3'UTR +263* Ins/Ins in the SAC population (52.1%), *3DL1(+)3DS1(+/-)HLA-B* $Bw4^{80I}$*(-)* genotypes in the SAC population (68.1%), *3DL1(+)3DS1(-)HLA-B* $Bw4^{80I}$*(-)* in the SAB population (45.3%) and *KIR2DS4f* carriage in the SAB population (84.9%) it was also not possible to draw any conclusions regarding an overall host genetic influence on HIV-1 VLS for the high and low prevalence populations.

A similar cumulative analysis of host genetic factors associated with VLS at the individual level was performed and showed that a predominance of protective genotypes associated with low VLS resistance was present in 48.0%, 22.0%, 22.3% and 13.2% of SAI, SAM, SAC and SAB populations, respectively, while a predominance of deleterious genotypes associated with high VLS was present in 30.0%, 56.0%, 60.6% and 59.7% of SAI, SAM, SAC and SAB populations, respectively. The results of this analysis suggest that the frequencies of immunogenetic factors associated with a low VLS following HIV-1

acquisition are significantly higher and the frequencies of immunogenetic factors associated with a high VLS following HIV-1 acquisition are significantly lower in healthy SAI populations compared to healthy SAB populations. Extrapolation of this finding to a low HIV-1 prevalence SAI population (0.8%) and a high HIV-1 prevalence SAB population (15.0%) implies that differences in host genetic factors associated with the infectivity of the donor may play a role in HIV-1 prevalence. Although statistical significance was not reached, the analysis also suggested that healthy SAM and SAC populations also have higher frequencies of immunogenetic factors associated with lower VLS. However, the frequencies of immunogenetic factors associated with high VLS were similar in the SAM, SAC and SAB populations and not statistically significantly different.

Other studies have postulated that a low HIV-1 prevalence in SAI, SAM and SAC populations is primarily due to low frequencies of concurrent sexual partnering practices that are high in the SAB population (Kenyon et al., 2013, Maughan-Brown, 2013). Taken together, the analysis performed in this study suggests that a low prevalence HIV-1 population, such as the SAI population, is likely influenced by resistance to HIV-1 infection at exposure and at a secondary level, the infectivity of donor partners is likely to be reduced due to a low VLS. In a high prevalence HIV-1 population, such as the SAB population, higher frequencies of susceptible individuals and lower frequencies of resistant individuals may be present and at a secondary level; lower frequencies of donor partners with low VLS and higher frequencies of donor partners with high VLS promote HIV-1 transmission. These same observations may apply to the low HIV-1 prevalence SAM and SAC populations, more so in terms of HIV-1 acquisition than transmission, but the statistical analyses lacked significance. It is of interest that the HIV-1 prevalence in adults between 15 and 49 in the mainland Indian population is 0.3% (http://www.who.int/gho/countries/ind/en/), while the HIV-1 prevalence in the same age category in SA was estimated at 1.0%, 4.6%, 0.6%, and 22.7% in SAI, SAM, SAC and SAB populations, respectively (www.hsrc.ac.za/en/research-outputs/view/6871). It is possible that similar protective host genetic factors are also prevalent in mainland Indian populations.

Importantly, there are limitations in this form of analysis. Firstly, the genetic associations reported in studies included from different populations possibly infected with different HIV-1 subtypes may not necessarily have the same phenotypic outcome in South African populations where subtype C infections predominate. African genetic studies as well

as limited studies performed in populations from India were used as well studies from Caucasian, African Black and Asian populations that would be relevant to the SAM population due to genetic admixture. The cumulative analysis of protective/deleterious genotypes associated with HIV-1 acquisition and VLS was done with equal weighting assigned to each of the variables, which is not necessarily true at a biological level. The influence of other genes that were not investigated in this study cannot be excluded. The genes selected for study in this work was based on published work with specific associations to HIV-1 acquisition and VLS. The sample size of the SAI and SAM study participants was smaller than the number of individuals included from the SAC and SAB populations, which may have affected statistical analyses. Due to the large numbers of genetic characterisations done per individual it was not feasible to increase the number of individuals genotyped. The results of this study have also been generated in healthy individuals and extrapolated to a HIV-1 infected population. Future investigations should include an application of the methods used in HIV-1 infected individuals, although cohorts in the SAI, SAM and SAC populations are more difficult to study due the low HIV-1 prevalence.

In conclusion, new information on the prevalence of individual host genetic factors related to HIV-1 acquisition and VLS has been generated for healthy SAI, SAM, SAC and SAB populations. The hypothetical analysis of the cumulative effect of individual protective/deleterious host genetic factors suggests that differences in the frequencies of these genes in the different populations may be influencing the rates of spread of HIV-1 within each population as indicated by the differences in HIV-1 prevalence. These findings have generated a hypothesis that requires further testing in HIV-1 infected cohorts to confirm the speculations based on genotyping of healthy individuals.

Chapter 6:
The cumulative effect of mutations/insertions in Gag contributes to disease progression in *HLA-B*58:01* and *HLA-B*58:02*-positive South African Black individuals with subtype C HIV-1 infection

6.1. Introduction

Immunogenetic studies in HIV-1 subtype C infected Black South African individuals have established that particular *HLA class I* alleles, such as *HLA-B*57*, *-B*58:01*, *-B*81:01*, *-A*02:05*, *-C*04:01* and *-B*42:01* are significantly associated with lower viral loads, higher CD4 counts and slower disease progression, whereas the opposite is true for deleterious alleles such as *HLA-B*45:01*, *-B*58:02*, *-B*18* and *-C*06:02* (Kiepiela et al., 2004). Of note, HLA-B*58:01 and -B*58:02 molecules are structurally similar and are differentiated by three amino acid changes at positions 94, 95 and 97, which map to the peptide binding cleft. Both receptors are grouped within the *HLA-B58* Supertype (Sidney et al., 2008) and are predicted to bind similar peptide antigens. According to the Los Alamos HIV immunology database, the peptide binding motifs are X-(ATSG)-XXXXXX-(WFY) and X-(ST)-XXX-(R)-XX-(F) for HLA-B*58:01 and -B*58:02, respectively (http://www.hiv.lanl.gov). In HIV-1 disease, it remains unclear why these receptors would associate with opposite clinical outcomes. It seems plausible that the amino acid differences in HLA-B*58:02 alter the peptide specificity and prevent efficient recognition of crucial epitopes necessary for detection and killing of infected cells, which are detectable by HLA-B*58:01. This phenomenon has been observed in *HLA-B*42:01* carriers who efficiently target $p24^{Gag}$ epitopes and have lower viral loads, in contrast to *HLA-B*42:02* carriers who have higher viral loads and do not respond to $p24^{Gag}$ epitopes (Kloverpris et al., 2012a). Notably, a single amino acid differentiates these two proteins.

Previous work has shown that CTL responses directed towards conserved proteins, such as Gag, are associated with lower viral loads in comparison to CTL responses to epitopes derived from variable proteins such as Env (Kiepiela et al., 2007, Wang et al., 2009, Masemola et al., 2004a). Following on from this, studies have demonstrated that Gag-specific HLA-B*58:01-restricted CTLs could be identified in individuals controlling HIV-1 replication, whereas, in non-controlling individuals, HLA-B*58:02-restricted CTLs were directed towards Env, and furthermore, selection pressure on viral epitopes was reduced in *HLA-B*58:02*, but not *HLA-B*58:01* carriers (Ngumbela et al., 2008). One explanation is the theoretical possibility that ineffective HLA-B*58:02-restricted CTL responses relate to thymic selection of naïve T cells. In an *HLA-B*57* (protective) and *-B*07:01* (deleterious) context, epitope prediction algorithms have been used to hypothesise that protective *HLA class I* alleles are defined by fewer events of self-peptide binding when compared to non-

protective alleles and would therefore permit the generation of a greater diversity of naïve T-cells in the thymus with increased probability of selecting HIV-1 specific T-cells (Kosmrlj et al., 2010). This could explain the greater breadth of anti-viral CTL responses and increased selective pressure on viral epitopes seen in individuals with protective *HLA class I* alleles. However, this remains to be proven experimentally, but if true, may also apply to the context of *HLA-B*58:01* and *-B*58:02*. Another possibility may relate to differences in effector function of CTLs restricted by HLA-B*58:01 and -B*58:02 that could impact on the breadth and immune selective pressures to HIV-1 epitopes. This was not confirmed in studies comparing protective HLA-B*58:01-restricted CTL responses to non-protective HLA-C*07-restricted CTL responses (Mkhwanazi et al., 2010). In their study the polyfunctionality measured by cytokine profiles in these two CTL sets was not significantly different; however, HLA-B might not be directly comparable to HLA-C in this context.

It was also observed in the SIV model that an increase in the number of Gag epitopes targeted correlated with lower viral load which is also evident from HIV-1 studies (Geldmacher et al., 2007). There are three HLA-B*58:01-restricted epitopes in Gag reported to be targeted in subtype C HIV-1 infected South African individuals (Kloverpris et al., 2012b), but no Gag epitopes are known to elicit HLA-B*58:02-restricted CTL responses. Putative HLA-B*58:02-restricted epitopes have been reported, but are unconfirmed (Spivak et al., 2010, Chopera et al., 2008). It is therefore likely that HLA-B*58:01-mediated immune pressure generates early protective CTL responses to incoming $p24^{Gag}$ and these responses remain sustained during the chronic stage of disease, whereas HLA-B*58:02 targets epitopes derived from late proteins expressed after integration such as Env, but fails to target Gag, resulting in poor immunity.

The reason why HLA-B*58:02 fails to target epitopes in Gag may simply be related to binding motif, but another possibility involves the accessibility of proteosomally derived peptides to the HLA class I molecule in the endoplasmic reticulum. Studies in HLA-B*57 and -B*58:01 have shown that although the KF11 epitope in $p24^{Gag}$ binds to both molecules, there is a failure in presentation of KF11 by HLA-B*58:01 (Kloverpris et al., 2013b), although an unconfirmed response to KF11 in a *HLA-B*58:01* carrier has been noted in a previous study (Chopera et al., 2011). The failure was delineated to an intracellular block in access of KF11 to HLA-B*58:01 molecules in the endoplasmic reticulum, since it was noted

that in individuals co-expressing HLA-B*58:01 and -B*57 molecules, KF11 was efficiently presented by HLA-B*57 but not by HLA-B*58:01.

It has been shown, subsequently, that the benefits attributed to protective *HLA class I* alleles, such as *HLA-B*57* and *-B*58:01*, can be due to fitness costs imposed on replicating virus when CTL escape mutations are introduced into conserved proteins (Crawford et al., 2007). This is further supported by the observation that escape mutations often revert to wild-type when virus is transmitted to *HLA class I* mismatched hosts (Leslie et al., 2004). Unfortunately, it has also been shown that the benefits of early escape mutations that reduce viral replicative capacities can be overcome by compensatory mutations that develop during the chronic stage of disease (Brockman et al., 2010, Gijsbers et al., 2013). This has been documented in *HLA-B*58:01* carriers with subtype C HIV-1 infection in South Africa (Chopera et al., 2011). Notably, over time compensatory mutations may become established in viruses circulating in the community as shown recently by a decline in the protective benefits of *HLA-B*57/58:01* alleles in Botswana compared to South Africa due to the longer duration of the epidemic in Botswana (Payne et al., 2014). Alternatively, viral replication can also be enhanced by mutations that improve the release of virions from infected cells through budding. This has been observed to occur via duplication of the PTAP amino-acid motif in $p6^{Gag}$ that enhances recruitment of cellular Tsg101 proteins that are required for efficient viral budding (VerPlank et al., 2001). There is also an alternative budding pathway that is dependent on a PYRE amino-acid motif in $p6^{Gag}$ that recruits ALG-2-interacting protein X (ALIX), normally absent in sub-Saharan subtype C viruses, however, in subtype C HIV-1 infected Indian individuals, similar 4-amino acid insertions were found to be more common in patients failing protease therapy and may represent adaptations of virus to restore fitness due to the development of protease resistance mutations (Neogi et al., 2014).

An important finding in individuals carrying protective *HLA class I* alleles is that immune pressure by other *HLA class I* alleles in the individual's repertoire, are co-operative or additive in the overall control of viral replication (Matthews et al., 2012, Leslie et al., 2010). Particular haplotypes, such as *HLA-C*04:01-B*81:01*, *-C*12:03-B*39:10*, *-A*74:01-B*53:03* associate with improved viral control compared to the singular alleles (Leslie et al., 2010). Similarly, slow disease progression based on CD4+ T cell count is reported to associate with haplotypes *HLA-A*74:01-B*15:03*, *-B*14:02-C*08:02* and *-A*74:01-B*15:03-C*02:02*, while accelerated disease progression associates with haplotypes *HLA-*

*A*30:02-B*45:01, -A*30:02-C*16:01, -B*53:01-C*04:01, -B*15:10-C*03:04* and *-B*58:01-C*03:02* (Sampathkumar et al., 2014).

Interestingly, an alternative consequence of viral escape mutations in $p24^{Gag}$ comes from evidence supporting the impact of Gag mutations on the ability of the intrinsic antiretroviral factor, Tripartite motif-containing protein 5 alpha (TRIM5α), to inactivate incoming viral cores (Granier et al., 2013). In addition, evidence of escape mutations in CTL epitopes that modulate NK cell recognition of infected target cells via killer-cell immunoglobulin-like receptor (KIR) interaction has come to light (Alter et al., 2011). The antiviral efficacy of NK cells are often underappreciated since NK cell responses precede CTL responses during early stages of acute HIV-1 infection (Alter et al., 2007) and KIRs in combination with cognate HLA class I ligands, such as KIR3DL1 and HLA-$Bw4^{80I}$, associate with reduced viral load (Jiang et al., 2013, Pelak et al., 2011, Martin et al., 2007). The synergistic or epistatic interaction between KIR3DS1 and HLA-$Bw4^{80I}$ also associates with low viral load (Pelak et al., 2011, Martin et al., 2002). *KIR3DS1* and *-3DL1* are similar in structure in the extracellular domains, however, functional studies have not shown direct interaction of KIR3DS1 with HLA-$Bw4^{80I}$ ligands (Carr et al., 2007, Gillespie et al., 2007, O'Connor et al., 2007). An alternative explanation may stem from a report that subsets of CTLs can also express KIR3DL1 and KIR3DS1, and KIR3DS1-positive CTLs are more effective in targeting HIV-infected target cells (Zipperlen et al., 2014). Unfortunately, HIV-1 attempts to evade NK cell detection by positive selection of mutations in HLA class I-restricted epitopes that affects KIR recognition. Mutations mapping to several viral proteins such as Gag, Vpr, Tat, Vpu, Env and Nef that mediate escape from NK cell recognition have been shown to influence NK cell activation governed by KIRs such as KIR2DS1, -2DS2, -2DS3, 2DS5, -2DL2, -2DL3 and -3DS1 (Alter et al., 2011). Since KIRs can either transduce activating or inhibitory signals, a plausible explanation for this finding is that mutations in HLA class I-restricted viral epitopes either strengthen inhibitory signals or dampen activation signals, thereby reducing NK cell activation potential. Other investigators have shown that sequence variation in the YL9 epitope in $p24^{Gag}$, restricted by HLA-C*03:04, can influence binding to KIR2DL2 inhibitory receptors (van Teijlingen et al., 2014) and similarly, HLA-B*57 and KIR3DL1 binding studies have shown the same effect of amino acid substitutions in the $p24^{Gag}$ TW10 epitope (Brackenridge et al., 2011).

In the current study, an extensive analysis of mutations in inferred Gag protein sequences from progressors with subtype C HIV-1 infection and carrying either *HLA-B*58:01* (n=5) or *-B*58:02* (n=13) was performed. The identification of immune escape mutations that contribute to a lack of control of HIV-1 replication, resulting in a $CD4^+$ T cell decline to <350 cells/ml, may give insight into the mechanisms of failure of protective *HLA-B*58:01* in sustaining effective virological control and a lack HLA-B*58:02-mediated control of viral replication. Plasma samples stored at baseline and 6 and 12 months after baseline sampling were used to extract viral RNA. The partial *gag* gene (nt 958-nt 2288, HXB2 numbering), spanning partial $p17^{Gag}$ ORF (codon 53) to the Gag stop codon, was sequenced. Mutations in deduced protein sequences were identified to investigate the short term evolution of Gag during late stage chronic HIV-1 infection. Gag sequences were examined with a specific focus on CTL escape mutations and compensatory mutations, as well as changes that may impact on NK cell surveillance or TRIM-5α antiretroviral activity.

6.2. Results

6.2.1. Study participant information

6.2.1.1. Analysis of $CD4^+$ T cell count, viral load and $CD4^+$ T cell decline

Measurements for $CD4^+$ T cell count, viral load (VL) and rate of $CD4^+$ T cell decline for the study participants as well as the availability of *gag* sequences are listed in **Table 6.1**. $CD4^+$ T cell counts and VL measurements at baseline, 6 months and 12 months were stratified according to *HLA-B*58:01* and *-B*58:02* carriage and compared (**Figure 6.1** and **Figure 6.2**). No significant differences were found ($p>0.05$). The rate of $CD4^+$ T cell decline was calculated using the earliest and last available $CD4^+$ T cell counts divided by the number of months. Results were stratified according to *HLA-B*58:01* and *-B*58:02* carriage were compared (**Figure 6.3**). Again, no significant differences were found ($p=0.554$). Since there were no significant differences in the $CD4^+$ T cell count or VL at baseline, 6 months or 12 months as well as in the rates of $CD4^+$ T cell decline for *HLA-B*58:01* and *-B*58:02* carriers, these two groups could be considered equivalent in terms of the stage of HIV-1 disease progression. Since *HLA-B*58:01* is a protective allele usually associated with lower VLs,

higher $CD4^+$ T cell counts and slower decline in $CD4^+$ T cell count, it was interesting that *HLA-B*58:01* carriers appeared clinically similar to the *HLA-B*58:02* carriers. However, the duration of infection is not known and this could mean that immune control has been lost in *HLA-B*58:01* carriers over a longer period of time. The possibility of viral mutation mediating escape from immune surveillance was therefore investigated in these two groups.

Table 6.1. Study participant information.

	Baseline			6 months			12 months			
Participant ID	$CD4^+$ T cell count	Log_{10} VL	*Gag* sequences	$CD4^+$ T cell count	Log_{10} VL	*Gag* sequences	$CD4^+$ T cell count	Log_{10} VL	*Gag* sequences	$CD4^+$ T cell decline*
1	208	4.17	Yes	165	3.99	Yes	-	-	-	-7.167
2	254	3.77	Yes	286	4.38	Yes	222	4.01	Yes	-2.667
3	268	4.48	Yes	-	-	Yes	243	4.83	Yes	-2.083
4	252	3.85	Yes	156	3.42	Yes	-	-	-	-16.0
5	241	3.70	Yes	203	4.07	-	200	4.02	Yes	-3.417
6	259	2.73	Yes	47	5.68	Yes	-	-	-	-35.33
7	226	3.84	Yes	154	4.30	Yes	-	-	-	-12.0
8	154	4.46	Yes	134	4.33	Yes	-	-	-	-3.333
9	243	4.04	Yes	162	4.01	Yes	-	-	-	-13.5
10	-	-	Yes	6	5.11	Yes	3	5.29	Yes	-0.5
11	239	4.69	Yes	133	3.90	Yes	-	-	-	-17.67
12	324	3.94	-	178	1.69	Yes	253	2.14	Yes	-5.917
13	246	4.23	Yes	216	3.85	Yes	91	3.66	Yes	-12.92
14	245	4.93	Yes	273	5.40	Yes	157	5.08	Yes	-7.333
15	207	3.09	Yes	191	3.39	Yes	200	3.73	Yes	-0.583
16	273	2.94	Yes	310	2.86	-	258	2.92	Yes	-1.25
17	307	2.41	-	247	3.10	Yes	246	3.51	Yes	-5.083
18	351	2.31	Yes	299	2.99	Yes	-	-	-	-8.667

$CD4^+$ T cell count measurements in cells/µl; VL, viral load measurements in copies/ml; "-", data not available; *HLA-B*58:01*-positive participants are shaded in grey; *$CD4^+$ T cell decline measured in cells/µl/month and calculated using the earliest and last available measurements for $CD4^+$ T cell count.

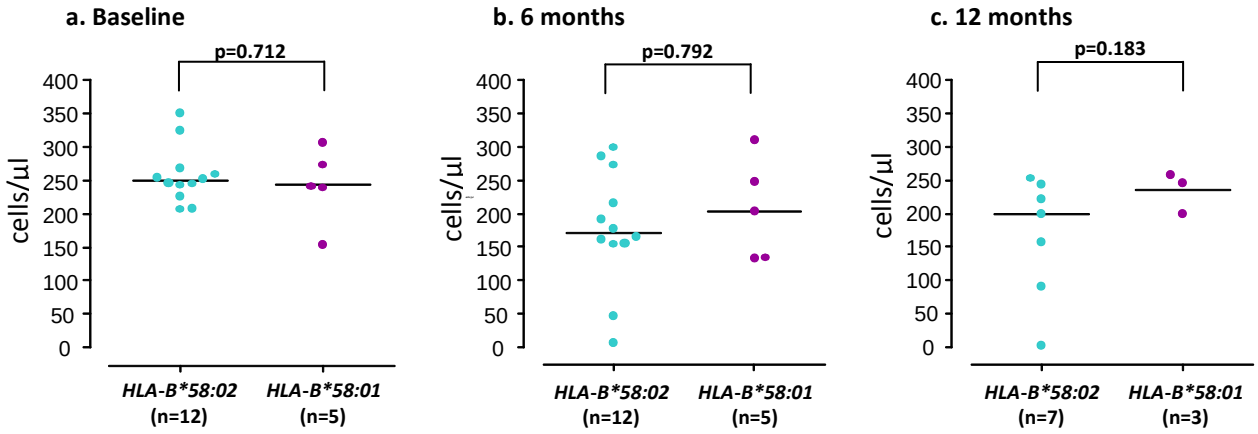

Figure 6.1. Comparison of $CD4^+$ T cell counts at (a) baseline, (b) 6 months and (c) 12 months for participants with either *HLA-B*58:01* or *HLA-B*58:02*.

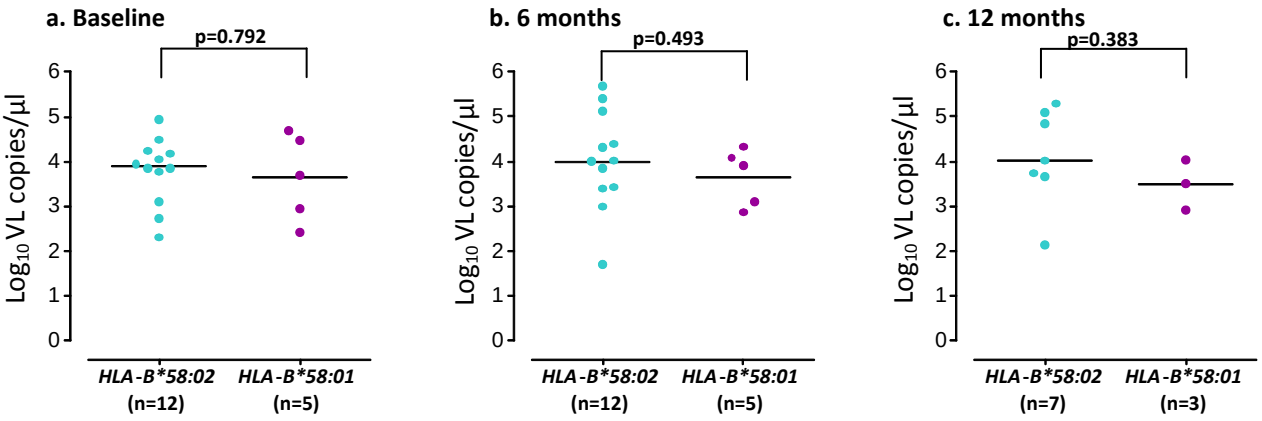

Figure 6.2. Comparison of Log_{10} viral load at (a) baseline, (b) 6 months and (c) 12 months for study participants with either *HLA-B*58:01* or *HLA-B*58:02*.

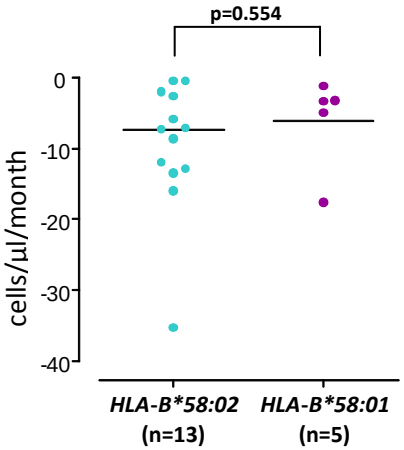

Figure 6.3. Comparison of $CD4^+$ T cell decline for study participants with either *HLA-B*58:01* or *HLA-B*58:02*. Rate of $CD4^+$ T cell decline was calculated as the difference between the earliest and last measurement for $CD4^+$ T cell count divided by the number of months.

6.2.1.2. Phylogenetic analysis of gag sequences

To confirm that *gag* sequences generated for each study participant were correctly assigned, a phylogenetic analysis was performed using alignments of sequences as input for phylogenetic tree construction using MEGA v5 software. A bootstrapped (n=1000) neighbour-joining tree was constructed using the Kimura 2-paramater model and showed that *gag* sequences generated for each study participant correctly clustered together (**Figure 6.4**).

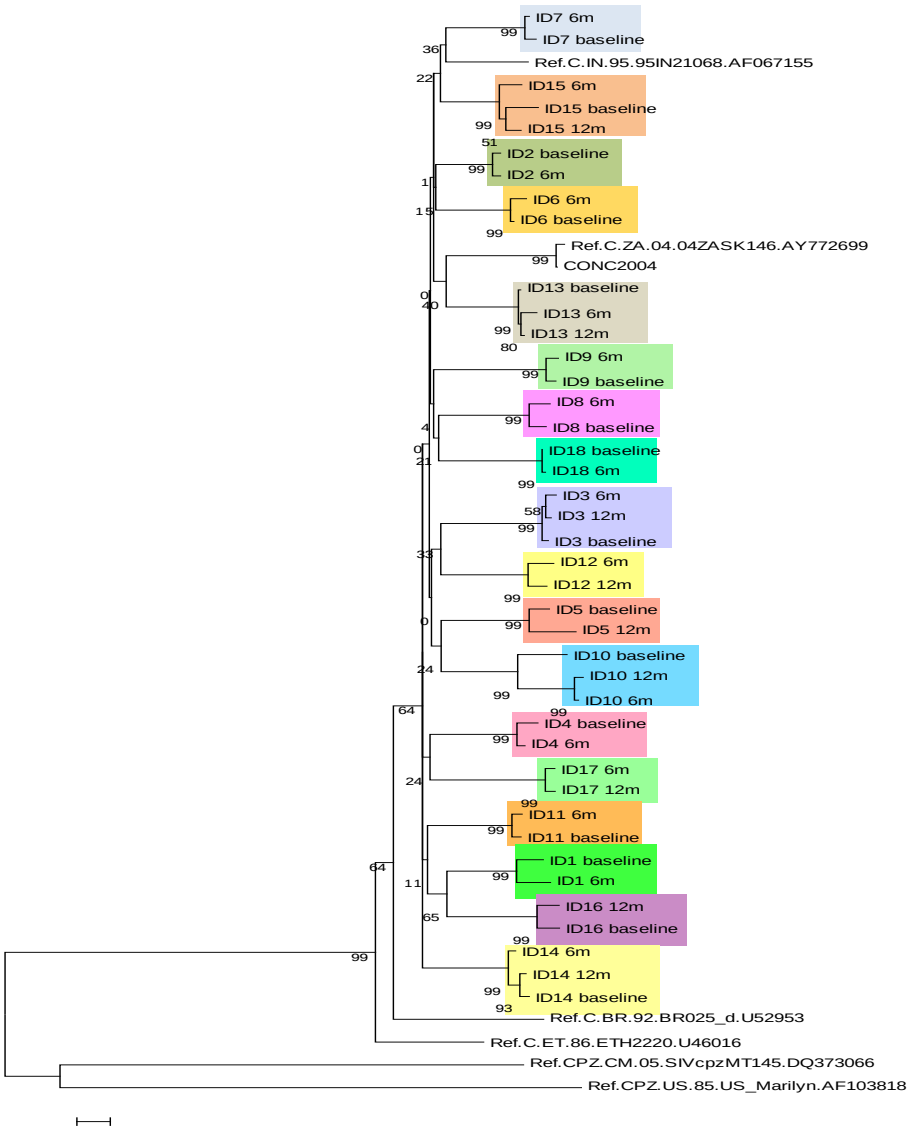

Figure 6.4. Neighbour-joining phylogenetic tree constructed using *gag* sequences generated for each of the study participants (sequences from individual participants are shown in shaded boxes). Subtype C HIV-1 reference sequences and SIV Chimpanzee outlier sequences were included. Numbers at the nodes indicate bootstrap scores (%).

6.2.2. Presence of protective/deleterious *HLA class I* alleles

An analysis of the distribution of known protective or deleterious *HLA class I* alleles associated with disease progression was performed based on high resolution *HLA class I -A*, *-B* and *-C* genotyping data generated by Roche 454 FLX deep sequencing methods (**Table 6.2**). A detailed analysis of singular *HLA class I* alleles at each locus is presented below, however in summary, the number of protective *HLA class I* alleles was 20 in the total of 30 alleles present (66.7%) in the 5 *HLA-B*58:01* carriers compared to 28 protective *HLA class I* alleles in a total of 76 alleles present (36.8%) in the 13 *HLA-B*58:02* carriers (two alleles were not counted due to unknown associations). Comparison of the two frequencies was significant (p=0.0149; OR=3.257[1.329-7.984]). Based on the additive effect of protective *HLA class* I alleles (Matthews et al., 2012, Leslie et al., 2010), it is unexpected that the *HLA-B*58:01* carriers would be at the same stage of disease progression as *HLA-B*58:02* carriers. The duration of infection is not known and the protective benefits of the additional alleles may have been lost, hence an analysis of viral mutations mediating escape from other HLA class I-restricted epitopes was undertaken.

6.2.2.1. Alleles at the HLA-A locus

HLA-A genotypes determined for each study participant, *HLA* Supertypes and allele frequencies reported for the SAB population are listed in **Table 6.2**. Twelve different *HLA-A* alleles were identified in the study population. *HLA-A*02:01,* found in 4 individuals, is reported to be a deleterious allele in South African and Japanese populations (Leslie et al., 2010, Naruto et al., 2012), while African studies report that *HLA-A*02:05,* also present in 4 individuals, is a protective allele (Novitsky et al., 2003, Leslie et al., 2010, Kiepiela et al., 2004, Saathoff et al., 2010). Of note, *HLA-A*02:05* is in high linkage disequilibrium with *HLA-B*58:01* (D'=0.73, p<0.001) (Paximadis et al., 2012) and therefore *HLA-B*58:01* carriers may often inherit an additional protective allele. Within the study participants, *HLA-A*02:05* was present in three *HLA-B*58:01* carriers and one *HLA-B*58:02* carrier. Two individuals carried *HLA-A*68:02* reported as deleterious in a South African study (Kiepiela et al., 2004), however, a study from Botswana found the opposite association (Novitsky et al., 2003). *HLA-A*68:02* is in low linkage disequilibrium (D'=0.33, p<0.001) with *HLA-B*15:10* (Paximadis et al., 2012) which has been reported to be a deleterious allele in African populations (Kiepiela et al., 2004, Leslie et al., 2010, Matthews et al., 2012). *HLA-A*03:01,*

present in 3 individuals, was found to be protective in Chinese populations (Zhang et al., 2013), but deleterious in African populations (Matthews et al., 2012, Carlson et al., 2012). *HLA-A*66:01*, present in 3 individuals and *-A*68:01*, present in 4 individuals, are both reported to be deleterious alleles in African populations (Leslie et al., 2010, Matthews et al., 2012, Kiepiela et al., 2004, Carlson et al., 2012). *HLA-A*03:01:01* is in moderate linkage disequilibrium with *HLA-B*58:02* (D'=0.46, p<0.001), while *HLA-A*68:01:01* is in high linkage disequilibrium with *HLA-B*58:02* (D'=0.67, p<0.001) (Paximadis et al., 2012) and therefore *HLA-B*58:02* carriers can often inherit a second deleterious allele. *HLA-A*80:01* present in one individual and *-A*36:01*, also present in one individual, have not been reported in any HIV-1 association studies. *HLA-A*30:02*, found in 5 individuals, (Carlson et al., 2012) and *-A*29:02*, also found in five individuals, (Tang et al., 2010) are protective in African populations. *HLA-A*30:01* was present in one study participant and is also a protective allele associated in African populations (Carlson et al., 2012). *HLA-A*23:01*, present in two study participants, is a protective allele in South African populations (Kiepiela et al., 2004).

Table 6.2. *HLA-A* alleles present in *HLA-B*58:01* and *-B*58:02*-positive progressors their association with HIV-1 disease progression, *HLA* Supertype, *HLA* ligand designation and allele frequencies known for the SAB population.

HLA-A										
ID	**Allele 1**	**P/D**	***HLA* Supertype**	***HLA* ligand**	**Allele frequency SAB (%)**	**Allele 2**	**P/D**	***HLA* Supertype**	***HLA* ligand**	**Allele frequency SAB (%)**
1	*A*02:01:01*	D	A02	non-Bw4	8.29	*A*66:01:01*	D	A03	non-Bw4	0.75
2	*A*02:05:01*	P	A02	non-Bw4	6.78	*A*30:02:01*	P	A01	non-Bw4	9.55
3	*A*23:01:01*	P	A24	$Bw4^{80I}$	8.04	*A*68:01:01*	D	A03	non-Bw4	3.27
4	*A*03:01:01*	D	A03	non-Bw4	6.53	*A*68:01:01*	D	A03	non-Bw4	3.27
5	*A*02:05:01*	P	A02	non-Bw4	6.78	*A*30:02:01*	P	A01	non-Bw4	9.55
7	*A*68:01:01*	D	A03	non-Bw4	3.27	*A*68:02:01*	D	A02	non-Bw4	8.54
8	*A*02:05:01*	P	A02	non-Bw4	6.78	*A*29:02:01*	P	A01 A24	non-Bw4	6.28
9	*A*03:01:01*	D	A03	non-Bw4	6.53	*A*30:02:01*	P	A01	non-Bw4	9.55
6	*A*23:01:01*	P	A24	$Bw4^{80I}$	8.04	*A*29:02:01*	P	A01 A24	non-Bw4	6.28
10	*A*03:01:01*	D	A03	non-Bw4	6.53	*A*36:01*	?	A01	non-Bw4	1.01
11	*A*29:02:01*	P	A01 A24	non-Bw4	6.28	*A*66:01:01*	D	A03	non-Bw4	0.75
12	*A*02:01:01*	D	A02	non-Bw4	8.29	*A*66:01:01*	D	A03	non-Bw4	0.75
13	*A*29:02:01*	P	A01 A24	non-Bw4	6.28	*A*30:02:01*	P	A01	non-Bw4	9.55
14	*A*29:02:01*	P	A01 A24	non-Bw4	6.28	*A*80:01*	?	A01	non-Bw4	0.50
15	*A*30:01:01*	P	A01 A03	non-Bw4	10.05	*A*30:02:01*	P	A01	non-Bw4	9.55
16	*A*02:01:01*	D	A02	non-Bw4	8.29	*A*68:02:01*	D	A02	non-Bw4	8.54
17	*A*02:05:01*	P	A02	non-Bw4	6.78	*A*23:01:01*	P	A24	$Bw4^{80I}$	8.04
18	*A*02:01:01*	D	A02	non-Bw4	8.29	*A*68:01:01*	D	A03	non-Bw4	3.27

HLA, human leukocyte antigen, P, protective, D, deleterious, "?", undetermined; *HLA-B*58:01*-positive participants shaded in grey.

6.2.2.2. Alleles at the HLA-B locus

HLA-B genotypes determined for each study participant, *HLA* Supertypes and allele frequencies reported for the SAB population are listed in **Table 6.3**. Ten different *HLA-B* alleles were identified in the study population. *HLA-B*58:01* (n=5) and *-B*58:02* (n=13), the subject of the current study, are both common alleles in the SAB population (Paximadis et al., 2012). Numerous studies in diverse populations, including Caucasian, African and Japanese individuals have associated *HLA-B*58:01* with low VL (Brockman et al., 2010, Carlson et al., 2012, Gao et al., 2010, Geldmacher et al., 2007, Kiepiela et al., 2004, Leslie et al., 2010, Matthews et al., 2012, Mlisana et al., 2014, Naruto et al., 2012, Saathoff et al., 2010), while *HLA-B*58:02,* an allele mostly found in African populations, associates with high VL (Carlson et al., 2012, Kiepiela et al., 2004, Leslie et al., 2010, Matthews et al., 2012, Mlisana et al., 2014, Saathoff et al., 2010). Other *HLA-B* alleles present in the study population included *HLA-B*08:01* (n=1) which is considered a deleterious allele in African populations (Leslie et al., 2010, Matthews et al., 2012, Kiepiela et al., 2004, Carlson et al., 2012). *HLA-B*15:03* (n=2) (Geldmacher et al., 2007, Leslie et al., 2010, Kiepiela et al., 2004, Saathoff et al., 2010, Frahm et al., 2006, Matthews et al., 2012) and *-B*15:10 (n=3)* (Kiepiela et al., 2004, Leslie et al., 2010, Matthews et al., 2012) are also considered deleterious alleles, although two studies showed the opposite association for these alleles (Mlisana et al., 2014, Novitsky et al., 2003). *HLA-B*39:10,* present in one individual, is a deleterious allele (Matthews et al., 2012, Tang et al., 2002), although one study reported the opposite finding (Leslie et al., 2010). *HLA-B*42:01,* found in 2 individuals, is considered protective (Carlson et al., 2012, Kiepiela et al., 2004, Leslie et al., 2010, Matthews et al., 2012, Saathoff et al., 2010), while *HLA-B*53:01,* also present in 2 individuals, was reported to be deleterious in African-American populations (Lazaryan et al., 2011). *HLA-B*44:03,* present in 4 individuals, is protective in African (Carlson et al., 2012, Kiepiela et al., 2004, Leslie et al., 2010, Matthews et al., 2012) and Asian (Naruto et al., 2012, Zhang et al., 2013) populations, while *-B*45:01,* found in 2 individuals, is deleterious in African populations (Carlson et al., 2012, Geldmacher et al., 2007, Kiepiela et al., 2004, Lazaryan et al., 2011, Leslie et al., 2010, Matthews et al., 2012, Saathoff et al., 2010, Tang et al., 2010).

Table 6.3. *HLA-B* alleles present in *HLA-B*58:01* and *-B*58:02*-positive progressors, their association with HIV-1 disease progression, *HLA* Supertype, *HLA* ligand designation and allele frequencies known for the SAB population**.**

HLA-B										
ID	**Allele 1**	**P/D**	***HLA* Supertype**	***HLA* ligand**	**Allele frequency SAB (%)**	**Allele 2**	**P/D**	***HLA* Supertype**	***HLA* ligand**	**Allele frequency SAB (%)**
1	*B*15:03:01*	D	B27	Bw6	8.40	*B*58:02*	D	B58	$Bw4^{80I}$	9.41
2	*B*45:01:01*	D	B44	Bw6	6.36	*B*58:02*	D	B58	$Bw4^{80I}$	9.41
3	*B*44:03:02*	P	B44	$Bw4^{80T}$	3.31	*B*58:02*	D	B58	$Bw4^{80I}$	9.41
4	*B*45:01:01*	D	B44	Bw6	6.36	*B*58:02*	D	B58	$Bw4^{80I}$	9.41
5	*B*08:01:01*	D	B08	Bw6	6.36	*B*58:01:01*	P	B58	$Bw4^{80I}$	8.14
7	*B*15:10:01*	D	B27	Bw6	7.89	*B*58:02*	D	B58	$Bw4^{80I}$	9.41
8	*B*15:03:01*	D	B27	Bw6	8.40	*B*58:01:01*	P	B58	$Bw4^{80I}$	8.14
9	*B*39:10*	P	B07	Bw6	2.54	*B*58:02*	D	B58	$Bw4^{80I}$	9.41
6	*B*44:03:02*	P	B44	$Bw4^{80T}$	3.31	*B*58:02*	D	B58	$Bw4^{80I}$	9.41
10	*B*53:01:01*	D	B07	$Bw4^{80I}$	4.33	*B*58:02*	D	B58	$Bw4^{80I}$	9.41
11	*B*58:01:01*	P	B58	$Bw4^{80I}$	8.14	*B*58:01:01*	P	B58	$Bw4^{80I}$	8.14
12	*B*53:01:01*	D	B07	$Bw4^{80I}$	4.33	*B*58:02*	D	B58	$Bw4^{80I}$	9.41
13	*B*44:03:02*	P	B44	$Bw4^{80T}$	3.31	*B*58:02*	D	B58	$Bw4^{80I}$	9.41
14	*B*44:03:02*	P	B44	$Bw4^{80T}$	3.31	*B*58:02*	D	B58	$Bw4^{80I}$	9.41
15	*B*42:01:01*	P	B07	Bw6	8.91	*B*58:02*	D	B58	$Bw4^{80I}$	9.41
16	*B*15:10:01*	D	B27	Bw6	7.89	*B*58:01:01*	P	B58	$Bw4^{80I}$	8.14
17	*B*15:10:01*	D	B27	Bw6	7.89	*B*58:01:01*	P	B58	$Bw4^{80I}$	8.14
18	*B*42:01:01*	P	B07	Bw6	8.91	*B*58:02*	D	B58	$Bw4^{80I}$	9.41

HLA, human leukocyte antigen, P, protective, D, deleterious, *HLA-B*58:01*-positive participants shaded in grey.

6.2.2.3. Alleles at the HLA-C locus

HLA-C genotypes determined for each study participant and allele frequencies reported for the SAB population are listed in **Table 6.4.** Ten different *HLA-C* alleles were identified in the study population. *HLA-C*07:01*, identified in 7 individuals, was found to be protective in African individuals (Kiepiela et al., 2004, Leslie et al., 2010), but not in African-Americans (International et al., 2010). *HLA-C*06:02*, present in 13 individuals, is deleterious allele in African populations (Kiepiela et al., 2004, Leslie et al., 2010, Matthews et al., 2012) and Japanese populations (Naruto et al., 2012). *HLA-C*04:01*, found in 3 individuals, is considered a protective allele in African populations (Matthews et al., 2012, Tang et al., 2002, Kiepiela et al., 2004, Leslie et al., 2010) and Japanese populations (Naruto et al., 2012). *HLA-C*02:10*, present in 2 individuals, have been found to be deleterious in a study from South Africa/Botswana (Matthews et al., 2012). *HLA-C*03:02*, found in 2 individuals, is protective in African populations (Matthews et al., 2012, Leslie et al., 2010). *HLA-C*03:04*, found in one study participant, is considered deleterious in African and Japanese populations (Kiepiela et al., 2004, Naruto et al., 2012). *HLA-C*08:04*, also present in one individual, was reported as a protective allele in a recent study in long-term surviving adolescents infected at birth in Zimbabwe (Shepherd et al., 2015). *HLA-C*16:01*, found in 3 individuals, is reported to be deleterious in African populations (Kiepiela et al., 2004,

Matthews et al., 2012). *HLA-C*12:03*, found in one individual, is considered protective allele in African populations (Carlson et al., 2012, Leslie et al., 2010, Matthews et al., 2012). *HLA-C*17:01*, found in two individuals, is also considered a protective allele in African populations (Carlson et al., 2012, Matthews et al., 2012, Tang et al., 2002).

Table 6.4. *HLA-C* alleles present in *HLA-B*58:01* and *-B*58:02*-positive progressors, their association with HIV-1 disease progression, *HLA* Supertype, *HLA* ligand designation and allele frequencies known for the SAB population**.**

HLA-C								
ID	**Allele 1**	**P/D**	***HLA* ligand**	**Allele frequency SAB (%)**	**Allele 2**	**P/D**	***HLA* ligand**	**Allele frequency SAB (%)**
1	*C*04:01:01*	P	C2	11.9	*C*04:01:01*	P	C2	11.9
2	*C*06:02:01*	D	C2	14.94	*C*16:01:01*	D	C1	6.58
3	*C*06:02:01*	D	C2	14.94	*C*07:01:01*	P	C1	7.59
4	*C*06:02:01*	D	C2	14.94	*C*16:01:01*	D	C1	6.58
5	*C*07:01:01*	P	C1	7.59	*C*07:01:01*	P	C1	7.59
7	*C*03:04:02*	D	C1	5.06	*C*06:02:01*	D	C2	14.94
8	*C*02:10*	D	C2	7.1*	*C*07:01:01*	P	C1	7.59
9	*C*06:02:01*	D	C2	14.94	*C*12:03:01*	P	C1	1.52
6	*C*06:02:01*	D	C2	14.94	*C*07:01:01*	P	C1	7.59
10	*C*04:01:01*	P	C2	11.9	*C*06:02:01*	D	C2	14.94
11	*C*02:10*	D	C2	7.1*	*C*03:02:01*	P	C1	1.52
12	*C*04:01:01*	P	C2	11.9	*C*06:02:01*	D	C2	14.94
13	*C*06:02:01*	D	C2	14.94	*C*07:01:01*	P	C1	7.59
14	*C*06:02:01*	D	C2	14.94	*C*07:01:01*	P	C1	7.59
15	*C*06:02:01*	D	C2	14.94	*C*17:01:01*	P	C2	11.4
16	*C*03:02:01*	P	C1	1.52	*C*08:04*	P	C1	2.53
17	*C*07:01:01*	P	C1	7.59	*C*16:01:01*	D	C1	6.58
18	*C*06:02:01*	D	C2	14.94	*C*17:01:01*	P	C2	11.4

HLA, human leukocyte antigen, P, protective, D, deleterious, *HLA-B*58:01*-positive participants shaded in grey.

6.2.3. *KIR* genotypes

The presence/absence of the 16 known *KIR* genes was determined by real-time PCR assay (Hong et al., 2011) and *KIR* genotypes were assigned according to the system of nomenclature used by the online Allele Frequency Net Database (AFND) (http://www.allelefrequencies.net) (**Table 6.5**). In *HLA-B*58:01* carriers, two individuals were genotyped as *Bx6*, while the remaining 3 were *AA1*, *Bx4* and *Bx91*. In *HLA-B*58:02* carriers, the *AA1* genotype was found in 3 individual, *Bx4* and *Bx5* genotypes were each found in two individuals, while the remaining individuals were classified as *Bx6, Bx20, Bx71, Bx92, Bx112* and *Bx393*. Although the difference was not significant, the average number of activating *KIRs* (i.e. *KIR2DS1-5* and *KIR3DS1*) was higher in *HLA-B*58:01* carriers compared to *HLA-B*58:02* carriers (4.0 vs. 2.76, respectively), while the average number of inhibitory *KIRs* detected in *HLA-B*58:01* carriers compared to *HLA-B*58:02* carriers was similar (7.2 vs. 7.1, respectively).

Table 6.5. *KIR* genes and *KIR* genotypes determined for the *HLA-B*58:01* and *-B*58:02*-positive progressors.

ID[1]	*KIR* gene[2]																*KIR* genotype[1]
	Inhibitory *KIRs*								**Activating *KIRs***						***KIR* pseudogenes**		
	2DL1	2DL2	2DL3	**2DL4**	2DL5	3DL1	**3DL2**	**3DL3**	2DS1	2DS2	2DS3	2DS4	2DS5	3DS1	2DP1	**3DP1**	
1	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	Bx6
2	■		■	■		■	■	■				■			■	■	AA1
3	■	■	■	■	■	■	■	■		■	■	■			■	■	Bx5
4	■		■	■		■	■	■				■			■	■	AA1
5	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	Bx6
6	■	■		■	■	■	■	■		■	■	■			■	■	Bx71
7	■		■	■		■	■	■				■			■	■	AA1
8	■	■	■	■		■	■	■		■		■			■	■	Bx4
9	■	■		■	■	■	■	■	■	■		■	■		■	■	Bx92
10	■	■	■	■	■	■	■	■		■	■	■			■	■	Bx5
11	■		■	■		■	■	■				■			■	■	AA1
12	■	■	■	■		■	■	■		■		■			■	■	Bx4
13	■	■		■	■	■	■	■		■	■	■	■		■	■	Bx112
14	■	■	■	■		■	■	■		■		■			■	■	Bx4
15	■		■	■	■	■	■	■	■	■		■	■		■	■	Bx393
16	■	■		■	■	■	■	■	■	■	■	■	■		■	■	Bx91
17	■	■	■	■	■	■	■	■				■	■	■	■	■	Bx6
18	■	■	■	■	■	■	■	■				■	■		■	■	Bx20

KIR, killer-cell immunoglobulin-like receptor. *KIRs* in bold are framework genes; [1]ID numbers and *KIR* genotypes of *HLA-B*58:01*-positive participants are shaded in grey; [2]The presence of a *KIR* gene is indicated by a blue shaded box under the respective *KIR* gene labelled column.

6.2.4. Mutational analysis of HIV-1 Gag

6.2.4.1. Mutations in p17^{Gag}

6.2.4.1.1. CTL epitopes

We were unable to analyse mutations in the RY10 (^{20}RLRPGGKKHY^{29}) epitope restricted by HLA-B*42:01 (Chopera et al., 2011), since the coding region for the first 52 amino acids of p17^{Gag} were not sequenced. However, partial p17^{Gag} sequences were available to analyse mutations in the highly antigenic region that includes the RY11 (^{76}RSLYNTVATLY^{86}), LY9 (^{78}LYNTVATLY^{86}), SL9 (^{77}SLYNTVATL^{85}), GY9 (^{70}GSEELRSLY^{79}), EV9 (^{73}ELRSLYNTV^{83}) and TI9 (^{84}TLYCVHQRI^{92}) epitopes restricted by HLA-A*30, -A*29/-B*44, -A*02, -A*01, -B*08 and -A*11, respectively (Tenzer et al., 2009).

In subtype C HIV-1 studies it has been reported that the RY11 epitope (**Figure 6.5**) is restricted by HLA-A*30:02 (Masemola et al., 2004b, Kiepiela et al., 2004, Chopera et al., 2008, Kiepiela et al., 2007) and Frahm *et al.* reported that RY11 is possibly also restricted by

HLA-B*58:01 (Frahm et al., 2005). Due to insertions/deletions that obscured sequence analysis, the baseline $p17^{Gag}$ sequence was not available for ID8. Many mutations were presented in this region and may be indicative of immune selection pressure due to the presence of relevant *HLA class I* alleles or other *HLA class I* alleles belonging to the same Supertype (Sidney et al., 2008). Alleles of the same Supertype that may target this region were present in the study participants and included *HLA-A*02:01*, *-A*02:05* and *-A*68:02* (A02 Supertype); *HLA-A*30:02*, *-A*30:01*, *-A*29:02*, *-A*36:01* and *-A*80:01* (A01 Supertype) and *HLA-B*44:03* and *-B*45:01* (B44 Supertype) (**Table 6.2**). *HLA-B*08:01*, that targets the EV9 epitope, was present in 1/18 study participants, but no participants were positive for *HLA-A*01* or *-A*11*, that targets the GY9 and TI9 epitopes, respectively.

The Y79F mutation $p17^{Gag}$ was the most common change observed in 10/18 (56%) of the study participants. This change is expected to alter the P2 anchor residue of the RY11 epitope restricted by *HLA-A*30*, however this is a TCR escape mutation with respect to the SL9 epitope restricted by *HLA-A*02*. R76K was a common mutation in *HLA-B*58:02* carriers, 5/13 (38%). Only one *HLA-B*58:01* carrier had R76K present at baseline, but this reverted by the 6 month time-point. L75I was also seen in 6 individuals. Two individuals, ID5 and ID9, did not show any other mutations besides L75I, while L75I was found in combination with R76K in three individuals (ID16, ID13 and ID15), although reversion to wild type occurred in ID16. R76K is a flanking mutation in the SL9 epitope and is not associated with failure of peptides to bind to *HLA-A*02* receptors (Brander et al., 1999), however, minor variations in proteosomal cleavage patterns, more so if downstream mutations were also present, have been reported suggesting that this change may be selected as a minor processing escape mutation (Tenzer et al., 2009). An association between L75I and *HLA-A*02:02* in subtype A HIV-1 infection has been reported, also suggesting a processing escape mutation (Peters et al., 2008).

The R76K and Y79F mutations were present in one *HLA-B*58:01* carrier (ID16) at baseline, but reverted to wild-type at the 12 month time-point, with the concomitant appearance of T81A and T84I mutations. R76K and T84V escape mutations have been observed in an *HLA-A*30:02* carrier (Chopera et al., 2011). T81A is unusual, but was found in subtype A HIV-1 infections associated with *HLA-A*02:01*. T84V is a more common escape mutation that is observed in *HLA-A*02* carriers (Tenzer et al., 2009). ID16 is both *HLA-A*02:01* and *-A*68:02* positive and both of these alleles are members of the A02

Supertype (Geldmacher et al., 2007). Changes in position 81 and 84 were also noted in an *HLA-A*02:01* positive subtype C infected individuals (Liang et al., 2011). Five *HLA-B*58:02* carriers (ID12, ID10, ID13, ID7 and ID15) also had R76K and Y79F mutations in combination with *HLA-A*02:01* (ID12), *-A*36:01* (ID10), *-A*29:02/-A*30:02/-B*44:03* (ID13), *-A*68:02* (ID7) and *-A*30:01/-A*30:02* (ID15), the most likely HLAs applying immune selective pressure. ID8 is a *HLA-A*02:05* and *-A*29:02* carrier, but did not show any mutations in this region. ID17 is a *HLA-A*02:05* carrier and only demonstrated a V82I escape mutation (Tenzer et al., 2009). Y79F alone was found in ID11 (*HLA-A*29:02*), ID2 (*HLA-A*02:05/-A30:02*) and ID14 (*HLA-A*29:02/-A*80:01/-B*44:03*). Surprisingly, Y79F reverted to wild-type at the 12 month time-point in ID14. ID1 is *HLA-A*02:01* positive and showed Y79F only at baseline and both Y79F and V82L at the 6 month time-point. ID18 unexpectedly did not show any mutations in this region, but is *HLA-A*02:01* positive. ID3 did not show any changes in this region but is a *HLA-B*44:03* carrier. Similarly, ID6 did not show any changes and is *HLA-A*29:02* and *-B*44:03* positive. LY9 is presented by *HLA-B*44:03* and *-A*29:02* (Masemola et al., 2004b, Chopera et al., 2008). Unexpectedly, ID4 had no changes in this epitope and is *HLA-B*45:01*-positive (B44 Supertype), while ID9 also had no changes and is a *HLA-A*30:02* carrier. It has been postulated that reversion of escape mutations to wild-type in the presence of the restricting HLA molecule may be associated with a lack of immune pressure due to immune dysfunction associated with late stage AIDS (Iversen et al., 2006, Tenzer et al., 2009). An unusual L75I mutation was present often in combination with R76K in 6 individuals (ID5, 16, 11, 13, 9 and 15). In ID16, reversion to wild-type was seen at the 12 month time-point. In ID11, L75F was observed at baseline and L75I at 6 month time-point. L75I was present in two individuals (ID5 and ID9) who did not have any other mutations in the epitope target region. ID5 is *HLA-B*58:01*-positive and L75I lies within the EV9 epitope which may represent an escape mutation in this case.

Of interest was the noticeable lack of R76K in *HLA-B*58:01* carriers compared to *HLA-B*58:02* carriers. R76K occurred in combination with Y79F in only one *HLA-B*58:01* carrier at baseline, but this reverted within a 6 month period. A closer examination of the position of these changes raised the possibility that the Y79F mutation potentially creates a new epitope that may be restricted by both HLA-B*58:01 and -B*58:02 molecules. A conserved threonine residue at position 72 could serve as the anchor residue with the phenylalanine change at position 79 serving as anchor residue at the carboxy terminus. This is based on the peptide binding motifs predicted for *HLA-B*58:01* and *-B*58:02*, i.e. X-

(ATSG)-XXXXXX-(WFY) and X-(ST)-XXX-(R)-XX-(F), respectively (http://www.hiv.lanl.gov). *HLA-B*58:02* has an additional requirement of arginine in a central position, which may explain the R76K change that predominated in *HLA-B*58:02* carriers. L75I may additionally contribute to escape by interfering with TCR recognition or this may indeed be an epitope processing mutation, since 2 individuals had no other changes except L75I, and a third individual with Y79F, showed evolution from L75F to L75I between baseline and 6 months. This possibility of a new epitope being generated by the L75F change may explain avoidance of this change in *HLA-B*58:01* carriers and secondary R76K changes in *HLA-B*58:02* carriers, but should be determined by functional assays. Interestingly, a R76K change in an *HLA-A*30:02/-B*58:0*1 carrier, has been observed, but there was an absence of R79F over the 13 month study period, suggesting that this change may not have evolved due to the presence of *HLA-B*58:01*. The phenomenon of mutations that escape from one HLA class I molecule but inadvertently create detectable epitopes restricted by another HLA class I molecule have been reported before as in two instances in the HIV-1 Nef protein, aptly termed "tug-of-war" mutations (Brumme et al., 2007). Mutations in reverse transcriptase due to drug resistance development have also been shown to influence CTL detection of modified epitopes (Mahnke and Clifford, 2006).

6.2.4.1.2. Insertions

Amino-acid insertions in $p17^{Gag}$ (**Figure 6.5**) were noted in 3 individuals (ID2, ID6 and ID10) at amino acid position 117 near the $p17^{Gag}/p24^{Gag}$ cleavage site. A 3 amino acid insertion (EQN) was present at all time-points in ID2. An insertion of 4 amino acids (KTQQ) was present at all time-points in ID6 and a similar 8 amino acid insertion (ASQQKTQQ) was detected at baseline in ID10, but was lost at the 6 and 12 months time-points. Other investigators have observed that similar mutations were related to restoration of viral replicative fitness due to the development of protease resistance mutations (Tamiya et al., 2004). The loss of the insertion in ID10 is possibly due to transmission of a protease resistant virus to a drug naïve recipient leading to reversion.

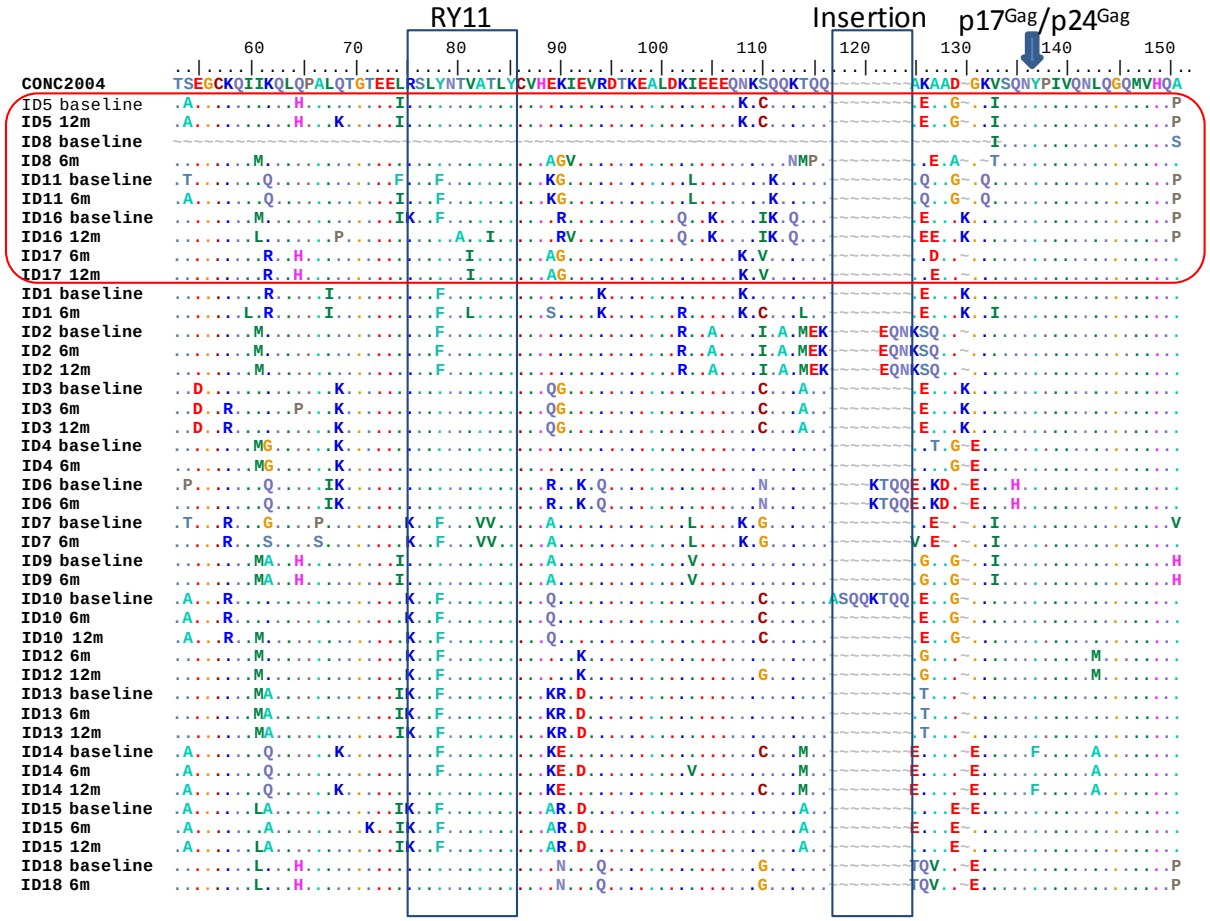

Figure 6.5. Protein sequence alignment of the partial p17^{Gag} open reading frame with a subtype C consensus reference showing the RY11 epitope and a region containing insertions. The arrow indicates the p17^{Gag}/p24^{Gag} cleavage site. Gaps introduced into the sequence are indicated by "~"; Identical amino acids to the consensus are indicated by "." *HLA-B*58:01*-positive individuals' sequences are shown in red box.

6.2.4.2. Mutations in p24^{Gag}

6.2.4.2.1. CTL epitopes

Analysis of the ISW9 (^{147}ISPRTLNAW^{155}) epitope in p24^{Gag} (**Figure 6.6**) known to be targeted by HLA-B*57 and -B*58:01 (Chopera et al., 2008) showed the presence of A146P in three *HLA-B*58:01* carriers (ID5, ID16 and ID11) and one *HLA-B*58:02* carrier (ID18). An unusual A146S mutation was present at baseline in one *HLA-B*58:01* carrier (ID8), but reverted by the 6 month time-point. Two *HLA-B*58:02* carriers also had unusual changes, i.e. A146V which was present at baseline but reverted by the 6 month time-point in ID7 and A146H that was established in ID9. The I147L mutation was present in only one *HLA-B*58:01* carrier (ID8) and two *HLA-B*58:02* carriers (ID6 and ID18). ISW9 is not a common epitope targeted by *HLA-B*58:01* carriers, hence the mutations observed in 4/5

*HLA-B*58:01* carriers may be an aspect of virological failure at late stages of chronic HIV-1 infection (Kloverpris et al., 2012b). The presence of mutations such as I146P in ISW9 in *HLA-B*58:02* carriers not known to respond to this epitope, may be genetic footprints undergoing reversion due to transmission from *HLA-B*57/58:01* donors (Chopera et al., 2008).

Mutations in the KF11 (^{162}KAFSPEVIPMF^{172}) epitope (**Figure 6.6**) were not present in *HLA-B*58:01* carriers, as expected, due to an intrinsic failure in presentation of this epitope (Kloverpris et al., 2013b, Kloverpris et al., 2012b), although in one report, an *HLA-B*58:01*-positive individual, who was *HLA-B*57*-negative, showed a response to this epitope (Chopera et al., 2011). Two *HLA-B*58:02* carriers (ID2 and ID6) carried V36I mutations within KF11, suggestive of possible weak CTL responses to this epitope that may mediate escape from TCR recognition. A weak CTL response to two peptides containing the KF11 epitope were reported in a subtype B HIV-1 infected individual who was *HLA-B*58:02*-positive (Spivak et al., 2010).

A p24^{Gag} epitope AT12 (^{179}ATPQDLNTMLNT^{190}) (**Figure 6.6**) predicted to be restricted by HLA-B*58:02 (Chopera et al., 2008), was examined and found unchanged in 10/13 *HLA-B*58:02* carriers. Three individuals, however, had polymorphisms at position 182 (Q182G in ID9; Q182S in ID14 and Q182T in ID15) possibly reflecting TCR contact escape mutations, but this epitope may be rarely targeted. Another p24^{Gag} epitope DW10 (^{203}DTINEEAAEW^{212}) (**Figure 6.6**), newly described as a HLA-B*58:01-restricted epitope was examined, but did not appear to be under selection pressure in our cohort. One *HLA-B*58:01* carrier (ID11) showed an E207D mutation at baseline but this reverted at the 6 month time-point. An E211D mutation was observed at baseline in an *HLA-B*58:02* carrier (ID12), but reverted at 12 months, while a second *HLA-B*58:02* carrier (ID10) had two established changes, D203E and A209I.

As anticipated, the well characterised T242N escape mutation in the TW10 (^{240}TSTLQEQIAW^{249}) epitope (**Figure 6.7**) of p24^{Gag} (Chopera et al., 2011, Leslie et al., 2004) was already established in all *HLA-B*58:01* carriers at baseline and also present at subsequent time-points. In *HLA-B*58:02* carriers, one T242N (ID18) and an unusual T242S (ID10) mutation was also found. A T242S mutation in an *HLA-B*58:01*-positive individual has been reported (Leslie et al., 2004) and in *HLA-B*58:02*-positive individuals these

changes most likely represent footprints from HIV-1 transmission from a *HLA-B*57* or -*B*58:01*-positive donor (Chopera et al., 2008). A transition from T242N to T242S before final reversion to 242T has been recorded in one *HLA-B*58:01*-negative individual, suggesting that T242S is an intermediate change during the reversion process (Chopera et al., 2008). In our study, surprisingly, one *HLA-B*58:02* positive individual (ID10), developed the 242S mutation at 6 months and maintained the change at 12 months, suggesting an alternative benefit of T242S in this individual. An E245D change was noted in one *HLA-B*58:02* individual, ID2, which was reported in the *HLA-B*58:01*-negative group and I247V, present in ID12, was observed in the *HLA-B*57*-positive group (Leslie et al., 2004). In subtype B (Miura et al., 2009) and subtype C (Leslie et al., 2004) HIV-1 studies, a guanine at position 248 is considered wild-type. However, in the cohort under current study, an alanine was the most common amino acid at position 248 (11/18). Sequence polymorphisms involving G (n=3), T (n=2) and Q (n=2) at position 248 of TW10 were noted, possibly influencing TCR escape recognition.

In two *HLA-B*58:01* (ID16 and ID17) carriers who were also *HLA-B*15:10*-positive, T303I and T303A changes were observed in YL9 epitope (^{296}YVDRFFKTL^{304}) in p24^{Gag} (**Figure 6.8**). CTL responses to YL9 in B*15 carriers has been reported previously (Masemola et al., 2004b). In *HLA-B*58:02* carriers, T303I was also seen in two individuals (ID12 and ID4) and T303C was present in one individual (ID10), however, none of these individuals were *HLA-B*15*-positive or -*C*03*-positive. CTL responses to YL9 in *HLA-C*03* carriers and T303AIC escape mutations that decreased viral fitness have been documented (Honeyborne et al., 2010). Notably, ID16 was also *HLA-C*03:02*-positive and the T303I change may therefore be escape from two HLA restricted CTL responses.

In *HLA-B*58:01* carriers, analysis of the QW9 epitope (^{308}QATQDVKNW^{316}) in p24^{Gag} (**Figure 6.8**), showed two established T310S mutations (ID16 and ID11) and one individual (ID5) acquired the change between baseline and 12 months. Positive selection of T310S and S309A mutations have been seen in *HLA-B*58:01* carriers (Carlson et al., 2012). All *HLA-B*58:01* carriers in our study also carried the S309A mutation. The T310S change was only seen at baseline in one *HLA-B*58:02* carrier (ID10), but reverted to wild-type by 6 months. The T303S mutation in ID10 at baseline was accompanied by an A309G mutation. The positive selection of A309X and N315G mutations in QW9 has been observed in *HLA-B*53:01*-positive individuals (Carlson et al., 2012). Reversion of both T303S and A309G to

wild-type at the 6 month time-point and the appearance of N315V and D312E in ID10 (*HLA-B*53:01*-positive) suggests immune selection pressure on this epitope. D312E was also present in ID12 who was also an *HLA-B*53:01* carrier. The E312D change sometimes found in *HLA-B*57* carriers, was present in all except one *HLA-B*58:01* carrier (ID16), where this change was present at baseline but reverted to 312E by 12 months. Other studies, however did not find E312D to be a significant change in *HLA-B*58:01* carriers (Miura et al., 2009). In *HLA-B*58:02* carriers, 312E was more common (8/13) and two individuals (ID10 and ID13) acquired D312E mutations after baseline and before the 6 month time-point.

6.2.4.2.2. Compensatory mutations

Since the T242N mutation in TW10 is known to confer a fitness cost to the virus (Crawford et al., 2007, Leslie et al., 2004), compensatory mutations such as H219X, I223X and M228X in $p24^{Gag}$ that restore viral replicative capacities, can subsequently develop (Gijsbers et al., 2013). The H219Q mutation maps to the CypA binding site (**Figure 6.7**) and has been found to enhance viral fitness (Gatanaga et al., 2006). This change was found in one *HLA-B*58:01* carrier (ID8) and three *HLA-B*58:02* carriers (ID12, ID3 and ID9). I223V was found in all *HLA-B*58:01* carriers, one of which was acquired between the 6 and 12 month time-point (ID17). I223V was also found in nine *HLA-B*58:02* carriers (ID1, ID18, ID10, ID2, ID3, ID6, ID7, ID9 and ID14), two of which acquired the change between baseline and 6 month time-points (ID1 and ID10). The two *HLA-B*58:02* carriers with the T242N and T242S changes also carried the I223V compensatory mutations. In *HLA-B*58:01* carriers, M228I and M228L were present in two individuals (ID16 and ID11). M228I was also present in two *HLA-B*58:02* carriers (ID12 and -3) as well as one M228L mutation (ID7). As expected, the M250I mutation, that causes a fitness cost to subtype C viruses and does not occur together with T242N (Chopera et al., 2012), was not present in any of the individuals with T242N, but was present in an *HLA-B*58:02* carrier who did not have T232N. Notably, partial compensatory mutations S252G and D260E described in other studies were also present in this individual (Chopera et al., 2012).

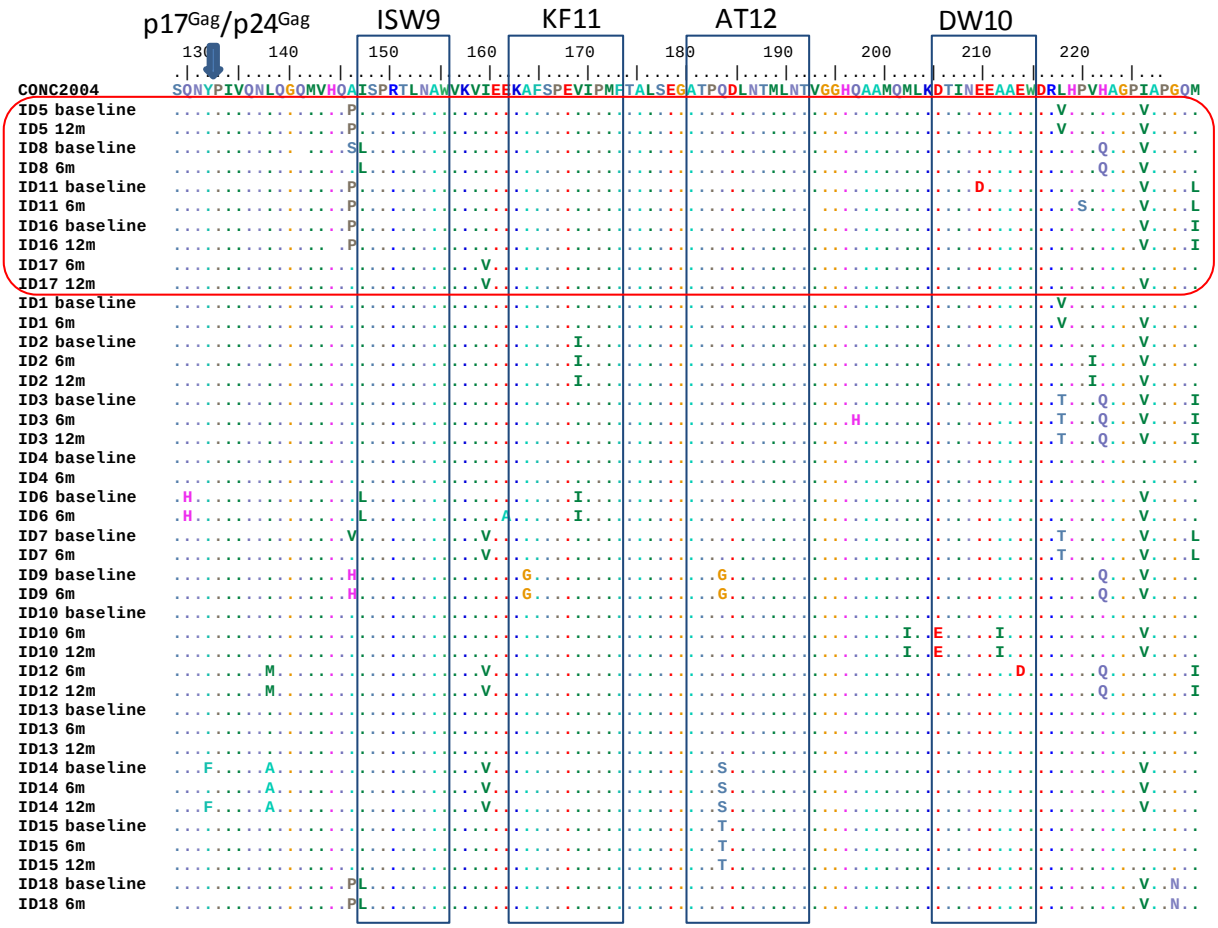

Figure 6.6. Protein sequence alignment of the partial p24^{Gag} open reading frame with a subtype C consensus reference showing the RY11, KF11, AT12 and DW10 epitopes. The arrow indicates the p17^{Gag}/p24^{Gag} cleavage site. Gaps introduced into the sequence are indicated by "~"; Identical amino acids to the consensus indicated by "."; *HLA-B*58:01*-positive individuals' sequences are shown in red box.

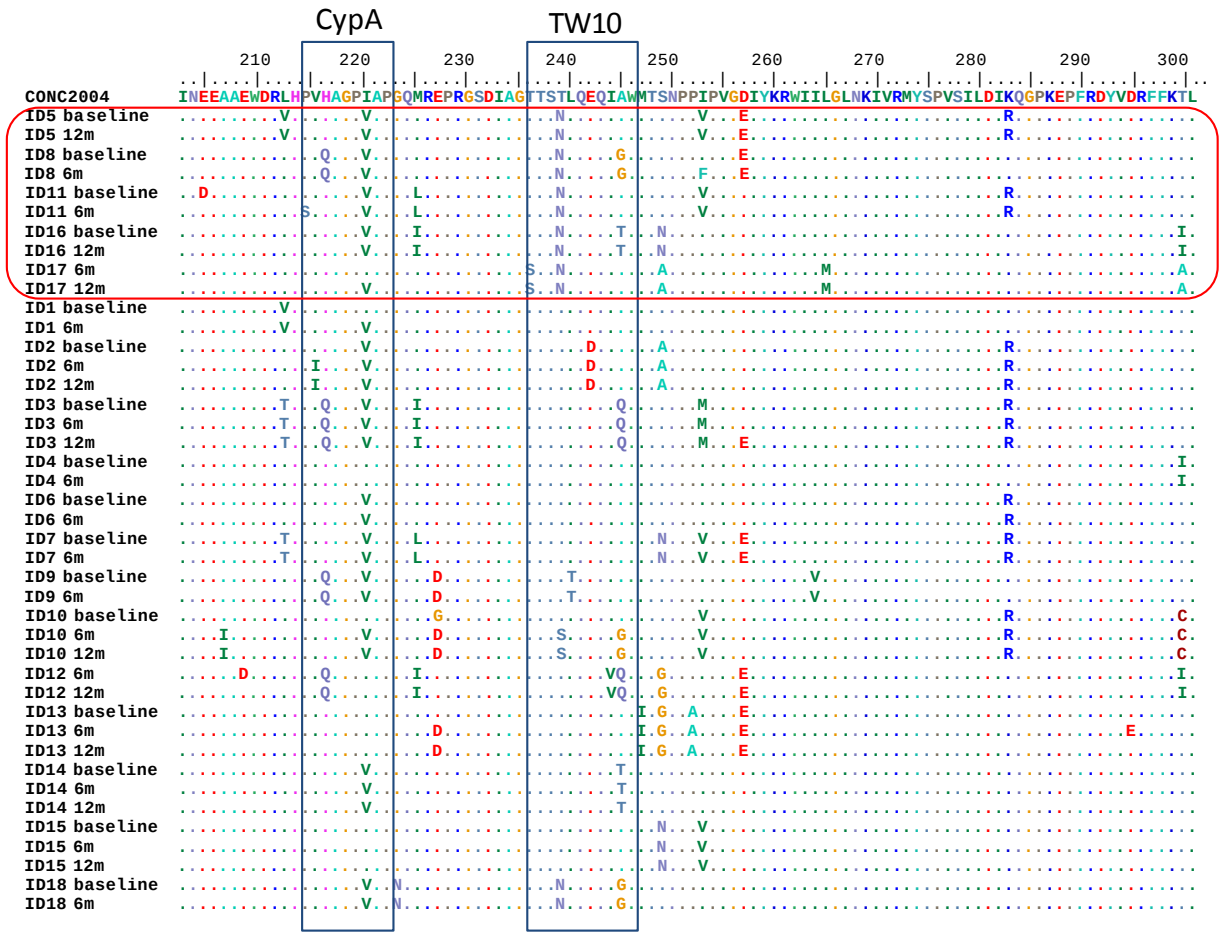

Figure 6.7. Protein sequence alignment of the partial p24^{Gag} open reading frame with a subtype C consensus reference showing the CypA binding site and the TW10 epitope. Gaps introduced into the sequence are indicated by "~"; Identical amino acids to the consensus indicated by "." ; *HLA-B*58:01*-positive individuals' sequences are shown in red box.

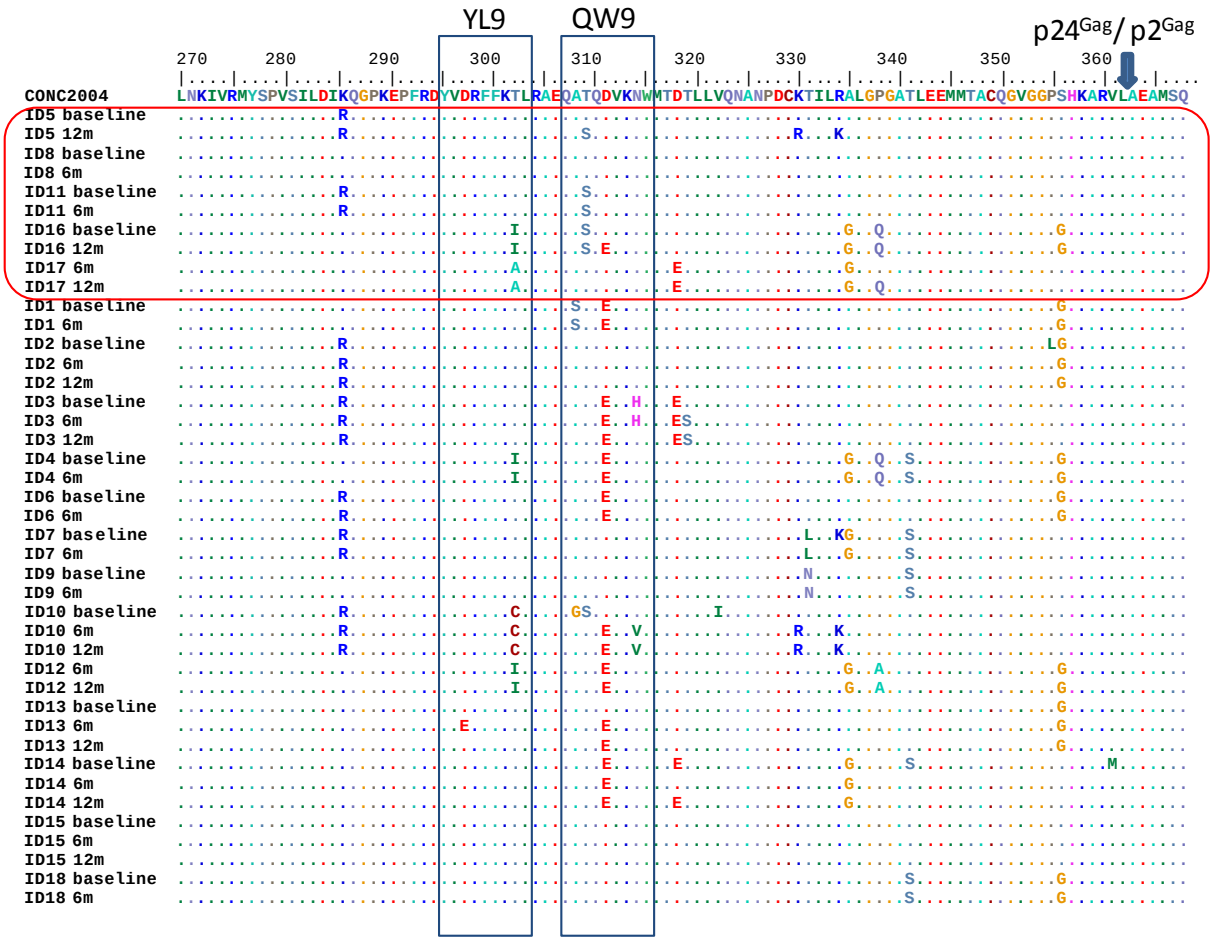

Figure 6.8. Protein sequence alignment of the partial p24^{Gag} open reading frame with a subtype C consensus reference showing the YL9 and QW9 epitopes. The arrow indicates the p24^{Gag}/p2^{Gag} cleavage site. Gaps introduced into the sequence are indicated by "~"; Identical amino acids to the consensus indicated by "."; *HLA-B*58:01*-positive individuals' sequences are shown in red box**.**

6.2.4.2.3. KIR binding

Mutations in the TW10 epitope (**Figure 6.7**) have also been shown to influence the strength of binding of KIR3DL1 to HLA-B*57 ligands (Brackenridge et al., 2011). In their report, an alanine at position 248, that we found as the consensus, was shown to marginally increase the binding affinity of KIR3DL1. Two individuals, ID16 and -14 had a threonine at position 248, which was reported to show the highest increase in binding affinity for KIR3DL1. A glutamine at position 248, however, which was present in two individuals, ID12 and -3, decreases the strength of binding of KIR3DL1. Mutations that increase the binding affinity of inhibitory receptors such as KIR3DL1 with HLA class I ligands, may be positively selected in order to prevent activation of antiviral NK cells. This may also apply to HLA-B*58:01 and -B*58:02 since these receptors are also Bw4 ligands for KIR3DL1 and may

explain the increased frequency of alanine at this position at end stage chronic HIV-1 infection.

Similarly, changes at position 303 of the YL9 epitope such as T303A and T303I have also been associated with increased binding of KIR2DL2 to HLA-C*03 molecules, while T303C reduced the binding affinity (van Teijlingen et al., 2014). ID16 had the T303I change and was also *KIR2DL2*-positive, suggesting potential escape from NK cell responses. ID17 with T303A was also *KIR2DL2*-positive, but did not have *HLA-C*03*; however both HLA-C molecules were C1 ligands. ID10 and ID12, with T303I and T303C, carried C2 ligands only and are not expected to bind KIR2DL2, even though they carried this KIR. ID4 had T303I, but was negative for *KIR2DL2*. Therefore, the presence of mutations at position 303 in the *HLA-B*58:02* carriers cannot be explained by known CTL epitope responses or escape from NK cell surveillance. These may likely be footprints of escape mutations generated in the donor partner.

6.2.4.2.4. TRIM5α sensitivity

Another possibility worth discussion is the effect of mutations in $p24^{Gag}$ on the ability of TRIM5α to restrict viral cores (Granier et al., 2013). Within the TW10 epitope (**Figure 6.7**), the T242N mutation alone or in combination with amino acid substitutions at position 248, was shown to increase the sensitivity to TRIM5α. Notably, T242N in the presence of 248G showed the highest sensitivity, while T242N in the presence of 248A was reduced. This may explain the presence of T242N and 248A in 3/5 *HLA-B*58:01* carriers. The presence of 248Q and 248T was not evaluated by the investigators, which were polymorphisms present in some of the $p24^{Gag}$ sequences. Most of the *HLA-B*58:02* carriers (7/13), however, carried a 248A mutation, which implies greater sensitivity to TRIM5α, and presumably a greater benefit to the virus, such as decreased NK cell detection, as mentioned earlier, was provided by this change. Interestingly, one *HLA-B*58:02* carrier (ID10) had 242T and 248A at baseline, which mutated to 242S and 248G at 6 months and this change was maintained at 12 months. Unfortunately, a 242S change was not evaluated, but it could be speculated that the two changes in combination may have reduced sensitivity to TRIM5α.

6.2.4.3. Insertions in p6^{Gag}

Analysis of p6^{Gag} protein sequences showed the presence of PTAP insertions in 4/5 *HLA-B*58:01* carriers and in 5/13 *HLA-B*58:02* carriers (**Figure 6.9**). The PTAP duplication is reported to be more common in subtype C viruses, but is also present in other HIV-1 subtypes (Flys et al., 2005). A putative CTL epitope restricted by *HLA-B*58:01* (^{453}EPTAPPAESFRF^{464}) that contains the PTAP motif has been predicted previously (Chopera et al., 2011) and it is therefore plausible that insertions in this region in *HLA-B*58:01* carriers may disrupt presentation of the epitope as a means of CTL escape. On the other hand PTAP duplications are known to enhance viral budding through recruitment of Tsg101 (VerPlank et al., 2001). It has been demonstrated that deletion of the first PTAP motif leads to a 50% reduction in Tsg101 binding compared to a 25% reduction when the second PTAP motif is deleted. In our study, we noticed that the first PTAP motif appeared to be deleted in ID5, however sequences from this individual at baseline and 12 month time-point, had an unusual 4 amino insertion (PYRE) just upstream of the Vpr binding site. In subtype C HIV-1 infected Indian individuals, similar 4-amino acid insertions were found to be more common in patients failing protease therapy and are thought to be restoration of fitness changes due to the development of protease resistance mutations (Neogi et al., 2014). The functional significance of this insertion is that it restores the ALIX-binding motif that serves as an alternative budding pathway, normally absent in sub-Saharan subtype C viruses. In terms of viral budding efficiency, in ID5, the original PTAP duplication may have become redundant due to the PYRE insertion. Loss of the first PTAP motif was most likely of no consequence to the virus, since the deletion of the second PTAP motif has been shown to associate with a lesser reduction in Tsg101 recruitment than deletion of the first PTAP motif.

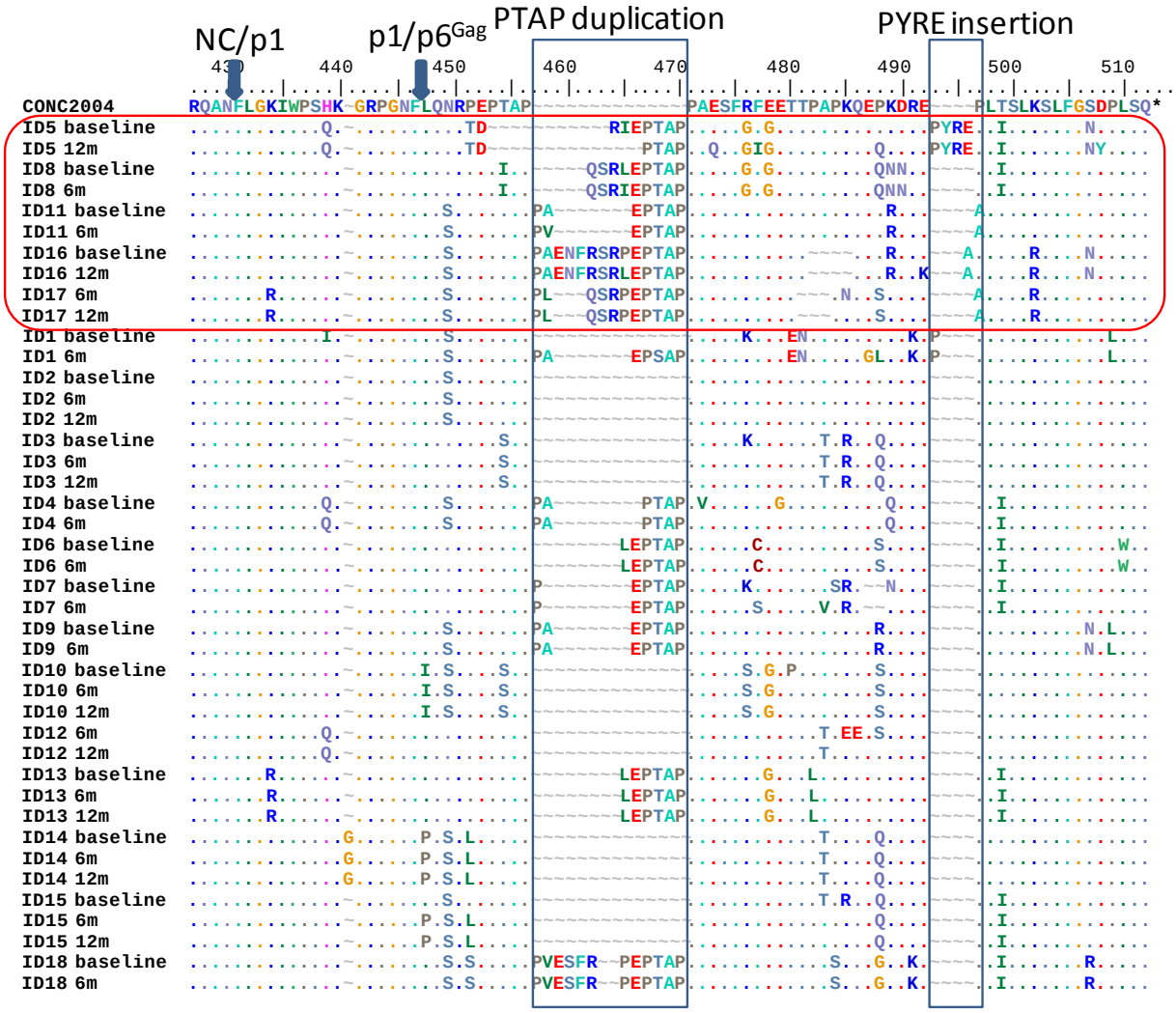

Figure 6.9. Protein sequence alignment of the p6^{Gag} open reading frame with a subtype C consensus reference showing the site of the PTAP duplication and the PYRE insertion. The arrows indicate the NC/p1 and p1/p6^{Gag} cleavage sites. Gaps introduced into the sequence are indicated by "~"; Identical amino acids to the consensus indicated by "."; *HLA-B*58:01*-positive individuals' sequences are shown in red box**.**

6.2.4.4. Insertions and p6^{Pol}

In addition, insertions in the p6^{Gag} region also the impact on the primary sequence of the p6^{Pol} transframe protein generated by the -1 frameshifting during Gag-Pol polyprotein synthesis (**Figure 6.10**). The possible significance of p6^{Pol} mutations is that this domain was found to be necessary for regulation of protease autocatalysis required for virion maturation (Paulus et al., 2004) as well as incorporation of Nef into virions (Costa et al., 2004). It seemed plausible to speculate that enhanced incorporation of Nef into virions may lead to increased downregulation of HLA class I molecules on the surface of newly infected cells.

However, other investigators have analysed p6^{Pol} mutants containing an ISSEQTRANSPT insertion in p6^{Pol} as a consequence of PTAP duplications in p6^{Gag}, and no differences in Nef incorporation compared to the wild-type strain were noted (Leiherer et al., 2009). This report, however, did demonstrate other consequences of the p6^{Pol} insertion, which included increased infectivity and enhanced replication in PBMC's. This implies that the insertions in p6^{Gag}, similar to insertions observed in the current study, that enhance virus budding may also impact on viral infectivity and replication kinetics mediated by p6^{Pol}.

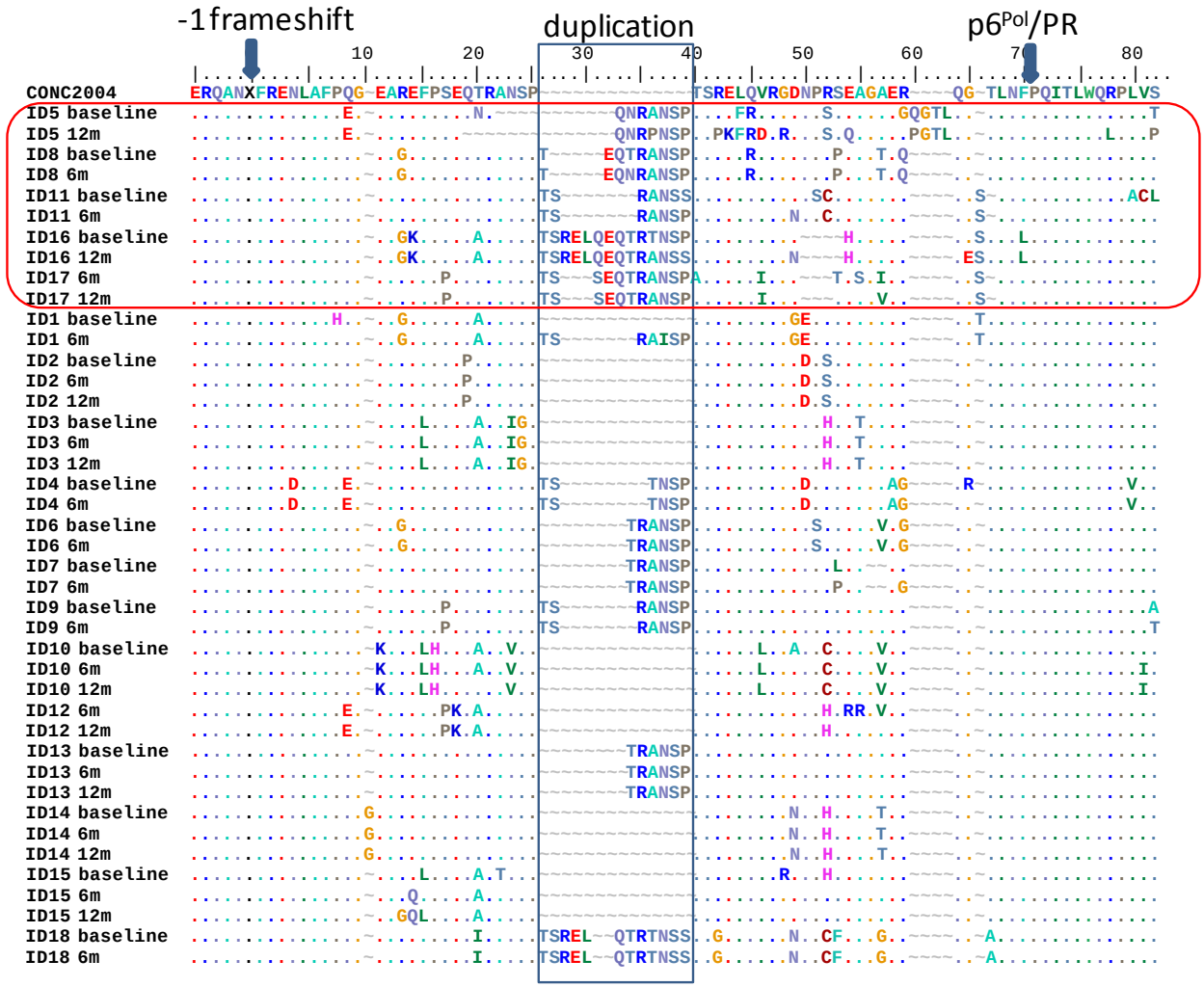

Figure 6.10. Protein sequence alignment of the p6^{Pol} open reading frame with a subtype C consensus reference showing the site of duplication. The arrows indicate the p6^{Pol}/PR cleavage site. Gaps introduced into the sequence are indicated by "~"; Identical amino acids to the consensus indicated by "."; *HLA-B*58:01*-positive individuals' sequences are shown in red box.

6.2.5. Summary of mutations in Gag

A summary of the mutations and insertions detected in Gag in the *HLA-B*58:01* and -*B*58:02* progressors is depicted in **Table 6.5**. Red squares indicate a change. Similar changes between the *HLA-B*58:01* and -*B*58:02* carriers are evident in the $p17^{Gag}$ RY11 epitope, compensatory mutations in $p24^{Gag}$, insertions in $p6^{Gag}$, NK escape mutations and TRIM5α sensitivity. Differences are evident between the *HLA-B*58:01* and -*B*58:02* progressors in the ISW9, TW10, YL9 and QW9 $p24^{Gag}$ epitopes. It is known that *HLA-B*58:02* carriers do not target $p24^{Gag}$, while the changes in $p24^{Gag}$ epitopes in *HLA-B*58:01* carriers are escape mutations in epitopes restricted by HLA-B*58:01. The summary of results suggest that similar escape mutations can develop in both *HLA-B*58:01* and -*B*58:02* carriers in targets that are independent of HLA-B*58:01 and -B*58:02-mediated CTL responses and these changes contribute to virological failure in both groups. However in *HLA-B*58:01* carriers, $p24^{Gag}$ is a major target for HLA-B*58:01-restricted CTLs, hence escape mutations develop in these epitopes and also contribute to virological failure. A lack of HLA-B*58:02-restricted epitopes in $p24^{Gag}$ explains why no escape mutations were detected in HLA-B*58:02 carriers and the virus therefore does not compromise fitness.

Table 6.6. Summary of mutations in Gag detected in *HLA-B*58:01* and *-B*58:02*-positive progressors.

	CTL escape mutations											**Fitness compensatory mutations**			**Insertions**			**NK Cell escape mutations**		**TRIM5α sensitivity**
	p17^{Gag}						p24^{Gag}					p24^{Gag}			p17^{Gag}	p6^{Gag}		p24^{Gag}		p24^{Gag}
ID	RY11						ISW9		TW10	YL9	QW9							*KIR3DL1*	*KIR2DL2*	
	L75	R76	Y79	T81	V82	T84	A146	I147	T242	T303	T310	H219	I223	M228				G248	T303	G248
5	I						P		N		S		V				PYRE	A		A
8							S	L	N			Q	V			PTAP				
11	I/F		F				P		N				V	L		PTAP		A		A
16	I	K	F	A		I	P		N	I	S		V	I		PTAP		T	I	
17					I				N	A	S		V			PTAP		A	A	A
1			F		I								V			PSAP		A		A
2			F										V		EQN			A		A
3												Q	V	I						
4										I						PTAP		A	I	A
6								L					V		KTQQ	PTAP		A		A
7		K	F			V	V						V	L		PTAP		A		A
9	I						H					Q	V			PTAP		A		A
10		K	F						S	C			V						I	
12		K	F							I		Q		I						
13	I	K	F													PTAP		A		A
14			F										V					T		
15	I	K	F															A		A
18							P	L	N				V			PTAP				

CTL, Cytotoxic T cell; NK, Natural killer; *KIR*, killer-cell Immunoglobulin-like receptor; TRIM5α, tripartite motif-5α; *HLA-B*58:01* carriers are shaded in grey.

6.3 Discussion.

In the current study, drug-naive HIV-1-positive South African Black individuals, followed up at 6 monthly intervals, who required treatment based on $CD4^+$ T cell counts declining to <350 cells/µl, were sub-selected for further investigation based on *HLA class I* genotype and the availability of two follow up plasma samples. Plasma samples stored prior to treatment were used for genetic characterisation of HIV-1 *gag* sequences. *HLA-B*58:01* and *-B*58:02*-positive participants were specifically selected for further study due to published reports that *HLA-B*58:01* associates with early Gag responses and low viraemia, while *HLA-B*58:02* associates with CTL responses to Env and rapid disease progression (Kiepiela et al., 2007, Ngumbela et al., 2008). However, *HLA-B*58:01* carriage does not guarantee slow disease progression since it has been shown that CTL escape mutations that incur fitness costs to the virus can be overcome by compensatory mutations that restore viral replicative capacities (Chopera et al., 2011). In contrast uncontrolled viral replication associates with *HLA-B*58:02*, which is not known to mediate antiviral CTL responses to Gag and preferentially targets variable proteins such as Env. Importantly, other investigators have reported multiple immune mechanisms inhibiting HIV-1 replication including the additive effect of the other *HLA class I* alleles present in the individual's repertoire (Leslie et al., 2010, Sampathkumar et al., 2014), NK cell-mediated targeting of infected cells (Alter et al., 2011) and the sensitivity of viral cores to the intrinsic antiretroviral factor, TRIM5α (Granier et al., 2013). It is the concerted effort of singular counter-measures mediated by adaptive and innate immune pressures on replicating virus, and the subsequent circumvention of these individual mechanisms, that determines the rate of disease progression. Via constant mutagenesis of viral proteins, HIV-1 reduces sensitivity to TRIM5α, evades NK cell detection and overcomes the debilitating effect of CTL escape mutations, particularly in Gag, that reduces viral fitness (Honeyborne et al., 2010, Crawford et al., 2007) by selecting compensatory mutations that restore viral replicative capacities. In addition, insertions in Gag can enhance viral budding, infectivity and replication kinetics (VerPlank et al., 2001), which in concert with immune evasion, further increases virus production.

An investigation of genetic alterations in HIV-1 *gag* sequences derived from viral RNA extracted from plasma virus in stored samples taken at baseline, 6 months or 12 months prior to initiation of antiretroviral therapy in *HLA-B*58:01* and *-B*58:02*-positive progressors, was carried out. In agreement with previous reports, all *HLA-B*58:01* carriers

demonstrated known CTL escape mutations in Gag epitopes, particularly the TW10 epitope, where the T242N mutation was present at baseline and subsequent time-points. This was not unexpected, since the study participants were in the late chronic stage of HIV-1 infection and T242N mutations have been reported to emerge as early as 2 weeks following acute infection, but usually within the first 3-4 months (Chopera et al., 2011). Mutations in other HLA-B*58:01-restricted epitopes, such as ISW9 and QL9 were also detected, suggesting that even though T242N prevents recognition of the TW10 epitope, the continuation of HLA-B*58:01 targeting of Gag, necessitates escape mutations in subdominant epitopes. Mutations in other HLA-B*58:01 epitopes were not observed in other similar studies (Chopera et al., 2011), however, the participants in the current study have most likely been infected for much longer than the follow up period of their study, allowing for more time for escape mutations to emerge.

In *HLA-B*58:02* carriers, despite the similar binding motifs predicted, no consistent mutations were observed when known HLA-B*58:01-restricted epitopes were examined, in agreement with reports that there are no HLA-B*58:02-restricted epitopes present in Gag (Ngumbela et al., 2008). One *HLA-B*58:02* carrier had escape mutations in ISW9 and TW10, but this was most likely genetic footprints of escape in a HLA-B*58:01 donor prior to transmission to this participant. As expected, no mutations were observed in the KF11 epitope, consistent with a previous report that even though this epitope binds to HLA-B*58:01, there is an intracellular block of access of the peptide to HLA-B*58:01 molecules prior to localisation to the cell membrane (Kloverpris et al., 2013b).

The additive effect of CTL pressure on Gag by other HLA class I molecules was well illustrated by analysis of the highly immunogenic $p17^{Gag}$ region between residues 70 and 92, containing several CTL epitopes (RY11, LY9, SL9, GY9, EV9 and TI9). This region is targeted by numerous HLA class I molecules including HLA-A*30, -A*29/-B*44, -A*02, -A*01, -B*08 and -A*11. *HLA-A*30*, *-A*29/-B*44*, *-A*02* alleles were all present in the study participants. R76K and Y79F were commonly detected mutations that mediate anchor residue changes, TCR contact changes or peptide processing interference. It was noted that R76K was only present in one *HLA-B*58:01*-positive individual, but reverted at 6 months, whereas 5 *HLA-B*58:02* carriers had this change. The possibility exists that the Y79F mutation could have been induced by other HLA class I-mediated pressures, but was less tolerated in *HLA-B*58:01* carriers, because the change potentially creates a new HLA-

B*58:01-restricted epitope. These viral variants would therefore be quickly eliminated. Y79F may also create a new HLA-B*58:02-restricted epitope, however based on binding motif predictions, an arginine is required at position 76. The frequent occurrence of an R76K mutation in combination with Y79F in *HLA-B*58:02* carriers provides speculation that the natural arginine at position 76 does generate a new HLA-B*58:02 epitope, that is subsequently overcome by a R76K change. Hence the two changes, Y79F and R76K, would provide escape from both additive HLA class I pressure as well as newly induced HLA-B*58:02 responses as a consequence of Y79F. However, functional studies are required to confirm that HLA-B*58:01 and -B*58:02-restricted CTLs respond to this potential epitope.

A contributory factor to non-progression in *HLA-B*58:01* carriers is the observation that this allele is often in linkage disequilibrium with other protective *HLA class I* alleles. In the current study, a significantly higher allele frequency of protective *HLA* class I alleles was observed in *HLA-B*58:01* carriers compared to *HLA-B*58:02* carriers (66.7% vs. 36.8%; p=0.0149; OR=3.257[1.329-7.984]). Leslie *et al.* reported improved HIV-1 control for the *HLA-C*12:03-B*39:10* haplotype, which are both protective alleles (Leslie et al., 2010). This haplotype was found in one *HLA-B*58:02* progressor, but may have been insufficient in controlling HIV-1 due to the development of escape mutations. An association of *HLA-A*30:02-B*45:01, -A*30:02-C*16:01, -B*53:01-C*04:01, -B*15:10-C*03:04* and *-B*58:01-C*03:02* haplotypes with accelerated disease progression has been reported (Sampathkumar et al., 2014). Both *HLA-A*30:02-B*45:01* and *-A*30:02-C*16:01* haplotypes were present in one *HLA-B*58:02* carrier and although *HLA-A*30:02* is a protective allele, both *HLA-B*45:01* and *HLA-C*16:01* are deleterious alleles. Two *HLA-B*58:02* carriers had the *HLA-B*53:01-C*04:01* haplotype while in *HLA-B*58:01* carriers, one *HLA-B*15:10-C*03:04* and two *-B*58:01-C*03:02* haplotypes were detected. Taken together the additive protective effect of *HLA class I* genotype may be compromised in both *HLA-B*58:01* and *-B*58:02* carriers, but more so in *HLA-B*58:02* carriers due to linkage disequilibrium with other deleterious alleles.

Immune surveillance by NK cells augments the antiviral CTL responses in HIV-1 infection. NK cell detection of HIV-1 infected cells is dependent on interaction of inhibitory or activating KIRs with HLA class I ligands. Interference with KIR3DL1 and KIR2DL2 interaction by mutations at the 7^{th} and 8^{th} position of HIV-1 peptides bound to the HLA ligand molecules have been reported (Brackenridge et al., 2011, van Teijlingen et al., 2014).

The mechanism of escape from NK cell detection can be mediated by mutations in the peptide that either promote increased strength of binding to inhibitory receptors or reduced strength of binding to activating receptors and this has been noted for other KIRs (Alter et al., 2011). Mutations in the 8th position of the TW10 epitope were detected in four *HLA-B*58:01* carriers, namely G248A and G248T, which have been shown to increase the binding affinity of KIR3DL1 to HLA-B*57 molecules (Brackenridge et al., 2011). HLA-B*58:01 and -B*58:02 molecules are ligands of KIR3DL1 due to the presence of the Bw4^{80I} epitope. G248T and G248Q changes were noted in one and two respective *HLA-B*58:02* carriers. These mutations possibly promote escape from KIR3DL1-mediated NK cell detection of virus-infected cells. Similarly, T303IAC mutations were present in the YL9 epitope which also interferes with KIR2DL2-mediated NK cell detection of HLA-C*03 ligands on infected cells (van Teijlingen et al., 2014). KIR2DL2 interacts with many HLA-C2 ligands, but the YL9 epitope is only restricted by HLA-B*15 and -C*03 (Masemola et al., 2004b, Honeyborne et al., 2010). One individual with a T303I mutation was *HLA-C*03*-positive suggesting escape from NK cell surveillance in this individual. T303IAC mutations are also reported to reduce viral fitness (Honeyborne et al., 2010), hence it was unexpected that these mutations were present in Gag from three individuals who did not carry *HLA-B*15* or *-C*03* alleles. Either compensatory mutations were also present or interactions with other KIRs may be in effect. Both *HLA-B*58:01* and *-B*58:02* carriers had similar numbers of inhibitory KIRs, however the number of activating KIRs tended to be higher in *HLA-B*58:01* carriers compared to *HLA-B*58:02* carriers. NK cell responses to HIV-1 peptides have been shown to be associated with a more activating KIR repertoire in the presence of at least one HLA-C1 allotype (Tiemessen et al., 2011). Notably, all 5 *HLA-B*58:01* carriers were *HLA-C1*-positive compared to 8/13 (61.5%) in the *HLA-B*58:02* group. This observation has implications of potential enhancement of NK cell mediated immune pressure on HIV-1 in *HLA-B*58:01* carriers, although this may not be true in all *HLA-B*58:01* carriers, since *KIR* and *HLA* genes are not necessarily inherited together. There is, however, the possibility that immune escape from NK cell surveillance may have contributed to disease progression in the *HLA-B*58:01* carriers in the current study.

Interference with the activity of intrinsic antiretroviral factor TRIM5α was also detected. Three *HLA-B*58:01* carriers demonstrated G248A mutations in combination with T242N together in the TW10 epitope. Recent work showing that the T242N mutation in combination with wild-type 248G renders the p24^{Gag} capsid more sensitive to TRIM5α has

come to light (Granier et al., 2013). The G248A change reduces sensitivity to TRIM5α and may be an alternative escape mechanism that HIV-1 employs to escape from immune pressure at an early stage of the viral replication cycle.

In addition to the additive effects of other *HLA class I* alleles, NK cell activity and TRIM5α sensitivity that drive the selection of escape mutations, we also observed a propensity for the virus to select for changes in Gag that enhance viral production. It is known that compensatory mutations develop during the chronic stage of disease that can restore viral replicative capacity mediated by the T242N mutation in TW10 (Brockman et al., 2010). Alternatively, compensatory mutations may be pre-existing at transmission as observed in a previous study (Chopera et al., 2011). This phenomenon of the accumulation of compensatory mutations in circulating virus over time at a population level has been observed in Botswana where the protective benefit of *HLA-B*57* and *-B*58:01* has been lost (Payne et al., 2014). In *HLA-B*58:01*-positive participants, the compensatory mutations, I223V, H219Q and M228I/L were present in 5/5, 1/5 and 2/5 individuals, respectively. Notably, I223V, H219Q and M228I/L was also found in 9/13, 3/13 and 3/13, respective *HLA-B*58:02* carriers. This finding suggests that these mutations were either pre-existing at transmission or that viral fitness compensatory mutations do not necessarily develop solely in response to low fitness escape mutations such as T242N in $p24^{Gag}$ TW10, but may represent a natural tendency for HIV-1 to increase replicative capacities in general. In further support of this, PTAP duplications in $p6^{Gag}$ are known to develop in subtype C viruses in order to enhance virus budding efficiency through the Tsg101-mediated pathway. PTAP duplications were more common in *HLA-B*58:01* carriers, suggesting a greater necessity for an increase in virus production due to the higher fitness cost incurred by escape mutations, such as T242N. A priority to enhance viral budding efficiency, was further exemplified by the observation of a rare PYRE insertion in $p6^{Gag}$ that restores the ALIX-mediated budding pathway, not common in sub-Saharan subtype C viruses. This same sample appeared to have lost the first of the PTAP duplications in preference for the PYRE change. PTAP duplications were also observed *HLA-B*58:02* carriers, but at a lower frequency. It is not known whether the PTAP duplications pre-existed or if they developed after infection. The duplication of the PTAP motif during the chronic stage was shown to occur in one *HLA-B*58:02* carrier between baseline and 6 months. This implies that PTAP duplications are not necessarily inherited at transmission, but may evolve following immune pressure that possibly reduces viral fitness. In addition, changes introduced into $p6^{Gag}$ by PTAP duplications are reflected by

changes in the $p6^{Pol}$ transframe protein due to overlapping reading frames. The corresponding changes in $p6^{Gag-Pol}$ have been shown to enhance viral infectivity and replication kinetics, particularly in PBMCs, and this may serve as an additional mechanism to further increase virus production (Leiherer et al., 2009).

This study however has a number of limitations. Besides the small sample sizes used in this study, other limitations include the absence of analysis of mutations in other viral proteins, such as Nef, that contribute to CTL-mediated immune control. We focused on Gag because it has been established that early Gag targeting associates with better immune control (Kiepiela et al., 2007) and it has also emerged that early control of HIV-1 infection can be mediated by CTLs and TRIM5α that target $p24^{Gag}$ capsid proteins derived from incoming cores of infecting virions (Kloverpris et al., 2013a, Granier et al., 2013). We also did not analyse Gag mutations in *HLA-B*58:01* and *-B*58:02* carriers with non-progression of HIV-1 disease, since we did not have access to samples from such individuals and due to lower viral loads, this may pose difficulties in amplifying *gag* sequences from plasma RNA. Lastly, there are many other host genetic factors besides *HLA* class I and *KIR* genotypes that can influence HIV-1 disease progression that we have not examined.

In conclusion, as others have shown, disease progression in *HLA-B*58:01* carriers associates with the presence of CTL escape mutations in HLA-B*58:01-restricted Gag epitopes that may transiently confer a fitness cost to replicating virus, but are subsequently overcome by the development of compensatory mutations. Alternatively, compensatory mutations may already be present in transmitting virus. In contrast, *HLA-B*58:02* carriers do not target Gag via HLA-B*58:02-restricted CTLs and therefore do not have the advantage of early immune control of HIV-1 replication. However, some additional observations can be added to this picture in terms of the contribution of escape mutations that impact on other antiviral pressures, including the additive effect of other *HLA class I* molecules, NK cell targeting and sensitivity to TRIM5α. Furthermore, increased virus production can also be mediated by selection of viral fitness compensatory mutations as well as insertions that enhance viral budding, infectivity and replication kinetics. It was observed that escape mutations from CTLs/NK cells and enhancements of viral production were more common in *HLA-B*58:01*-positive progressors than in *HLA-B*58:02*-positive progressors. Since *HLA-B*58:01* is more frequently in linkage with other protective *HLA class I* alleles, it is speculated that the increased CTL-mediated immune pressure on replicating virus resulted in

more frequent selection of escape mutations that incur viral fitness costs, necessitating genetic alterations in Gag to overcome these defects. However, in both *HLA-B*58:01* and -*B*58:02* carriers, disease progression was likely a consequence of the combined influence of accumulating escape mutations from multiple immune targets as well as enhancing mutations that increase virus production.

Chapter 7:
Concluding remarks

South Africa is a unique geographical region well suited for the study of the influence of immunogenetic variation on HIV-1 disease, since four distinct populations are present (SAI, SAM, SAC and SAB) and a single viral subtype predominates (subtype C HIV-1). For reasons unknown, HIV-1 prevalence is disproportionate between the populations. Low prevalence populations include the SAI, SAM and SAC populations (0.8%, 3.1% and 0.3%, respectively), while the SAB population has a high prevalence (15.0%). Naturally many host genetic differences are expected to be present between the populations, but if one set of host genes could be selected that differs even between individuals of the same population and is also integral in host defense against pathogens, it would be the *HLA class I -A, -B* and *-C* set of genes. These genes demonstrate the highest degrees of genetic heterogeneity in the human genome as evidenced by the number of allelic variants characterised (*HLA-A* alleles, n=2,995; *HLA-B* alleles, n=3,760 and *HLA-C* alleles, n=2,553), yet humans only carry two *HLA-A, B* and *-C* alleles. The encoded proteins are expressed on the surface of nucleated cells and bind small peptide antigens derived from proteins processed in the cytoplasm. HLA class I molecules serve as ligands for three different receptors expressed by cells of the immune system. Interaction with the T cell receptor and CD8 molecule modulates CTL activity, interaction with KIRs modulates NK cell activity and interaction with leukocyte immunoglobulin-like receptors (LILRs) modulates the activation of antigen presenting cells (macrophages and dendritic cells). Allelic variants of HLA class I molecules have different peptide antigen specificities which modulates CTL activity and different binding affinities for KIRs and LILRs, likewise modulates NK cell and antigen presenting cell activities. These differences can translate into variation in disease susceptibility or severity. An additional role for HLA class I molecules in precipitating immune hypersensitivity reactions to antiretroviral drugs in susceptible individuals has been shown.

Whereas many genotypic characterisation studies have been done in the SAC and SAB populations, the SAI and SAM populations are relatively understudied and there is therefore a lack of *HLA class I* genotypic information for these populations. New data was generated for the SAI and SAM populations containing the frequencies of important immunogenetic factors such as *HLA class I* genotypes as well as other host genetic factors reported to influence HIV-1 acquisition or viral load, including the *CCR5Δ32* variant, *KIR* genotypes, *HLA-Bw4*^{80I}*/Bw4*^{80T}*/C1/C2* ligands, *HLA-C 3'UTR* +263 insertion/deletion variants and the p2-M/T variant of the HLA-B signal peptide. Data available for the SAC and

SAB populations were included for comparative purposes. Generating the data for these populations also presented the opportunity for novel assays to be designed such as the real-time PCR assay for the determination of *CCR5Δ32* genotypes and for existing real-time PCR assays to be modified and improved such as the *KIR* genotyping assay and *HLA-Bw4*^{80I}*/Bw4*^{80T}*/C1/C2* ligand assay.

There is a rich diversity in the *HLA class I* alleles detected in the SAI, SAM, SAC and SAB populations. At the *HLA-A, -B* and *-C* loci, 42, 61 and 38 different alleles were detected. The presence of specific *HLA class I* alleles such as *HLA-B*57:01* that associates with immune hypersensitivity reaction to abacavir and *HLA-B*35:05, -C*04* and *-C*08* that are associated with immune hypersensitivity reaction to nevirapine could be determined from the data collected. *HLA-B*57:01* was prevalent with allele frequencies of 8.0%, 12.0% and 8.2% in SAM, SAI and SAC individuals, respectively and absent in the SAB population, while *HLA-B*35:05* was present in 6.0% and 1.0% in SAM and SAC populations, respectively and absent in SAC and SAB populations. *HLA-C*04* was present at frequencies of 22.0%, 24.0%, 18.6% and 23.7% in SAM, SAI, SAC and SAB populations, respectively, while *HLA-C*08* was present at frequencies of 10.0%, 8.0%, 6.9% and 10.2% in SAM, SAI, SAC and SAB populations, respectively. The number of HIV-1 infected South Africans potentially at risk of developing ABC IHR and NVP IHR, was estimated at 14,235 and 2,212,700 people, respectively, from the known allele frequencies and available demographic data for the South African population. Since treatment of HIV-1 infected individuals is underway on a large scale in South Africa and both ABC and NVP are in current use, the substantial number of individuals that may be at risk warranted the development of a novel real-time AS-PCR assay for the detection of the alleles associated with immune hypersensitivity reaction.

Susceptibility and resistance to HIV-1 infection has been associated with particular *HLA class I* alleles and specific *HLA class I* alleles have also been shown to associate with HIV-1 viral load. However, even though singular alleles can be classified as favourable or deleterious, humans carry 6 *HLA class I* alleles and therefore the impact on immunity must be additive. To address this, a scoring system was used to determine whether the *HLA class I* allelic variants present in an individual were cumulatively protective, deleterious or neutral. This was done for *HLA class I* alleles reported to be associated with either susceptibility or resistance to HIV-1 infection and also for alleles reported to be associated with a high or a low viral load. The analysis showed no significant differences between the populations for

overall *HLA* genotypes considered resistant (44.0%, 32.0%, 39.4% and 30.2% in SAI, SAM, SAC and SAB populations, respectively), however, the frequency of overall *HLA* genotypes considered susceptible was significantly higher in the SAB population compared to the other populations (32.0%, 46.0%, 40.4% and 57.9% in SAI, SAM, SAC and SAB populations, respectively). For genotypes associated with viral load, the analysis determined that overall HLA genotypes associated with a low viral load was significantly higher in the SAI population compared to the other populations (66.0%, 44.0%, 39.4% and 40.3% in SAI, SAM, SAC and SAB populations, respectively, while overall HLA genotypes associated with a high viral load were significantly lower in the SAI population compared to the other populations (16.0%, 24.0%, 33.0% and 29.6% in SAI, SAM, SAC and SAB populations, respectively). This result on its own suggested that based on *HLA class I* genotype, the SAB population is the most susceptible to HIV-1 infection, and that the SAI population has the lowest likelihood of transmitting HIV-1 due to genotypes associated with low viral loads. However, the effects of other host genetic factors in modulating HIV-1 susceptibility or viral load were not considered.

Additional genes or gene combinations reported to be associated with HIV-1 acquisition/viral load were included in a broader cumulative analysis incorporating *HLA class I* genotype, *KIR* genotype, *CCR5Δ32* variants, *HLA-Bw4*^{80I}*/Bw4*^{80T}*/C1/C2* ligands, *HLA-C 3'UTR* +263 insertion/deletion variants and the p2-M/T variant of the HLA-B signal peptide. Results from this regarding HIV-1 acquisition suggested that the SAB population may be significantly more susceptible to HIV-1 infection in comparison to the SAI population (SAB vs. SAI, 42.8% vs. 14.0%, p=0.0002), while the SAI population may be significantly more resistant to HIV-1 infection in comparison to the SAB population (SAI vs. SAB, 68.0% vs. 31.4%, p<0.0001). Similarly, results regarding HIV-1 VLS suggested that in the SAB population, significantly more individuals with overall genotypes associated with high HIV-1 VLS may be present compared to the SAI population (SAB vs. SAI, 59.7% vs. 30.0%, p=0.0003) and significantly more individuals with genotypes associated with low HIV-1 VLS may be present in the SAI population in comparison to the SAB population (SAI vs. SAB, 48.0% vs. 13.2%, p<0.0001). Based on extrapolation from data analysed in healthy individuals, it was concluded that host genetic differences regarding HIV-1 resistance/susceptibility and high/low VLS may contribute to the differential HIV-1 prevalence rates seen in the four South African populations. This hypothesis naturally requires further testing in HIV-1 infected populations for verification.

The contribution of host genotype to the immune response directed towards HIV-1, once infected, is only representative of a component of the complex and dynamic relationship between the host and the virus and the influence of environmental factors. The role of HLA class I molecules in mediating viral evolution was studied in HIV-1 infected individuals, in the context of two alleles, *HLA-B*58:01* and *-B*58:02,* found in all of the populations, but most common in the SAB population. *HLA-B*58:01* is present at allele frequencies of 5.0%, 2.0%, 2.1% and 8.8% in SAI, SAM, SAC and SAB populations, respectively, while *HLA-B*58:02* is present at allele frequencies of 1.0%, 5.0%, 0.5% and 9.4% in SAI, SAM, SAC and SAB populations, respectively. Many studies show that *HLA-B*58:01* is a protective allele associated with low viral load and higher $CD4^+$ T cell count, while *HLA-B*58:02* associates with fast disease progression. However, unlike the presence of *HLA class I* alleles that are positively predictive of adverse drug reactions, carriage of *HLA-B*58:01* or *-B*58:02* does not necessarily serve as an absolute predictor of slow or fast disease progression. *HLA-B*58:01* can be present in individuals who do not present with low viral load and higher $CD4^+$ T cell counts. It appears that all untreated HIV-1-positive individuals ultimately reach the stage of disease when treatment is required, but it is the time period that is variable. Since viral mutagenesis causes the host immune response to become compromised, it is the host genetic makeup, particularly *HLA class I* genotype, which dictates the length of time the virus takes to accumulate the genetic alterations necessary to overwhelm the host.

To establish whether patterns of immune escape might be different between HLA-*B*58:01* and *-B*58:02*-positive individuals at the stage of disease when treatment is required, mutations in the Gag protein were investigated in *HLA-B*58:01* (n=5) and *-B*58:02* (n=13) carriers from an untreated HIV-1 infected SAB cohort. *Gag* sequences were generated from viral RNA extracted from plasma samples at baseline, 6 months and 12 months that were available prior to initiation of treatment. The study presented an opportunity for the novel design of the PCR primers needed for reverse transcription of viral RNA, DNA PCR and sequencing of the *gag* gene. An analysis of deduced Gag protein sequences was done to identify amino acid changes. Epitopes known to be targeted by HLA-B*58:01-restricted CTLs were examined and mutations were identified in the ISW9, TW10, YL9 and QW9 epitopes in $p24^{Gag}$. No HLA-B*58:02-restricted CTL epitopes were evident. CTL escape mutations often reduce viral replicative capacity and known compensatory mutations, H219Q, I223 and M228I/L in $p24^{Gag}$ were detected in both *HLA-B*58:01* and *-B*58:02* carriers. As other investigators have shown previously, disease progression in *HLA-B*58:01*

carriers associates with the presence of CTL escape mutations in *HLA-B*58:01*-restricted Gag epitopes and these changes are accompanied by compensatory mutations that restore viral fitness, while in contrast, *HLA-B*58:02* carriers do not target Gag. Escape mutations in other HLA class I-restricted epitopes were present in $p17^{Gag}$ in both *HLA-B*58:01* and -*B*58:02* carriers illustrating the additive effect of immune pressure by other HLA class I molecules, but these responses were subsequently lost. Mutations in the $p24^{Gag}$ TW10 epitope that affects KIR3DL1 binding and mutations in the YL9 epitope that affects KIR2DL2 were present in both *HLA-B*58:01* and *-B*58:02* carriers and may represent escape mutations from NK cells. In general, escape mutations from CTLs/NK cells were more frequent in *HLA-B*58:01* progressors than in *HLA-B*58:02* progressors. This is possibly due to the more frequent linkage of *HLA-B*58:01* with other protective *HLA class I* alleles. Due to increased CTL-mediated immune pressure, more frequent selection of CTL escape mutations that impose fitness costs may occur, necessitating genetic alterations in Gag to overcome these defects. Although the duration of infection was not known, in both *HLA-B*58:01* and -*B*58:02* carriers, disease progression was likely a consequence of accumulating mutations mediating escape from multiple immune targets.

In conclusion, the information presented in this thesis gives a broader insight into the immune mechanisms influencing HIV-1 acquisition and VLS determination and demonstrates how protective/deleterious immunogenetic factors can vary between populations. Extrapolation from studies in healthy populations suggested that host genetic differences between populations may affect HIV-1 prevalence. The importance of HLA class I alleles associated with immune hypersensitivity was highlighted. Lastly, the protective effect of HLA class I alleles associated with slow disease progression were shown to be lost due to viral escape mutations and allelic variation may simply influence the time it takes to progress. The data presented here could be useful for the development of novel epitope-based vaccines that take advantage of both innate and adaptive immune defense mechanisms, however host genetic differences between the populations in South Africa must be considered.

Appendices

Appendix A

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M130873

NAME: (Principal Investigator)	Mr Shayne Loubser
DEPARTMENT:	Centre for HIV and STIs/Virology National Institute for Communicable Diseases
PROJECT TITLE:	The Multiple Roles of Human Leukocyte Antigen (HLA) in Human Immunodeficiency Virus Type 1 (HIV-1) Immunity and Treatment
DATE CONSIDERED:	Ad hoc
DECISION:	Approved unconditionally
CONDITIONS:	
SUPERVISOR:	Prof C Tiemessen
APPROVED BY:	Professor PE Cleaton-Jones, Chairperson, HREC (Medical)

DATE OF APPROVAL: 04/09/2013

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Secretary in Room 10004, 10th floor, Senate House, University.

I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit the application to the Committee. I agree to submit a yearly progress report.

Principal Investigator Signature M130873Date

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

Appendix B

List of studies associating *HLA class I* alleles with HIV-1 acquisition.

HLA-A	Study	Country
01	Peterson *et al.*, 2013	Kenya
01:01	Liu *et al.*, 2012; Koehler *et al.*, 2010; Mori *et al.*, 2014	China; Tanzania; Thailand
02	Shankarkumar *et al.*, 2009; Singh *et al.*, 2008	India; India
02:01	Singh *et al.*, 2008; Koehler *et al.*, 2010	India; Tanzania
02:02	Koehler *et al.*, 2010; Macdonald *et al.*, 2000	Tanzania; Kenya
02:05	Koehler *et al.*, 2010; Macdonald *et al.*, 2000	Tanzania; Kenya
02:11	Shankarkumar *et al.*, 2012	India
02:14	Macdonald *et al.*, 2000	Kenya
03	Zhang *et al.*, 2013	China
03:01	Koehler *et al.*, 2001; Liu *et al.*, 2012	Tanzania; China
11	Beyrer *et al.*, 1999; Selvaraj *et al.*, 2006; Sriwanthana *et al.*, 2001	Thailand; India; Thailand
11:01	Raghavan *et al.*, 2009; Sriwanthana *et al.*, 2001	Thailand; Thailand; India; Thailand
23:01	Peterson *et al.*, 2013; Macdonald *et al.*, 2000; Koehler *et al.*, 2010; Olvera *et al.*, 2015	Kenya; Kenya; Tanzania; Peru
24	Singh *et al.*, 2008	India
24:02	Macdonald *et al.*, 2000	Kenya
24:07	Umapathy *et al.*, 2007	India
29:02	Koehler *et al.*, 2010	Tanzania
30:01	Koehler *et al.*, 2010	Tanzania
30:02	Koehler *et al.*, 2010	Tanzania
31	Figueiredo *et al.*, 2008	Brazil
32	Habegger *et al.*, 2013	Argentina
32:04	Umapathy *et al.*, 2007	India
34:02	Koehler *et al.*, 2010	Tanzania
36	Tang *et al.*, 2008	Zambia
36:01	Koehler *et al.*, 2010; Tang *et al.*, 2008	Tanzania; Zambia
66:01	Koehler *et al.*, 2010	Tanzania
68:01	Shankarkumar *et al.*, 2003; Umapathy *et al.*, 2007	India; India
68:02	Macdonald *et al.*, 2000; Song *et al.*, 2011; Koehler *et al.*, 2010	Kenya; Zambia; Tanzania
74:01	Koehler *et al.*, 2010	Tanzania
HLA-B	Study	Country
07	Xu *et al.*, 2004	China
07:02	Peterson *et al.*, 2013; Koehler *et al.*, 2010	Kenya; Tanzania
07:05	Peterson *et al.*, 2013	Kenya
08:01	Shankarkumar *et al.*, 2012; Koehler *et al.*, 2010; Shankarkumar *et al.*, 2009	India; Tanzania; India
08:04	Shankarkumar *et al.*, 2009	India
13	Zhang et al., 2013	China
13:01	Shankarkumar *et al.*, 2003	India
14	Song *et al.*, 2011	Zambia
15	Umapathy *et al.*, 2007	India
15:01	Liu *et al.*, 2012; Huang *et al.*, 2009; Umapathy *et al.*, 2007	China; China; India
15:03	Koehler *et al.*, 2007	Tanzania
15:10	Koehler *et al.*, 2007	Tanzania
18	Beyrer *et al.*, 1999; Chaudhari *et al.*, 2013	Thailand; India
18:01	Chaudhari *et al.*, 2013; Macdonald *et al.*, 2000; Shankarkumar *et al.*, 2003; Koehler *et al.*, 2010	India; Kenya; India; Tanzania
27	Gao *et al.*, 2010	USA/Europe
35	Chaudhari *et al.*, 2013; Gao *et al.*, 2010; Keynan *et al.*, 2015; Singh *et al.*, 2008; Umapathy *et al.*, 2007; Xu *et al.*, 2004	India; USA/Europe; Canada; India; India; China
35:01	Koehler *et al.*, 2010; Peterson *et al.*, 2013	China; Tanzania; Kenya
35:02	Peterson *et al.*, 2013	China; Kenya
35:03	Peterson *et al.*, 2013	China; Kenya
35:20	Shankarkumar *et al.*, 2003; Umapathy *et al.*, 2007	India; India
35:43	Olvera *et al.*, 2015	Peru
37	Huang *et al.*, 2009	China
39	Chaudhari *et al.*, 2013	India
40	Chaudhari *et al.*, 2013; Selvaraj *et al.*, 2006	India; India
40:01	Chaudhari *et al.*, 2013; Shankarkumar *et al.*, 2003;	India; India
40:02	Olvera et al., 2015	Zimbabwe
40:06	Chaudhari *et al.*, 2013; Raghavan *et al.*, 2009	India; India
41	Figueiredo *et al.*, 2008	Brazil
42:01	Peterson *et al.*, 2013; Song *et al.*, 2011; Koehler *et al.*, 2010	Kenya; Zambia; Tanzania
44	Habegger *et al.*, 2013; Song *et al.*, 2011; Xu *et al.*, 2004	Argentina; China
44:02	Chaudhari *et al.*, 2013	India
45:01	Koehler *et al.*, 2010	Tanzania
50:01	Chaudhari *et al.*, 2013	India
46	Huang *et al.*, 2009; Xu *et al.*, 2004	China; China

49:01	Koehler *et al.*, 2010	Tanzania
51	Peterson *et al.*, 2013; Keynan *et al.*, 2015; Singh *et al.*, 2008; Song *et al.*, 2011; Umapathy *et al.*, 2007	Kenya; Canada; India; Zambia; India
51:01	Liu *et al.*, 2012; Shankarkumar *et al.*, 2012; Umapathy *et al.*, 2007; Shankarkumar *et al.*, 2009	China; India; India; India
51:02	Shankarkumar *et al.*, 2009	India
51:04	Shankarkumar *et al.*, 2009	India
52	Huang *et al.*, 2009	China
52:01	Liu et al., 2012; Shankarkumar *et al.*, 2009	China; India
53	Gao *et al.*, 2010	USA/Europe
53:01	Koehler *et al.*, 2010; Peterson *et al.*, 2013	Tanzania; Kenya
54:01	Shankarkumar *et al.*, 2009	India
55	Xu *et al.*, 2004	China
55:01	Habegger *et al.*, 2013; Shankarkumar *et al.*, 2012; Shankarkumar *et al.*, 2009	Argentina; India; India
55:02	Shankarkumar *et al.*, 2009	India
56:01	Peterson *et al.*, 2013; Shankarkumar *et al.*, 2012; Shankarkumar *et al.*, 2009; Umapathy *et al.*, 2007	Kenya; India; India; India
56:03	Shankarkumar *et al.*, 2009	India
56:04	Shankarkumar *et al.*, 2009	India
57	Gao *et al.*, 2010; Tang *et al*., 2008	USA/Europe; Zambia
57:01	Umapathy *et al.*, 2007; Shankarkumar *et al.*, 2012; Shankarkumar *et al.*, 2009	India; India; India
57:02	Shankarkumar *et al.*, 2009	India
57:03	Koehler *et al.*, 2010; Olvera *et al.*, 2015	Tanzania; Peru
57:04	Shankarkumar *et al.*, 2009	India
58:01	Koehler *et al.*, 2010; Shankarkumar *et al.*, 2009	Tanzania; India
58:02	Tang *et al.*, 2010; Koehler *et al.*, 2010; Shankarkumar *et al.*, 2009	Zambia; Tanzania; India
67	Huang *et al.*, 2009	China
78	Xu *et al.*, 2004	China
81:01	Koehler *et al.*, 2010	Tanzania
HLA-C	Study	Country
02:10	Peterson *et al.*, 2013; Koehler *et al.*, 2010	Kenya; Tanzania
03	Zhang et al., 2013	China
03:04	Koehler *et al.*, 2010	Tanzania
04:01	Koehler *et al.*, 2010	Tanzania
05:01	Umapathy *et al.*, 2007	India
06	Zhang *et al.*, 2013	China
06:02	Peterson *et al.*, 2013; Koehler *et al.*, 2010	Kenya; Tanzania
07:01	Peterson *et al.*, 2013; Koehler *et al.*, 2010	Kenya; Tanzania
07:02	Koehler *et al.*, 2010; Peterson *et al.*, 2013	Tanzania; Kenya
07:04	Koehler *et al.*, 2010	Tanzania
08:01	Umapathy *et al.*, 2007	India
08:02	Koehler *et al.*, 2010	Tanzania
12:02	Umapathy *et al.*, 2007	India
15	Umapathy *et al.*, 2007	India
15:02	Liu *et al.*, 2012	China
15:07	Shankarkumar *et al.*, 2003; Umapathy *et al.*, 2007	India; India
16:01	Koehler *et al.*, 2010	Tanzania
17:01	Koehler *et al.*, 2010	Tanzania
18	Tang *et al.*, 2008	Zambia
18:01	Tang *et al.*, 2010; Koehler *et al.*, 2010	Zambia; Tanzania
18:02	Tang *et al.*, 2010	Zambia

Black ink = resistant; red ink = susceptible; yellow highlight = study shows opposite association; *HLA*, human leukocyte antigen.

Appendix C

List of studies associating *HLA class I* alleles with viral load setpoint in HIV-1 infection.

HLA-A	Study	Country
01	Kiepiela *et al.*, 2004; Leslie *et al.*, 2010; Naruto *et al.*, 2012; Zhang *et al.*, 2013	RSA; RSA; Japan; China
01:01	Kiepiela *et al.*, 2004; Leslie *et al.*, 2010; Naruto *et al.*, 2012; Mori *et al.*, 2014	RSA; RSA; Japan; Thailand
02:01	Leslie *et al.*, 2010; Naruto *et al.*, 2012; Kiepiela *et al.*, 2004; Mori *et al.*, 2014; Novitsky *et al.*, 2003; Olvera *et al.*, 2015; Zhang *et al.*, 2013	RSA; Japan; RSA; Thailand; Botswana; Peru; China
02:02	Novitsky *et al.*, 2003	Botswana
02:03	Novitsky *et al.*, 2003; Mori *et al.*, 2014	Botswana; Thailand
02:04	Novitsky *et al.*, 2003	Botswana
02:05	Novitsky *et al.*, 2003; Leslie *et al.*, 2010; Kiepiela *et al.*, 2004; Saathof *et al.*, 2010	Botswana; RSA; RSA; Tanzania
02:06	Novitsky *et al.*, 2003; Naruto *et al.*, 2012; Mori *et al.*, 2014	Botswana; Japan; Thailand
02:07	Novitsky *et al.*, 2003; Mori *et al.*, 2014; Naruto *et al.*, 2012	Botswana; Thailand; Japan
02:22	Olvera *et al.*, 2015	Peru
03	Matthews *et al.*, 2012; Zhang *et al.*, 2013	RSA/Botswana; China
03:01	Carlson *et al.*, 2012; Kiepiela *et al.*, 2004	RSA/Botswana/Malawi/Zimbabwe; RSA
11	Zhang *et al.*, 2013	China
11:01	Mori *et al.*, 2014; Naruto *et al.*, 2012; Olvera *et al.*, 2015	Thailand; Japan; Peru
11:02	Mori *et al.*, 2014	Thailand
23	Zhang *et al.*, 2013	China
23:01	Kiepiela *et al.*, 2004; Novitsky *et al.*, 2003; Mackelprang *et al.*, 2015	RSA; Botswana; Botswana/RSA/Zambia/Kenya/Rwanda/Tanzania/Uganda
24	Zhang *et al.*, 2013	China
24:02	Kiepiela *et al.*, 2004; Naruto *et al.*, 2012; Mori *et al.*, 2014; Novitsky *et al.*, 2003	RSA; Japan; Thailand; Botswana
24:03	Novitsky *et al.*, 2003	Botswana
24:04	Novitsky *et al.*, 2003	Botswana
24:07	Mori *et al.*, 2014	Thailand
24:10	Mori *et al.*, 2014	Thailand
25	Int. HIV controllers study *et al.*, 2010; Catano *et al.*, 2008	USA/Europe; USA
26	Zhang et al., 2013	China
26:01	Leslie *et al.*, 2010; Naruto *et al.*, 2012; Mori *et al.*, 2014	RSA; Japan; Thailand
26:02	Naruto *et al.*, 2012	Japan
26:03	Naruto *et al.*, 2012	Japan
29	Kiepiela *et al.*, 2004; Tang *et al.*, 2010; Zhang *et al.*, 2013	RSA; Zambia; China
30	Zhang et al., 2013	China
30:01	Carlson *et al.*, 2012; Kiepiela *et al.*, 2004; Mori *et al.*, 2014; Novitsky *et al.*, 2003; Mackelprang *et al.*, 2015	RSA/Botswana/Malawi/Zimbabwe; RSA; Thailand; Botswana; Botswana/RSA/Zambia/Kenya/Rwanda/Tanzania/Uganda
30:02	Carlson *et al.*, 2012; Kiepiela *et al.*, 2004; Novitsky *et al.*, 2003	RSA/Botswana/Malawi/Zimbabwe; RSA; Botswana
30:03	Novitsky *et al.*, 2003	Botswana
30:04	Leslie *et al.*, 2010	RSA
31	Zhang *et al.*, 2013	China
31:01	Carlson *et al.*, 2012; Naruto *et al.*, 2012; Mori *et al.*, 2014	RSA/Botswana/Malawi/Zimbabwe; Japan; Thailand
32	Zhang *et al.*, 2013	China
32:01	Lazaryan *et al.*, 2011; Mori *et al.*, 2014	USA(African-American); Thailand
33	Zhang *et al.*, 2013	China
33:03	Leslie *et al.*, 2010; Matthews *et al.*, 2012; Carlson *et al.*, 2012; Naruto *et al.*, 2012; Mori *et al.*, 2014; Olvera *et al.*, 2015; Mackelprang *et al.*, 2015	RSA, RSA/Botswana; RSA/Botswana/Malawi/Zimbabwe; Japan; Thailand; Peru; Botswana/RSA/Zambia/Kenya/Rwanda/Tanzania/Uganda
34:01	Mori *et al.*, 2014; Catano *et al.*, 2008	Thailand; USA
34:02	Carlson *et al.*, 2012; Leslie *et al.*, 2010; Kiepiela *et al.*, 2004	RSA/Botswana/Malawi/Zimbabwe; RSA; RSA
36	Tang *et al.*, 2008; Tang *et al.*, 2010; Zhang *et al.*, 2013	Zambia; Zambia; China
36:01	Carlson *et al.*, 2012; Tang *et al.*, 2002	RSA/Botswana/Malawi/Zimbabwe; Zambia
43:01	Kiepiela *et al.*, 2004; Catano *et al.*, 2008	RSA; USA(Caucasian)
66	Kiepiela *et al.*, 2004; Matthews *et al.*; 2012; Zhang *et al.*, 2013	RSA; RSA/Botswana; China
66:01	Leslie *et al.*, 2010; Carlson *et al.*, 2012; Catano *et al.*, 2008	RSA, RSA/Botswana/Malawi/Zimbabwe; USA(Caucasian)

66:03	Carlson *et al.*, 2012	RSA/Botswana/Malawi/Zimbabwe
68	Tang *et al.*, 2002; Zhang et al., 2013	Zambia; China
68:01	Leslie *et al.*, 2010; Matthews *et al.*, 2012; Kiepiela *et al.*, 2004; Carlson *et al.*, 2012; Mori *et al.*, 2014	RSA; RSA/Botswana; RSA; RSA/Botswana/Malawi/Zimbabwe; Thailand
68:02	Novitsky *et al.*, 2003; Kiepiela *et al.*, 2004	Botswana; RSA
69	Zhang *et al.*, 2013	China
69:01	Novitsky *et al.*, 2003	Botswana
74	Mathews *et al.*, 2012; Tang *et al.*, 2010	RSA/Botswana; Zambia
74:01	Carlson *et al.*, 2012; Kiepiela *et al.*, 2004; Lazaryan *et al.*, 2011; Mori *et al.*, 2014	RSA/Botswana/Malawi/Zimbabwe; RSA; USA(African-American); Thailand
80	Mathews *et al.*, 2012	RSA/Botswana
HLA-B	Study	Country
07	Emu *et al.*, 2008	USA
07:02	Geldmacher *et al.*, 2007; Leslie *et al.*, 2010; Kiepiela *et al.*, 2004; Naruto *et al.*, 2012; Mclaren *et al.*, 2012	Tanzania; RSA, RSA; Japan; USA(African-American)
07:05	Leslie *et al.*, 2010; Mclaren *et al.*, 2012; Mori *et al.*, 2014	RSA; USA(African-American); Thailand
08	Matthews *et al.*, 2012	RSA/Botswana
08:01	Leslie *et al.*, 2010; Kiepiela *et al.*, 2004; Carlson *et al.*, 2012; Mclaren *et al.*, 2012; Mori *et al.*, 2014; Olvera *et al.*, 2015	RSA; RSA; RSA/Botswana/Malawi/Zimbabwe; USA(African-American); Thailand; Peru
13	Emu *et al.*, 2008; Honeyborne *et al.*, 2007; Mathews *et al.*, 2012; Tang *et al.*, 2002; Tang *et al.*, 2010; Zhang *et al.*, 2013	USA; RSA; RSA/Botswana; Zambia; Zambia; China
13:01	Brockman *et al.*, 2010; Leslie *et al.*, 2010; Naruto *et al.*, 2012; Mori *et al.*, 2014	USA/Australia/Germany; RSA; Japan; Thailand
13:02	Mclaren *et al.*, 2012; Mori *et al.*, 2014	USA(African-American); Thailand
14	Kiepiela *et al.*, 2004; Lazaryan *et al.*, 2011; Mathews *et al.*, 2012	RSA; USA(African-American); RSA/Botswana
14:01	Lazaryan *et al.*, 2011; Mathews *et al.*, 2012; Novitsky *et al.*, 2003; Mclaren *et al.*, 2012; Mackelprang *et al.*, 2015	USA(African-American); RSA/Botswana; Botswana; USA(African-American); Botswana/RSA/Zambia/Kenya/Rwanda/Tanzania/Uganda
14:02	Lazaryan *et al.*, 2011; Mathews2012; Novitsky2003; Mclaren *et al.*, 2012; Mori *et al.*, 2014	USA(African-American); RSA/Botswana; Botswana; USA(African-American); Thailand
15:01	Mathews *et al.*, 2012; Naruto *et al.*, 2012; Mori *et al.*, 2014; Mclaren *et al.*, 2012; Olvera *et al.*, 2015	RSA/Botswana; Japan; Thailand; USA(African-American); Peru
15:02	Mori *et al.*, 2014	Thailand
15:03	Novitsky *et al.*, 2003; Geldmacher *et al.*, 2007; Leslie *et al.*, 2010; Kiepiela *et al.*, 2004; Saathof *et al.*, 2010; Frahm *et al.*, 2006; Mathews *et al.*, 2012; Mclaren *et al.*, 2012	Botswana; Tanzania; RSA; RSA; Tanzania; USA/RSA; RSA/Botswana; USA(African-American)
15:04	Mori *et al.*, 2014	Thailand
15:07	Mori *et al.*, 2014	Thailand
15:09	Novitsky *et al.*, 2003	Botswana
15:10	Novitsky *et al.*, 2003; Leslie *et al.*, 2010; Kiepiela *et al.*, 2004; Mathews *et al.*, 2012; Mclaren *et al.*, 2012	Botswana; RSA; RSA; RSA/Botswana; USA(African-American)
15:11	Naruto *et al.*, 2012	Japan
15:12	Mori *et al.*, 2014	Thailand
15:16	Frahm *et al.*, 2005; Carlson *et al.*, 2012; Lazaryan *et al.*, 2006; Mclaren *et al.*, 2012; Novitsky *et al.*, 2003; Olvera *et al.*, 2015	USA/RSA; RSA/Botswana/Malawi/Zimbabwe Zambia/Rwanda; USA(African-American); Botswana; Peru
15:17	Frahm *et al.*, 2005; Carlson *et al.*, 2012; Lazaryan *et al.*, 2006; Novitsky *et al.*, 2003	USA/RSA; RSA/Botswana/Malawi/Zimbabwe Zambia/Rwanda; Botswana
15:18	Naruto *et al.*, 2012; Novitsky *et al.*, 2003	Japan; Botswana
15:21	Mori *et al.*, 2014	Thailand
15:25	Mori *et al.*, 2014	Thailand
15:27	Mori *et al.*, 2014	Thailand
15:35	Mori *et al.*, 2014	Thailand
16	Mathews *et al.*, 2012	RSA/Botswana
18	Kiepiela *et al.*, 2004; Li *et al.*, 2014; Matthews *et al.*, 2012; Novitsky *et al.*, 2003	RSA; Kenya/Rwanda/Uganda/Zambia; RSA/Botswana; Botswana
18:01	Carlson *et al.*, 2012; Geldmacher *et al.*, 2007; Leslie *et al.*, 2010; Saathof *et al.*, 2010; Mclaren *et al.*, 2012; Mori *et al.*, 2014; Olvera *et al.*, 2015	RSA/Botswana/Malawi/Zimbabwe; Tanzania; RSA; Tanzania; USA(African-American); Thailand; Peru
18:02	Mori *et al.*, 2014	Thailand
24:02	Mori *et al.*, 2014	Thailand
24:07	Gao *et al.*, 2010	USA/Europe
24:10	Mori *et al.*, 2014	Thailand
26:01	Mori *et al.*, 2014	Thailand
27	Gao *et al.*, 2010	USA/Europe
27:01	Novitsky *et al.*, 2003	Botswana
27:02	Novitsky *et al.*, 2003	Botswana
27:03	Novitsky *et al.*, 2003; Mackelprang *et al.*, 2015	Botswana; Botswana/RSA/Zambia/Kenya/Rwanda/Tanzania/Uganda
27:04	Novitsky *et al.*, 2003; Mori *et al.*, 2014	Botswana; Thailand
27:05	Novitsky *et al.*, 2003; Gao *et al.*, 2010; Int. HIV	Botswana; USA/Europe; USA/European; USA(African-

	controllers study *et al.*, 2010; Mclaren *et al.*, 2012	American)
27:06	Novitsky *et al.*, 2003; Mori *et al.*, 2014	Botswana; Thailand
27:07	Novitsky *et al.*, 2003	Botswana
27:08	Novitsky *et al.*, 2003	Botswana
33	Mathews *et al.*, 2012	RSA/Botswana
34:01	Mori *et al.*, 2014	Thailand
35	Gao *et al.*, 2010; Int. HIV controllers study *et al.*, 2010	USA/Europe; USA/Europe
35:01	Carlson *et al.*, 2012; Lazaryan *et al.*, 2011; Naruto *et al.*, 2012; Mclaren *et al.*, 2012; Mori *et al.*, 2014; Olvera *et al.*, 2015	RSA/Botswana/Malawi/Zimbabwe; USA(African-American); Japan; USA(African-American); Thailand; Peru
35:02	Carlson *et al.*, 2012; Lazaryan *et al.*, 2011	RSA/Botswana/Malawi/Zimbabwe; USA(African-American)
35:03	Carlson *et al.*, 2012; Lazaryan *et al.*, 2011; Mori *et al.*, 2014	RSA/Botswana/Malawi/Zimbabwe; Thailand USA(African-American)
35:05	Carlson *et al.*, 2012; Lazaryan *et al.*, 2011; Mori *et al.*, 2014; Olvera *et al.*, 2015	RSA/Botswana/Malawi/Zimbabwe; USA(African-American); Thailand; Peru
35:09	Olvera *et al.*, 2015	Peru
37:01	Novitsky *et al.*, 2003; Mclaren *et al.*, 2012; Naruto *et al.*, 2012	Botswana; USA(African-American); Japan
38:01	Novitsky *et al.*, 2003	Botswana
38:02	Novitsky *et al.*, 2003; Mori *et al.*, 2014	Botswana; Thailand
39	Mathews *et al.*, 2012; Tang *et al.*, 2002	RSA/Botswana; Zambia
39:01	Mori *et al.*, 2014; Naruto *et al.*, 2012	Thailand; Japan
39:02	Novitsky *et al.*, 2003	Botswana
39:03	Novitsky *et al.*, 2003; Olvera *et al.*, 2015	Botswana; Peru
39:04	Novitsky *et al.*, 2003	Botswana
39:06	Tang *et al.*, 2002	Zambia
39:10	Mathews *et al.*, 2012; Tang *et al.*, 2002; Leslie *et al.*, 2010; Carlson *et al.*, 2012; Mclaren et al., 2012	RSA/Botswana; Zambia; RSA; RSA/Botswana/Malawi/Zimbabwe; USA(African-American)
39:13	Olvera *et al.*, 2015	Peru
39:14	Olvera *et al.*, 2015	Peru
40:01	Mori *et al.*, 2014; Naruto *et al.*, 2012; Mclaren *et al.*, 2012; Novitsky *et al.*, 2003	Thailand; Japan; USA(African-American); Botswana
40:02	Mori *et al.*, 2014; Naruto *et al.*, 2012	Thailand; Japan
40:04	Olvera *et al.*, 2015	Peru
40:06	Mori *et al.*, 2014; Naruto *et al.*, 2012; Novitsky *et al.*, 2003	Thailand; Japan; Botswana
41	Tang *et al.*, 2002	Zambia
41:01	Leslie *et al.*, 2010; Novitsky *et al.*, 2003	RSA; Botswana
41:02	Mclaren *et al.*, 2012	USA(African-American)
42	Matthews *et al.*, 2012	RSA/Botswana
42:01	Leslie *et al.*, 2010; Matthews *et al.*, 2012; Kiepiela *et al.*, 2004; Carlson *et al.*, 2012; Saathof *et al.*, 2010; Mclaren *et al.*, 2012	RSA; RSA/Botswana; RSA; RSA/Botswana/Malawi/Zimbabwe; Tanzania; USA(African-American)
42:02	Mclaren *et al.*, 2012	USA(African-American)
44	Matthews *et al.*, 2012	RSA/Botswana
44:02	Zhang *et al.*, 2013; Mclaren *et al.*, 2012; Novitsky *et al.*, 2003	China; USA(African-American); Botswana
44:03	Leslie *et al.*, 2010; Kiepiela *et al.*, 2004; Carlson *et al.*, 2012; Zhang *et al.*, 2013; Naruto *et al.*, 2012; Mclaren *et al.*, 2012; Mori *et al.*, 2014; Novitsky *et al.*, 2003	RSA; RSA; RSA/Botswana/Malawi/Zimbabwe; China; Japan; USA(African-American); Thailand; Botswana
44:05	Zhang *et al.*, 2013	China
44:29	Zhang *et al.*, 2013	China
45	Lazaryan *et al.*, 2011; Li *et al.*, 2014; Matthews *et al.*, 2012; Tang *et al.*, 2002; Tang et al., 2010	USA(African-American); Kenya/Rwanda/Uganda/Zambia; RSA/Botswana; Zambia; Zambia
45:01	Leslie *et al.*, 2010; Kiepiela *et al.*, 2004; Carlson *et al.*, 2012; Saathof *et al.*, 2010; Geldmacher *et al.*, 2007; Mclaren *et al.*, 2012	RSA; RSA; RSA/Botswana/Malawi/Zimbabwe; Tanzania; Tanzania; USA(African-American)
46:01	Mori *et al.*, 2014; Naruto *et al.*, 2012	Thailand; Japan
48:01	Novitsky *et al.*, 2003; Mori *et al.*, 2014; Naruto *et al.*, 2012	Botswana; Thailand; Japan
48:02	Novitsky *et al.*, 2003	Botswana
49:01	Leslie *et al.*, 2010; Mclaren *et al.*, 2012; Novitsky *et al.*, 2003	RSA; USA(African-American); Botswana
50:01	Mclaren *et al.*, 2012; Novitsky *et al.*, 2003	USA(African-American); Botswana
50:02	Carlson *et al.*, 2012	RSA/Botswana/Malawi/Zimbabwe
51	Leslie *et al.*, 2010	Japan; RSA; RSA/Botswana/Malawi/Zimbabwe
51:01	Naruto *et al.*, 2012; Carlson *et al.*, 2012; Mclaren *et al.*, 2012; Mori *et al.*, 2014	Japan; RSA/Botswana/Malawi/Zimbabwe; USA(African-American); Thailand
51:02	Mori *et al.*, 2014	Thailand
52	Int. HIV controllers study *et al.*, 2010; Naruto *et al.*, 2012	USA/Europe; Japan
52:01	Mclaren *et al.*, 2012; Naruto *et al.*, 2012; Mori *et al.*, 2014	USA(African-American); Japan; Thailand
53	Lazaryan *et al.*, 2011; Li *et al.*, 2014	USA(African-American); Kenya/Rwanda/Uganda/Zambia

53:01	Carlson., 2012; Mclaren *et al.*, 2012; Mackelprang *et al.*, 2015	RSA/Botswana/Malawi/Zimbabwe; USA(African-American); Botswana/RSA/Zambia/Kenya/Rwanda/Tanzania/Uganda
54	Dorak *et al.*, 2003	USA(Caucasian), Holland
54:01	Mori *et al.*, 2014; Naruto *et al.*, 2012	Thailand; Japan
55	Dorak *et al.*, 2003	USA(Caucasian), Holland
55:02	Mori *et al.*, 2014; Naruto *et al.*, 2012	Thailand; Japan
56	Dorak *et al.*, 2003	USA(Caucasian), Holland
56:01	Mori *et al.*, 2014; Naruto *et al.*, 2012	Thailand; Japan
56:02	Mori *et al.*, 2014	Thailand
57	Silva *et al.*, 2010; Emu *et al.*, 2008; *Gao et al.*, 2010; Int. HIV controllers study *et al.*, 2010; Kiepiela *et al.*, 2004; Li *et al.*, 2014; Mathews *et al.*, 2012; Tang *et al.*, 2002; Tang *et al.*, 2008; Tang *et al.*, 2010	Brazil; USA; USA/Europe; USA(African-American); RSA; Kenya/Rwanda/Uganda/Zambia; RSA/Botswana; Zambia; Zambia; Zambia
57:01	Brockman *et al.*, 2010; Leslie *et al.*, 2010; Carlson *et al.*, 2012; Lazaryan *et al.*, 2011; Matthews *et al.*, 2012; Gao *et al.*, 2010; Int. HIV controllers study *et al.*, 2010; Mclaren et al., 2012; Mori *et al.*, 2014; Novitsky *et al.*, 2003	USA/Australia/Germany; RSA; RSA/Botswana/Malawi/Zimbabwe; RSA, USA(African-American); RSA/Botswana; USA/Europe; USA/European; USA(African-American); Thailand; Botswana
57:02	Brockman *et al.*, 2010; Tang *et al.*, 2010; Leslie *et al.*, 2010; Carlson *et al.*, 2012; Lazaryan *et al.*, 2011; Matthews *et al.*, 2012; Gao *et al.*, 2010; Mclaren et al., 2012; Novitsky *et al.*, 2003	USA/Australia/Germany; Zambia; RSA; RSA/Botswana/Malawi/Zimbabwe; USA(African-American); RSA/Botswana; USA/Europe; USA(African-American); Botswana
57:03	Carlson *et al.*, 2012; Int. HIV controllers study *et al.*, 2010; Lazaryan *et al.*, 2006; Lazaryan *et al.*, 2011; Limou *et al.*, 2013; Mclaren et al., 2012; Olvera *et al.*, 2015; Saathof *et al.*, 2010	RSA/Botswana/Malawi/Zimbabwe; USA(African-American); Zambia/Rwanda; USA(African-American); USA(African-American); USA(African-American); Peru; Tanzania
58	Emu *et al.*, 2012; Novitsky *et al.*, 2003	USA; Botswana
58:01	Brockman *et al.*, 2010; Carlson *et al.*, 2012; Geldmacher *et al.*, 2007; Leslie *et al.*, 2010; Kiepiela *et al.*, 2004; Saathof *et al.*, 2010; Matthews *et al.*, 2012; Gao *et al.*, 2010; Naruto *et al.*, 2012; Mclaren *et al.*, 2012; Mori *et al.*, 2014	USA/Australia/Germany; RSA/Botswana/Malawi/Zimbabwe; Tanzania; RSA; RSA; Tanzania; RSA/Botswana; USA/Europe; Japan; USA(African-American); Thailand
58:02	Carlson *et al.*, 2012; Leslie *et al.*, 2010; Kiepiela *et al.*, 2004; Saathof *et al.*, 2010; Matthews *et al.*, 2012; Tang *et al.*, 2010; Lazaryan *et al.*, 2006; Mclaren *et al.*, 2012	RSA/Botswana/Malawi/Zimbabwe; RSA; RSA; Tanzania; RSA/Botswana; Zambia; Zambia/Rwanda; USA(African-American)
59:01	Naruto *et al.*, 2012	Japan
67	Zhang *et al.*, 2013	China
67:01	Naruto *et al.*, 2012	Japan
73:01	Novitsky *et al.*, 2003	Botswana
78:01	Mclaren *et al.*, 2012	USA(African-American)
81	Emu *et al.*, 2008; Matthews *et al.*, 2012; Tang *et al.*, 2008; Tang *et al.*, 2010	USA; RSA/Botswana; Zambia; Zambia
81:01	Carlson *et al.*, 2012; Geldmacher *et al.*, 2007; Tang *et al.*, 2010; Leslie *et al.*, 2010; Kiepiela *et al.*, 2004; Saathof *et al.*, 2010; Mclaren et al., 2012	RSA/Botswana/Malawi/Zimbabwe; Tanzania; Zambia; RSA; RSA; Tanzania; USA(African-American)

HLA-C	Study	Country
01	Zhang *et al.*, 2013	China
01:02	Naruto *et al.*, 2012; Mori *et al.*, 2014	Japan; Thailand
01:03	Naruto *et al.*, 2012	Japan
02	Matthews *et al.*, 2012	RSA/Botswana
02:02	Kiepiela *et al.*, 2004; Mackelprang *et al.*, 2015	RSA; Botswana/RSA/Zambia/Kenya/Rwanda/Tanzania/Uganda
03	Zhang *et al.*, 2013	China
03:02	Leslie *et al.*, 2010; Matthews *et al.*, 2012; Mori *et al.*, 2014; Naruto *et al.*, 2012	RSA; RSA/Botswana; Thailand; Japan
03:03	Mori *et al.*, 2014; Naruto *et al.*, 2012	Thailand; Japan
03:04	Kiepiela *et al.*, 2004; Naruto *et al.*, 2012; Mori *et al.*, 2014	RSA; Japan; Thailand
04	Matthews *et al.*, 2012	RSA/Botswana
04:01	Tang *et al.*, 2002; Leslie *et al.*, 2010; Kiepiela *et al.*, 2004; Naruto *et al.*, 2012; Mori *et al.*, 2014; Olvera et al., 2015	Zambia; RSA, RSA; Japan; Thailand; Peru
04:03	Tang *et al.*, 2002; Mori *et al.*, 2014	Zambia; Thailand
04:06	Mori *et al.*, 2014	Zambia; Thailand
05:01	Leslie *et al.*, 2010	RSA
06	Matthews *et al.*, 2012; Zhang *et al.*, 2013	RSA/Botswana; China
06:02	Leslie *et al.*, 2010; Kiepiela *et al.*, 2004; Naruto *et al.*, 2012; Mori *et al.*, 2014	RSA; RSA; Japan; Thailand
07	Int. HIV controllers study *et al.*, 2010	USA/Europe
07:01	Leslie *et al.*, 2010; Kiepiela *et al.*, 2004; Mori *et al.*, 2014	RSA; RSA; Thailand
07:02	Leslie *et al.*, 2010; Kiepiela *et al.*, 2004; Naruto *et al.*, 2012; Mori *et al.*, 2014; Olvera *et al.*, 2015	RSA; RSA; Japan; Thailand; Peru
07:04	Leslie *et al.*, 2010; Naruto *et al.*, 2012; Mori *et al.*, 2014	RSA; Japan; Thailand
08	Lazaryan *et al.*, 2011; Matthews *et al.*, 2012	USA(African-American); RSA/Botswana

08:01	Tang *et al.*, 2002; Naruto *et al.*, 2012; Leslie *et al.*, 2010; Kiepiela *et al.*, 2004; Mori *et al.*, 2014; Olvera *et al.*, 2015	Japan; RSA; RSA; Thailand; Peru
08:02	Tang *et al.*, 2002; Leslie *et al.*, 2010	Zambia; RSA
08:03	Tang *et al.*, 2002; Naruto *et al.*, 2012	Zambia; Japan
12	Matthews *et al.*, 2012; Tang *et al.*, 2002	RSA/Botswana; Zambia
12:02	Naruto *et al.*, 2012; Mori *et al.*, 2014	Japan; Thailand
12:03	Carlson *et al.*, 2012; Leslie *et al.*, 2010; Mori *et al.*, 2014	RSA/Botswana/Malawi/Zimbabwe; RSA; Thailand
14:02	Carlson *et al.*, 2012; Naruto *et al.*, 2012; Mori *et al.*, 2014	RSA/Botswana/Malawi/Zimbabwe; Japan; Thailand
14:03	Naruto *et al.*, 2012	Japan
15:02	Mori *et al.*, 2014; Naruto *et al.*, 2012	Thailand; Japan
16	Tang *et al.*, 2002	Zambia
16:01	Matthews *et al.*, 2012; Kiepiela *et al.*, 2004	RSA/Botswana; RSA
16:02	Matthews *et al.*, 2012	RSA/Botswana
17	Matthews *et al.*, 2012	RSA/Botswana
17:01	Tang *et al.*, 2002; Kiepiela *et al.*, 2004; Carlson *et al.*, 2012	Zambia; RSA; RSA/Botswana/Malawi/Zimbabwe
18	Silva *et al.*, 2010; Lazaryan *et al.*, 2011; Matthews *et al.*, 2012; Tang *et al.*, 2002; Tang *et al.*, 2008; Tang *et al.*, 2010	Brazil; USA(African-American); RSA/Botswana; Zambia; Zambia; Zambia
18:01	Kiepiela *et al.*, 2004	RSA
18:02	Kiepiela *et al.*, 2004	RSA

Black ink = low viral load setpoint; red print = high viral load setpoint; yellow highlight = study shows opposite association; *HLA*, human leukocyte antigen.

Appendix D

Turnitin Originality Report

SLoubserPhDThesisTurnItIn260215.docx by Alistair Loubser
From PhD_Research
(9T5NEoFKOF6336g28aNnCeg30dOsUdRa91SaP18F5PMsvNimlyM44I5a4kyMuzKwR1M3lV3d8wUdr5SYgJ6y7Kv6021K6br6KOM)

- Processed on 26-Feb-2015 11:38 AM SAST
- ID: 509905589
- Word Count: 48553

Similarity Index 16%
Similarity by Source
Internet Sources: 5%
Publications: 15%
Student Papers: 2%

Sources:

1

9% match (publications)
Loubser, S., M. Paximadis, N. Gentle, A. Puren, C. M. Gray, and C. T. Tiemessen. "Frequencies of immune hypersensitivity reaction-associated HLA class I alleles in healthy South African Indian and mixed ancestry populations determined by a novel real-time PCR assay : Real-time PCR assay to identify immune hypersensitivity reaction-associated HLA alleles", Tissue Antigens, 2014.

2

1% match (publications)
Goulder, PhilipÂ J.R., and BruceÂ D. Walker. "HIV and HLA Class I: An Evolving Relationship", Immunity, 2012.

3

< 1% match (Internet from 08-Sep-2014)
http://quantitativepcr.info/qpcr/tag/gene-expression/page/3/

4

< 1% match (Internet from 18-Feb-2012)
http://www.pbtfus.org/medcomm/research/papers/Witt_etal_2011_CancerCell.pdf

5

< 1% match (publications)
Olvera, Alex, Susana Pérez-Álvarez, Javier Ibarrondo, Carmela Ganoza, Javier R. Lama, Aldo Lucchetti, Steven Cate, William Hildebrand, Nicole Bernard, Lupe Gomez, Jorge Sanchez, and Christian Brander. "The HLA-C*04 : 01/KIR2DS4 gene combination and human leukocyte antigen alleles with high population frequency drive rate of HIV disease progression", AIDS, 2015.

6

< 1% match (Internet from 13-Nov-2009)
http://www.freepatentsonline.com/6045990.html

7

< 1% match (Internet from 29-Mar-2010)
http://iji.sums.ac.ir/spring2007/1stiji_vol4_no2_2007.pdf

8

< 1% match (Internet from 07-Nov-2014)
http://www.katerik.ru/print.html

9

< 1% match (Internet from 19-Oct-2011)
http://people.trentu.ca/barrysaville/Folder%20for%20Web%20Page/gene%20dis.pdf

10

< 1% match (Internet from 14-Feb-2013)
http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2519667/?tool=pmcentrez

References

ABI-RACHED, L., MOESTA, A. K., RAJALINGAM, R., GUETHLEIN, L. A. & PARHAM, P. 2010. Human-specific evolution and adaptation led to major qualitative differences in the variable receptors of human and chimpanzee natural killer cells. *PLoS Genet,* 6**,** e1001192.

ADAMS, E. J. & LUOMA, A. M. 2013. The adaptable major histocompatibility complex (MHC) fold: structure and function of nonclassical and MHC class I-like molecules. *Annu Rev Immunol,* 31**,** 529-61.

ALTER, G., HECKERMAN, D., SCHNEIDEWIND, A., FADDA, L., KADIE, C. M., CARLSON, J. M., ONIANGUE-NDZA, C., MARTIN, M., LI, B., KHAKOO, S. I., CARRINGTON, M., ALLEN, T. M. & ALTFELD, M. 2011. HIV-1 adaptation to NK-cell-mediated immune pressure. *Nature,* 476**,** 96-100.

ALTER, G., TEIGEN, N., AHERN, R., STREECK, H., MEIER, A., ROSENBERG, E. S. & ALTFELD, M. 2007. Evolution of innate and adaptive effector cell functions during acute HIV-1 infection. *J Infect Dis,* 195**,** 1452-60.

ALVES, L. G., RAJALINGAM, R. & CANAVEZ, F. 2009. A novel real-time PCR method for KIR genotyping. *Tissue Antigens,* 73**,** 188-91.

ANDERSEN, M. H., SCHRAMA, D., THOR STRATEN, P. & BECKER, J. C. 2006. Cytotoxic T cells. *J Invest Dermatol,* 126**,** 32-41.

ARHEL, N. J., SOUQUERE-BESSE, S., MUNIER, S., SOUQUE, P., GUADAGNINI, S., RUTHERFORD, S., PREVOST, M. C., ALLEN, T. D. & CHARNEAU, P. 2007. HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. *EMBO J,* 26**,** 3025-37.

ARIEN, K. K., JESPERS, V. & VANHAM, G. 2011. HIV sexual transmission and microbicides. *Rev Med Virol*.

ARTS, E. J. & HAZUDA, D. J. 2012. HIV-1 antiretroviral drug therapy. *Cold Spring Harb Perspect Med,* 2**,** a007161.

AUVERT, B., TALJAARD, D., LAGARDE, E., SOBNGWI-TAMBEKOU, J., SITTA, R. & PUREN, A. 2005. Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: the ANRS 1265 Trial. *PLoS Med,* 2**,** e298.

AYALA-GARCIA, I., HERNANDEZ-SEGURA, A. M., CASTELL-RODRIGUEZ, A., ALVAREZ PEREZ, S. J., TELLEZ, B. H. & RAMIREZ-GONZALEZ, M. D. 2005. Participation of epidermal langerhans cells in human pathology and their potential as targets for drug development: a review of literature. *Proc West Pharmacol Soc,* 48**,** 13-20.

BADULLI, C., SESTINI, R., SBARSI, I., BARONCELLI, M., PIZZOCHERO, C., MARTINETTI, M. & PORFIRIO, B. 2012. Tag SNPs of the ancestral haplotype 57.1 do not substitute HLA-B*57:01 typing for eligibility to abacavir treatment in the Italian population. *Pharmacogenomics,* 13**,** 247-9.

BALAMURUGAN, A., ALI, A., BOUCAU, J., LE GALL, S., NG, H. L. & YANG, O. O. 2013. HIV-1 gag cytotoxic T lymphocyte epitopes vary in presentation kinetics relative to HLA class I downregulation. *J Virol,* 87**,** 8726-34.

BARIN, F., M'BOUP, S., DENIS, F., KANKI, P., ALLAN, J. S., LEE, T. H. & ESSEX, M. 1985. Serological evidence for virus related to simian T-lymphotropic retrovirus III in residents of west Africa. *Lancet,* 2**,** 1387-9.

BARMANIA, F., POTGIETER, M. & PEPPER, M. S. 2013. Mutations in C-C chemokine receptor type 5 (CCR5) in South African individuals. *Int J Infect Dis,* 17**,** e1148-53.

BARRE-SINOUSSI, F. 2009. HIV: a discovery opening the road to novel scientific knowledge and global health improvement (Nobel lecture). *Angew Chem Int Ed Engl,* 48**,** 5809-14.

BARRE-SINOUSSI, F., CHERMANN, J. C., REY, F., NUGEYRE, M. T., 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**,** 868-71.

BELL, C. C., FAULKNER, L., MARTINSSON, K., FARRELL, J., ALFIREVIC, A., TUGWOOD, J., PIRMOHAMED, M., NAISBITT, D. J. & PARK, B. K. 2013. T-Cells from HLA-B*57:01+ Human Subjects Are Activated with Abacavir through Two Independent Pathways and Induce Cell Death by Multiple Mechanisms. *Chem Res Toxicol,* 26**,** 759-66.

BELL, N. M. & LEVER, A. M. 2013. HIV Gag polyprotein: processing and early viral particle assembly. *Trends Microbiol,* 21**,** 136-44.

BHARADWAJ, M., ILLING, P., THEODOSSIS, A., PURCELL, A. W., ROSSJOHN, J. & MCCLUSKEY, J. 2012. Drug hypersensitivity and human leukocyte antigens of the major histocompatibility complex. *Annu Rev Pharmacol Toxicol,* 52**,** 401-31.

BIENIASZ, P. D. 2012. An overview of intracellular interactions between immunodeficiency viruses and their hosts. *AIDS,* 26**,** 1243-54.

BLACK, V., HOFFMAN, R. M., SUGAR, C. A., MENON, P., VENTER, F., CURRIER, J. S. & REES, H. 2008. Safety and efficacy of initiating highly active antiretroviral therapy in an integrated antenatal and HIV clinic in Johannesburg, South Africa. *J Acquir Immune Defic Syndr,* 49**,** 276-81.

BORROW, P., LEWICKI, H., HAHN, B. H., SHAW, G. M. & OLDSTONE, M. B. 1994. Virus-specific CD8+ cytotoxic T-lymphocyte activity associated with control of viremia in primary human immunodeficiency virus type 1 infection. *J Virol,* 68**,** 6103-10.

BRACKENRIDGE, S., EVANS, E. J., TOEBES, M., GOONETILLEKE, N., LIU, M. K., DI GLERIA, K., SCHUMACHER, T. N., DAVIS, S. J., MCMICHAEL, A. J. & GILLESPIE, G. M. 2011. An early HIV mutation within an HLA-B*57-restricted T cell epitope abrogates binding to the killer inhibitory receptor 3DL1. *J Virol,* 85**,** 5415-22.

BRANDER, C., YANG, O. O., JONES, N. G., LEE, Y., GOULDER, P., JOHNSON, R. P., TROCHA, A., COLBERT, D., HAY, C., BUCHBINDER, S., BERGMANN, C. C., ZWEERINK, H. J., WOLINSKY, S., BLATTNER, W. A., KALAMS, S. A. & WALKER, B. D. 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. *J Virol,* 73**,** 10191-8.

BRENCHLEY, J. M., SCHACKER, T. W., RUFF, L. E., PRICE, D. A., TAYLOR, J. H., BEILMAN, G. J., NGUYEN, P. L., KHORUTS, A., LARSON, M., HAASE, A. T. & DOUEK, D. C. 2004. CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. *J Exp Med,* 200**,** 749-59.

BROCKMAN, M. A., BRUMME, Z. L., BRUMME, C. J., MIURA, T., SELA, J., ROSATO, P. C., KADIE, C. M., CARLSON, J. M., MARKLE, T. J., STREECK, H., KELLEHER, A. D., MARKOWITZ, M., JESSEN, H., ROSENBERG, E., ALTFELD, M., HARRIGAN, P. R., HECKERMAN, D., WALKER, B. D. & ALLEN, T. M. 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. *J Virol,* 84**,** 11937-49.

BRUMME, Z. L., BRUMME, C. J., HECKERMAN, D., KORBER, B. T., DANIELS, M., CARLSON, J., KADIE, C., BHATTACHARYA, T., CHUI, C., SZINGER, J., MO, T., HOGG, R. S., MONTANER, J. S., FRAHM, N., BRANDER, C., WALKER, B. D. & HARRIGAN, P. R. 2007. Evidence of differential HLA class I-mediated viral evolution in functional and accessory/regulatory genes of HIV-1. *PLoS Pathog,* 3**,** e94.

CAMPBELL, K. S. & HASEGAWA, J. 2013. Natural killer cell biology: An update and future directions. *J Allergy Clin Immunol,* 132**,** 536-44.

CAMPBELL, M. S., KAHLE, E. M., CELUM, C., LINGAPPA, J. R., KAPIGA, S., MUJUGIRA, A., MUGO, N. R., FIFE, K. H., MULLINS, J. I., BAETEN, J. M. & PARTNERS IN PREVENTION, H. S. V. H. I. V. T. S. T. 2013. Plasma viral loads during early HIV-1 infection are similar in subtype C- and non-subtype C-infected African seroconverters. *J Infect Dis,* 207**,** 1166-70.

CARIAS, A. M., MCCOOMBE, S., MCRAVEN, M., ANDERSON, M., GALLOWAY, N., VANDERGRIFT, N., FOUGHT, A. J., LURAIN, J., DUPLANTIS, M., VEAZEY, R. S. & HOPE, T. J. 2013. Defining the interaction of HIV-1 with the mucosal barriers of the female reproductive tract. *J Virol,* 87**,** 11388-400.

CARLSON, J. M., LISTGARTEN, J., PFEIFER, N., TAN, V., KADIE, C., WALKER, B. D., NDUNG'U, T., SHAPIRO, R., FRATER, J., BRUMME, Z. L., GOULDER, P. J. & HECKERMAN, D. 2012. Widespread impact of HLA restriction on immune control and escape pathways of HIV-1. *J Virol,* 86**,** 5230-43.

CAROSELLA, E. D., MOREAU, P., LEMAOULT, J. & ROUAS-FREISS, N. 2008. HLA-G: from biology to clinical benefits. *Trends Immunol,* 29**,** 125-32.

CARR, D. F., CHAPONDA, M., JORGENSEN, A. L., CASTRO, E. C., VAN OOSTERHOUT, J. J., KHOO, S. H., LALLOO, D. G., HEYDERMAN, R. S., ALFIREVIC, A. & PIRMOHAMED, M. 2013a. Association of human leukocyte antigen alleles and nevirapine hypersensitivity in a Malawian HIV-infected population. *Clin Infect Dis,* 56**,** 1330-9.

CARR, D. F., CHAPONDA, M., JORGENSEN, A. L., CASTRO, E. C., VAN OOSTERHOUT, J. J., KHOO, S. H., LALLOO, D. G., HEYDERMAN, R. S., ALFIREVIC, A. & PIRMOHAMED, M. 2013b. Association of Human Leukocyte Antigen Alleles and Nevirapine Hypersensitivity in a Malawian HIV-Infected Population. *Clin Infect Dis*.

CARR, W. H., ROSEN, D. B., ARASE, H., NIXON, D. F., MICHAELSSON, J. & LANIER, L. L. 2007. Cutting Edge: KIR3DS1, a gene implicated in resistance to progression to AIDS, encodes a DAP12-associated receptor expressed on NK cells that triggers NK cell activation. *J Immunol,* 178**,** 647-51.

CARRINGTON, M. & ALTER, G. 2012. Innate immune control of HIV. *Cold Spring Harb Perspect Med,* 2**,** a007070.

CASE, K. 1986. Nomenclature: human immunodeficiency virus. *Ann Intern Med,* 105**,** 133.

CELLA, M., LONGO, A., FERRARA, G. B., STROMINGER, J. L. & COLONNA, M. 1994. NK3-specific natural killer cells are selectively inhibited by Bw4-positive HLA alleles with isoleucine 80. *J Exp Med,* 180**,** 1235-42.

CELUM, C. L., BUCHBINDER, S. P., DONNELL, D., DOUGLAS, J. M., JR., MAYER, K., KOBLIN, B., MARMOR, M., BOZEMAN, S., GRANT, R. M., FLORES, J. & SHEPPARD, H. W. 2001. Early human immunodeficiency virus (HIV) infection in the HIV Network for Prevention Trials Vaccine Preparedness Cohort: risk behaviors, symptoms, and early plasma and genital tract virus load. *J Infect Dis,* 183**,** 23-35.

CENTERS FOR DISEASE, C. 1981. Pneumocystis pneumonia--Los Angeles. *MMWR Morb Mortal Wkly Rep,* 30**,** 250-2.

CHAHROUDI, A., BOSINGER, S. E., VANDERFORD, T. H., PAIARDINI, M. & SILVESTRI, G. 2012. Natural SIV hosts: showing AIDS the door. *Science,* 335**,** 1188-93.

CHAN, H. W., MILLER, J. S., MOORE, M. B. & LUTZ, C. T. 2005. Epigenetic control of highly homologous killer Ig-like receptor gene alleles. *J Immunol,* 175**,** 5966-74.

CHANG, S. Y., CHA, H. R., IGARASHI, O., RENNERT, P. D., KISSENPFENNIG, A., MALISSEN, B., NANNO, M., KIYONO, H. & KWEON, M. N. 2008. Cutting edge: Langerin+ dendritic cells in the mesenteric lymph node set the stage for skin and gut immune system cross-talk. *J Immunol,* 180**,** 4361-5.

CHANTARANGSU, S., MUSHIRODA, T., MAHASIRIMONGKOL, S., KIERTIBURANAKUL, S., SUNGKANUPARPH, S., MANOSUTHI, W., TANTISIRIWAT, W., CHAROENYINGWATTANA, A., SURA, T., CHANTRATITA, W. & NAKAMURA, Y. 2009. HLA-B*3505 allele is a strong predictor for nevirapine-induced skin adverse drug reactions in HIV-infected Thai patients. *Pharmacogenet Genomics,* 19**,** 139-46.

CHECKLEY, M. A., LUTTGE, B. G. & FREED, E. O. 2011. HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation. *J Mol Biol,* 410**,** 582-608.

CHESSMAN, D., KOSTENKO, L., LETHBORG, T., PURCELL, A. W., WILLIAMSON, N. A., CHEN, Z., KJER-NIELSEN, L., MIFSUD, N. A., TAIT, B. D., HOLDSWORTH, R., ALMEIDA, C. A., NOLAN, D., MACDONALD, W. A., ARCHBOLD, J. K., KELLERHER, A. D., MARRIOTT, D., MALLAL, S., BHARADWAJ, M., ROSSJOHN, J. & MCCLUSKEY, J. 2008. Human leukocyte antigen class I-restricted activation of CD8+ T cells provides the immunogenetic basis of a systemic drug hypersensitivity. *Immunity,* 28**,** 822-32.

CHOPERA, D. R., COTTON, L. A., ZAWAIRA, A., MANN, J. K., NGANDU, N. K., NTALE, R., CARLSON, J. M., MLISANA, K., WOODMAN, Z., DE ASSIS ROSA, D., MARTIN, E., MIURA, T., PEREYRA, F., WALKER, B. D., GRAY, C. M., MARTIN, D. P., NDUNG'U, T., BROCKMAN, M. A., KARIM, S. A., BRUMME, Z. L., WILLIAMSON, C. & TEAM, C. S. 2012. Intersubtype differences in the effect of a rare p24 gag mutation on HIV-1 replicative fitness. *J Virol,* 86**,** 13423-33.

CHOPERA, D. R., MLOTSHWA, M., WOODMAN, Z., MLISANA, K., DE ASSIS ROSA, D., MARTIN, D. P., ABDOOL KARIM, S., GRAY, C. M., WILLIAMSON, C. & TEAM, C. S. 2011. Virological and immunological factors associated with HIV-1 differential disease progression in HLA-B 58:01-positive individuals. *J Virol,* 85**,** 7070-80.

CHOPERA, D. R., WOODMAN, Z., MLISANA, K., MLOTSHWA, M., MARTIN, D. P., SEOIGHE, C., TREURNICHT, F., DE ROSA, D. A., HIDE, W., KARIM, S. A., GRAY, C. M., WILLIAMSON, C. & TEAM, C. S. 2008. Transmission of HIV-1 CTL escape variants provides HLA-mismatched recipients with a survival advantage. *PLoS Pathog,* 4**,** e1000033.

CLAVEL, F., GUETARD, D., BRUN-VEZINET, F., CHAMARET, S., REY, M. A., SANTOS-FERREIRA, M. O., LAURENT, A. G., DAUGUET, C., KATLAMA, C., ROUZIOUX, C. & ET AL. 1986. Isolation of a new human retrovirus from West African patients with AIDS. *Science,* 233**,** 343-6.

COHEN, M. S., HELLMANN, N., LEVY, J. A., DECOCK, K. & LANGE, J. 2008. The spread, treatment, and prevention of HIV-1: evolution of a global pandemic. *J Clin Invest,* 118**,** 1244-54.

COHEN, M. S., SHAW, G. M., MCMICHAEL, A. J. & HAYNES, B. F. 2011. Acute HIV-1 Infection. *N Engl J Med,* 364**,** 1943-54.

COIRAS, M., LOPEZ-HUERTAS, M. R., PEREZ-OLMEDA, M. & ALCAMI, J. 2009. Understanding HIV-1 latency provides clues for the eradication of long-term reservoirs. *Nat Rev Microbiol,* 7**,** 798-812.

COLOMBO, S., RAUCH, A., ROTGER, M., FELLAY, J., MARTINEZ, R., FUX, C., THURNHEER, C., GUNTHARD, H. F., GOLDSTEIN, D. B., FURRER, H., TELENTI, A. & SWISS, H. I. V. C. S. 2008. The HCP5 single-nucleotide polymorphism: a simple screening tool for prediction of hypersensitivity reaction to abacavir. *J Infect Dis,* 198**,** 864-7.

COSTA, L. J., ZHENG, Y. H., SABOTIC, J., MAK, J., FACKLER, O. T. & PETERLIN, B. M. 2004. Nef binds p6* in GagPol during replication of human immunodeficiency virus type 1. *J Virol,* 78**,** 5311-23.

CRAWFORD, H., PRADO, J. G., LESLIE, A., HUE, S., HONEYBORNE, I., REDDY, S., VAN DER STOK, M., MNCUBE, Z., BRANDER, C., ROUSSEAU, C., MULLINS, J. I., KASLOW, R., GOEPFERT, P., ALLEN, S., HUNTER, E., MULENGA, J., KIEPIELA, P., WALKER, B. D. & GOULDER, P. J. 2007. Compensatory mutation partially restores fitness and delays reversion of escape mutation within the immunodominant HLA-B*5703-restricted Gag epitope in chronic human immunodeficiency virus type 1 infection. *J Virol,* 81**,** 8346-51.

DE WIT, E., DELPORT, W., RUGAMIKA, C. E., MEINTJES, A., MOLLER, M., VAN HELDEN, P. D., SEOIGHE, C. & HOAL, E. G. 2010. Genome-wide analysis of the structure of the South African Coloured Population in the Western Cape. *Hum Genet,* 128**,** 145-53.

DEEKS, S. G. & WALKER, B. D. 2007. Human immunodeficiency virus controllers: mechanisms of durable virus control in the absence of antiretroviral therapy. *Immunity,* 27**,** 406-16.

DEGUINE, J. & BOUSSO, P. 2013. Dynamics of NK cell interactions in vivo. *Immunol Rev,* 251**,** 154-9.

DELLO RUSSO, C., LISI, L., FABBIANI, M., GAGLIARDI, D., FANTI, I., DI GIAMBENEDETTO, S., CAUDA, R. & NAVARRA, P. 2014. Detection of HLA-B*57:01 by real-time PCR: implementation into routine clinical practice and additional validation data. *Pharmacogenomics,* 15**,** 319-27.

DONG, B. J., ZHENG, Y., HUGHES, M. D., FRYMOYER, A., VEROTTA, D., LIZAK, P., SAWE, F., CURRIER, J. S., LOCKMAN, S., AWEEKA, F. T. & TEAM, A. C. T. G. S. 2012. Nevirapine pharmacokinetics and risk of rash and hepatitis among HIV-infected sub-Saharan African women. *AIDS,* 26**,** 833-41.

DOUEK, D. C., BRENCHLEY, J. M., BETTS, M. R., AMBROZAK, D. R., HILL, B. J., OKAMOTO, Y., CASAZZA, J. P., KURUPPU, J., KUNSTMAN, K., WOLINSKY, S., GROSSMAN, Z., DYBUL, M., OXENIUS, A., PRICE, D. A., CONNORS, M. & KOUP, R. A. 2002. HIV preferentially infects HIV-specific CD4+ T cells. *Nature,* 417**,** 95-8.

ENGELMAN, A. & CHEREPANOV, P. 2012. The structural biology of HIV-1: mechanistic and therapeutic insights. *Nat Rev Microbiol,* 10**,** 279-90.

FAHRBACH, K. M., BARRY, S. M., AYEHUNIE, S., LAMORE, S., KLAUSNER, M. & HOPE, T. J. 2007. Activated CD34-derived Langerhans cells mediate transinfection with human immunodeficiency virus. *J Virol,* 81**,** 6858-68.

FARIA, N. R., RAMBAUT, A., SUCHARD, M. A., BAELE, G., BEDFORD, T., WARD, M. J., TATEM, A. J., SOUSA, J. D., ARINAMINPATHY, N., PEPIN, J., POSADA, D., PEETERS, M., PYBUS, O. G. & LEMEY, P. 2014. HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. *Science,* 346**,** 56-61.

FAURE, M. & LONG, E. O. 2002. KIR2DL4 (CD158d), an NK cell-activating receptor with inhibitory potential. *J Immunol,* 168**,** 6208-14.

FIEBIG, E. W., WRIGHT, D. J., RAWAL, B. D., GARRETT, P. E., SCHUMACHER, R. T., PEDDADA, L., HELDEBRANT, C., SMITH, R., CONRAD, A., KLEINMAN, S. H. & BUSCH, M. P. 2003. Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. *AIDS,* 17**,** 1871-9.

FLYS, T., MARLOWE, N., HACKETT, J., PARKIN, N., SCHUMAKER, M., HOLZMAYER, V., HAY, P. & ESHLEMAN, S. H. 2005. Analysis of PTAP duplications in the gag p6 region of subtype C HIV type 1. *AIDS Res Hum Retroviruses,* 21**,** 739-41.

FORTUNY, C., DUICULESCU, D., CHENG, K., GARGES, H. P., COTTON, M., TAMARIRT, D. P., FORD, S. L., WIRE, M. B., GIVENS, N., ROSS, L. L., LOU, Y., PERGER, T. & SIEVERS, J. 2014. Pharmacokinetics and 48-week safety and antiviral activity of fosamprenavir-containing regimens in HIV-infected 2- to 18-year-old children. *Pediatr Infect Dis J,* 33**,** 50-6.

FRAHM, N., ADAMS, S., KIEPIELA, P., LINDE, C. H., HEWITT, H. S., LICHTERFELD, M., SANGO, K., BROWN, N. V., PAE, E., WURCEL, A. G., ALTFELD, M., FEENEY, M. E., ALLEN, T. M., ROACH, T., ST JOHN, M. A., DAAR, E. S., ROSENBERG, E., KORBER, B., MARINCOLA, F., WALKER, B. D., GOULDER, P. J. & BRANDER, C. 2005. HLA-B63 presents HLA-B57/B58-restricted cytotoxic T-lymphocyte epitopes and is associated with low human immunodeficiency virus load. *J Virol,* 79**,** 10218-25.

FRAHM, N., KIEPIELA, P., ADAMS, S., LINDE, C. H., HEWITT, H. S., SANGO, K., FEENEY, M. E., ADDO, M. M., LICHTERFELD, M., LAHAIE, M. P., PAE, E., WURCEL, A. G., ROACH, T., ST JOHN, M. A., ALTFELD, M., MARINCOLA, F. M., MOORE, C., MALLAL, S., CARRINGTON, M., HECKERMAN, D., ALLEN, T. M., MULLINS, J. I., KORBER, B. T., GOULDER, P. J., WALKER, B. D. & BRANDER, C. 2006. Control of human immunodeficiency virus replication by cytotoxic T lymphocytes targeting subdominant epitopes. *Nat Immunol,* 7**,** 173-8.

GALLO, R. C., SALAHUDDIN, S. Z., POPOVIC, M., SHEARER, G. M., KAPLAN, M., HAYNES, B. F., PALKER, T. J., REDFIELD, R., OLESKE, J., SAFAI, B. & ET AL. 1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science,* 224**,** 500-3.

GANSER-PORNILLOS, B. K., YEAGER, M. & PORNILLOS, O. 2012. Assembly and architecture of HIV. *Adv Exp Med Biol,* 726**,** 441-65.

GAO, S., GUI, X. E., LIANG, K., LIU, Z., HU, J. & DONG, B. 2012. HLA-dependent hypersensitivity reaction to nevirapine in Chinese Han HIV-infected patients. *AIDS Res Hum Retroviruses,* 28**,** 540-3.

GAO, X., O'BRIEN, T. R., WELZEL, T. M., MARTI, D., QI, Y., GOEDERT, J. J., PHAIR, J., PFEIFFER, R. & CARRINGTON, M. 2010. HLA-B alleles associate consistently with HIV heterosexual transmission, viral load, and progression to AIDS, but not susceptibility to infection. *AIDS,* 24**,** 1835-40.

GATANAGA, H., DAS, D., SUZUKI, Y., YEH, D. D., HUSSAIN, K. A., GHOSH, A. K. & MITSUYA, H. 2006. Altered HIV-1 Gag protein interactions with cyclophilin A (CypA) on the acquisition of H219Q and H219P substitutions in the CypA binding loop. *J Biol Chem,* 281**,** 1241-50.

GATANAGA, H., YAZAKI, H., TANUMA, J., HONDA, M., GENKA, I., TERUYA, K., TACHIKAWA, N., KIKUCHI, Y. & OKA, S. 2007. HLA-Cw8 primarily associated with hypersensitivity to nevirapine. *AIDS,* 21**,** 264-5.

GAUDIERI, S., DESANTIS, D., MCKINNON, E., MOORE, C., NOLAN, D., WITT, C. S., MALLAL, S. A. & CHRISTIANSEN, F. T. 2005. Killer immunoglobulin-like receptors and HLA act both independently and synergistically to modify HIV disease progression. *Genes Immun,* 6**,** 683-90.

GEDIL, M. A., STEINER, N. K. & HURLEY, C. K. 2005. Genomic characterization of KIR2DL4 in families and unrelated individuals reveals extensive diversity in exon and intron sequences including a common frameshift variation occurring in several alleles. *Tissue Antigens,* 65**,** 402-18.

GEIJTENBEEK, T. B. & VAN KOOYK, Y. 2003. DC-SIGN: a novel HIV receptor on DCs that mediates HIV-1 transmission. *Curr Top Microbiol Immunol,* 276**,** 31-54.

GELDMACHER, C., CURRIER, J. R., HERRMANN, E., HAULE, A., KUTA, E., MCCUTCHAN, F., NJOVU, L., GEIS, S., HOFFMANN, O., MABOKO, L., WILLIAMSON, C., BIRX, D., MEYERHANS, A., COX, J. & HOELSCHER, M. 2007. CD8 T-cell recognition of multiple epitopes within specific Gag regions is associated with maintenance of a low steady-state viremia in human immunodeficiency virus type 1-seropositive patients. *J Virol,* 81**,** 2440-8.

GENTLE, N. L., PAXIMADIS, M., PUREN, A. & TIEMESSEN, C. T. 2013. Genetic variability in markers of HLA-C expression in two diverse South African populations. *PLoS One,* 8**,** e67780.

GESKUS, R. B., PRINS, M., HUBERT, J. B., MIEDEMA, F., BERKHOUT, B., ROUZIOUX, C., DELFRAISSY, J. F. & MEYER, L. 2007. The HIV RNA setpoint theory revisited. *Retrovirology,* 4**,** 65.

GIJSBERS, E. F., FEENSTRA, K. A., VAN NUENEN, A. C., NAVIS, M., HERINGA, J., SCHUITEMAKER, H. & KOOTSTRA, N. A. 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.

GILLESPIE, G. M., BASHIROVA, A., DONG, T., MCVICAR, D. W., ROWLAND-JONES, S. L. & CARRINGTON, M. 2007. Lack of KIR3DS1 binding to MHC class I Bw4 tetramers in complex with CD8+ T cell epitopes. *AIDS Res Hum Retroviruses,* 23**,** 451-5.

GOMEZ-LOZANO, N., DE PABLO, R., PUENTE, S. & VILCHES, C. 2003. Recognition of HLA-G by the NK cell receptor KIR2DL4 is not essential for human reproduction. *Eur J Immunol,* 33**,** 639-44.

GOMEZ-LOZANO, N., ESTEFANIA, E., WILLIAMS, F., HALFPENNY, I., MIDDLETON, D., SOLIS, R. & VILCHES, C. 2005. The silent KIR3DP1 gene (CD158c) is transcribed and might encode a secreted receptor in a minority of humans, in whom the KIR3DP1, KIR2DL4 and KIR3DL1/KIR3DS1 genes are duplicated. *Eur J Immunol,* 35**,** 16-24.

GONZALEZ-GALARZA, F. F., CHRISTMAS, S., MIDDLETON, D. & JONES, A. R. 2011. Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. *Nucleic Acids Res,* 39**,** D913-9.

GOODRIDGE, J. P., BURIAN, A., LEE, N. & GERAGHTY, D. E. 2013. HLA-F and MHC Class I Open Conformers Are Ligands for NK Cell Ig-like Receptors. *J Immunol*.

GOODRIDGE, J. P., LATHBURY, L. J., STEINER, N. K., SHULSE, C. N., PULLIKOTIL, P., SEIDAH, N. G., HURLEY, C. K., CHRISTIANSEN, F. T. & WITT, C. S. 2007. Three common alleles of KIR2DL4 (CD158d) encode constitutively expressed, inducible and secreted receptors in NK cells. *Eur J Immunol,* 37**,** 199-211.

GRAEF, T., MOESTA, A. K., NORMAN, P. J., ABI-RACHED, L., VAGO, L., OLDER AGUILAR, A. M., GLEIMER, M., HAMMOND, J. A., GUETHLEIN, L. A., BUSHNELL, D. A., ROBINSON, P. J. & PARHAM, P. 2009. KIR2DS4 is a product

of gene conversion with KIR3DL2 that introduced specificity for HLA-A*11 while diminishing avidity for HLA-C. *J Exp Med,* 206**,** 2557-72.

GRANIER, C., BATTIVELLI, E., LECUROUX, C., VENET, A., LAMBOTTE, O., SCHMITT-BOULANGER, M., DELAUGERRE, C., MOLINA, J. M., CHAKRABARTI, L. A., CLAVEL, F. & HANCE, A. J. 2013. Pressure from TRIM5alpha contributes to control of HIV-1 replication by individuals expressing protective HLA-B alleles. *J Virol,* 87**,** 10368-80.

GRAS, G. & KAUL, M. 2010. Molecular mechanisms of neuroinvasion by monocytes-macrophages in HIV-1 infection. *Retrovirology,* 7**,** 30.

GRILO, N. M., ANTUNES, A. M., CAIXAS, U., MARINHO, A. T., CHARNEIRA, C., CONCEICAO OLIVEIRA, M., MONTEIRO, E. C., MATILDE MARQUES, M. & PEREIRA, S. A. 2013. Monitoring abacavir bioactivation in humans: Screening for an aldehyde metabolite. *Toxicol Lett*.

GUERINI, F. R., LO CAPUTO, S., GORI, A., BANDERA, A., MAZZOTTA, F., UGLIETTI, A., ZANZOTTERA, M., MASERATI, R. & CLERICI, M. 2011. Under representation of the inhibitory KIR3DL1 molecule and the KIR3DL1+/BW4+ complex in HIV exposed seronegative individuals. *J Infect Dis,* 203**,** 1235-9.

GUMPERZ, J. E., LITWIN, V., PHILLIPS, J. H., LANIER, L. L. & PARHAM, P. 1995. The Bw4 public epitope of HLA-B molecules confers reactivity with natural killer cell clones that express NKB1, a putative HLA receptor. *J Exp Med,* 181**,** 1133-44.

HABEGGER DE SORRENTINO, A., SINCHI, J. L., MARINIC, K., LOPEZ, R. & ILIOVICH, E. 2013. KIR-HLA-A and B alleles of the Bw4 epitope against HIV infection in discordant heterosexual couples in Chaco Argentina. *Immunology,* 140**,** 273-9.

HEMELAAR, J. 2012. The origin and diversity of the HIV-1 pandemic. *Trends Mol Med,* 18**,** 182-92.

HETHERINGTON, S., HUGHES, A. R., MOSTELLER, M., SHORTINO, D., BAKER, K. L., SPREEN, W., LAI, E., DAVIES, K., HANDLEY, A., DOW, D. J., FLING, M. E., STOCUM, M., BOWMAN, C., THURMOND, L. M. & ROSES, A. D. 2002. Genetic variations in HLA-B region and hypersensitivity reactions to abacavir. *Lancet,* 359**,** 1121-2.

HLADIK, F. & HOPE, T. J. 2009. HIV infection of the genital mucosa in women. *Curr HIV/AIDS Rep,* 6**,** 20-8.

HLADIK, F. & MCELRATH, M. J. 2008. Setting the stage: host invasion by HIV. *Nat Rev Immunol,* 8**,** 447-57.

HONEYBORNE, I., CODONER, F. M., LESLIE, A., TUDOR-WILLIAMS, G., LUZZI, G., NDUNG'U, T., WALKER, B. D., GOULDER, P. J. & PRADO, J. G. 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. *J Virol,* 84**,** 11279-88.

HONEYBORNE, I., PRENDERGAST, A., PEREYRA, F., LESLIE, A., CRAWFORD, H., PAYNE, R., REDDY, S., BISHOP, K., MOODLEY, E., NAIR, K., VAN DER STOK, M., MCCARTHY, N., ROUSSEAU, C. M., ADDO, M., MULLINS, J. I., BRANDER, C., KIEPIELA, P., WALKER, B. D. & GOULDER, P. J. 2007. Control of human immunodeficiency virus type 1 is associated with HLA-B*13 and targeting of multiple gag-specific CD8+ T-cell epitopes. *J Virol,* 81**,** 3667-72.

HONG, H. A., LOUBSER, A. S., DE ASSIS ROSA, D., NARANBHAI, V., CARR, W., PAXIMADIS, M., LEWIS, D. A., TIEMESSEN, C. T. & GRAY, C. M. 2011. Killer-cell immunoglobulin-like receptor genotyping and HLA killer-cell immunoglobulin-

like receptor-ligand identification by real-time polymerase chain reaction. *Tissue Antigens,* 78**,** 185-94.

HONG., H. A., PAXIMADIS., M., GRAY., G. E., KUHN., L. & TIEMESSEN., C. T. 2013. KIR2DS4 allelic variants: differential effects on in utero and intrapartum HIV-1 mother-to-child transmission. *Clinical Immunology,* doi: 10.1016/j.clim.2013.09.005.

HOU, L., CHEN, M., NG, J. & HURLEY, C. K. 2012. Conserved KIR allele-level haplotypes are altered by microvariation in individuals with European ancestry. *Genes Immun,* 13**,** 47-58.

HSU, K. C., CHIDA, S., GERAGHTY, D. E. & DUPONT, B. 2002. The killer cell immunoglobulin-like receptor (KIR) genomic region: gene-order, haplotypes and allelic polymorphism. *Immunol Rev,* 190**,** 40-52.

ILLING, P. T., VIVIAN, J. P., DUDEK, N. L., KOSTENKO, L., CHEN, Z., BHARADWAJ, M., MILES, J. J., KJER-NIELSEN, L., GRAS, S., WILLIAMSON, N. A., BURROWS, S. R., PURCELL, A. W., ROSSJOHN, J. & MCCLUSKEY, J. 2012. Immune self-reactivity triggered by drug-modified HLA-peptide repertoire. *Nature,* 486**,** 554-8.

INTERNATIONAL, H. I. V. C. S., PEREYRA, F., JIA, X., MCLAREN, P. J., TELENTI, A., DE BAKKER, P. I., WALKER, B. D., RIPKE, S., BRUMME, C. J., PULIT, S. L., CARRINGTON, M., KADIE, C. M., CARLSON, J. M., HECKERMAN, D., GRAHAM, R. R., PLENGE, R. M., DEEKS, S. G., GIANNINY, L., CRAWFORD, G., SULLIVAN, J., GONZALEZ, E., DAVIES, L., CAMARGO, A., MOORE, J. M., BEATTIE, N., GUPTA, S., CRENSHAW, A., BURTT, N. P., GUIDUCCI, C., GUPTA, N., GAO, X., QI, Y., YUKI, Y., PIECHOCKA-TROCHA, A., CUTRELL, E., ROSENBERG, R., MOSS, K. L., LEMAY, P., O'LEARY, J., SCHAEFER, T., VERMA, P., TOTH, I., BLOCK, B., BAKER, B., ROTHCHILD, A., LIAN, J., PROUDFOOT, J., ALVINO, D. M., VINE, S., ADDO, M. M., ALLEN, T. M., ALTFELD, M., HENN, M. R., LE GALL, S., STREECK, H., HAAS, D. W., KURITZKES, D. R., ROBBINS, G. K., SHAFER, R. W., GULICK, R. M., SHIKUMA, C. M., HAUBRICH, R., RIDDLER, S., SAX, P. E., DAAR, E. S., RIBAUDO, H. J., AGAN, B., AGARWAL, S., AHERN, R. L., ALLEN, B. L., ALTIDOR, S., ALTSCHULER, E. L., AMBARDAR, S., ANASTOS, K., ANDERSON, B., ANDERSON, V., ANDRADY, U., ANTONISKIS, D., BANGSBERG, D., BARBARO, D., BARRIE, W., BARTCZAK, J., BARTON, S., BASDEN, P., BASGOZ, N., BAZNER, S., BELLOS, N. C., BENSON, A. M., BERGER, J., BERNARD, N. F., BERNARD, A. M., BIRCH, C., BODNER, S. J., BOLAN, R. K., BOUDREAUX, E. T., BRADLEY, M., BRAUN, J. F., BRNDJAR, J. E., BROWN, S. J., BROWN, K., et al. 2010. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. *Science,* 330**,** 1551-7.

IVERSEN, A. K., STEWART-JONES, G., LEARN, G. H., CHRISTIE, N., SYLVESTER-HVIID, C., ARMITAGE, A. E., KAUL, R., BEATTIE, T., LEE, J. K., LI, Y., CHOTIYARNWONG, P., DONG, T., XU, X., LUSCHER, M. A., MACDONALD, K., ULLUM, H., KLARLUND-PEDERSEN, B., SKINHOJ, P., FUGGER, L., BUUS, S., MULLINS, J. I., JONES, E. Y., VAN DER MERWE, P. A. & MCMICHAEL, A. J. 2006. Conflicting selective forces affect T cell receptor contacts in an immunodominant human immunodeficiency virus epitope. *Nat Immunol,* 7**,** 179-89.

JENNES, W., VERHEYDEN, S., DEMANET, C., ADJE-TOURE, C. A., VUYLSTEKE, B., NKENGASONG, J. N. & KESTENS, L. 2006. Cutting edge: resistance to HIV-1 infection among African female sex workers is associated with inhibitory KIR in the absence of their HLA ligands. *J Immunol,* 177**,** 6588-92.

JENNES, W., VERHEYDEN, S., MERTENS, J. W., CAMARA, M., SEYDI, M., DIEYE, T. N., MBOUP, S., DEMANET, C. & KESTENS, L. 2013. Inhibitory KIR/HLA incompatibility between sexual partners confers protection against HIV-1 transmission. *Blood,* 121**,** 1157-64.

JIANG, W., JOHNSON, C., JAYARAMAN, J., SIMECEK, N., NOBLE, J., MOFFATT, M. F., COOKSON, W. O., TROWSDALE, J. & TRAHERNE, J. A. 2012. Copy number variation leads to considerable diversity for B but not A haplotypes of the human KIR genes encoding NK cell receptors. *Genome Res,* 22**,** 1845-54.

JIANG, Y., CHEN, O., CUI, C., ZHAO, B., HAN, X., ZHANG, Z., LIU, J., XU, J., HU, Q., LIAO, C. & SHANG, H. 2013. KIR3DS1/L1 and HLA-Bw4-80I are associated with HIV disease progression among HIV typical progressors and long-term nonprogressors. *BMC Infect Dis,* 13**,** 405.

JOST, S. & ALTFELD, M. 2013. Control of human viral infections by natural killer cells. *Annu Rev Immunol,* 31**,** 163-94.

KARN, J. & STOLTZFUS, C. M. 2012. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harb Perspect Med,* 2**,** a006916.

KAWAMURA, T., GULDEN, F. O., SUGAYA, M., MCNAMARA, D. T., BORRIS, D. L., LEDERMAN, M. M., ORENSTEIN, J. M., ZIMMERMAN, P. A. & BLAUVELT, A. 2003. R5 HIV productively infects Langerhans cells, and infection levels are regulated by compound CCR5 polymorphisms. *Proc Natl Acad Sci U S A,* 100**,** 8401-6.

KEELE, B. F., GIORGI, E. E., SALAZAR-GONZALEZ, J. F., DECKER, J. M., PHAM, K. T., SALAZAR, M. G., SUN, C., GRAYSON, T., WANG, S., LI, H., WEI, X., JIANG, C., KIRCHHERR, J. L., GAO, F., ANDERSON, J. A., PING, L. H., SWANSTROM, R., TOMARAS, G. D., BLATTNER, W. A., GOEPFERT, P. A., KILBY, J. M., SAAG, M. S., DELWART, E. L., BUSCH, M. P., COHEN, M. S., MONTEFIORI, D. C., HAYNES, B. F., GASCHEN, B., ATHREYA, G. S., LEE, H. Y., WOOD, N., SEOIGHE, C., PERELSON, A. S., BHATTACHARYA, T., KORBER, B. T., HAHN, B. H. & SHAW, G. M. 2008. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc Natl Acad Sci U S A,* 105**,** 7552-7.

KENYON, C., BUYZE, J. & COLEBUNDERS, R. 2013. HIV prevalence by race co-varies closely with concurrency and number of sex partners in South Africa. *PLoS One,* 8**,** e64080.

KHAKOO, S. I., THIO, C. L., MARTIN, M. P., BROOKS, C. R., GAO, X., ASTEMBORSKI, J., CHENG, J., GOEDERT, J. J., VLAHOV, D., HILGARTNER, M., COX, S., LITTLE, A. M., ALEXANDER, G. J., CRAMP, M. E., O'BRIEN, S. J., ROSENBERG, W. M., THOMAS, D. L. & CARRINGTON, M. 2004. HLA and NK cell inhibitory receptor genes in resolving hepatitis C virus infection. *Science,* 305**,** 872-4.

KIEPIELA, P., LESLIE, A. J., HONEYBORNE, I., RAMDUTH, D., THOBAKGALE, C., CHETTY, S., RATHNAVALU, P., MOORE, C., PFAFFEROTT, K. J., HILTON, L., ZIMBWA, P., MOORE, S., ALLEN, T., BRANDER, C., ADDO, M. M., ALTFELD, M., JAMES, I., MALLAL, S., BUNCE, M., BARBER, L. D., SZINGER, J., DAY, C., KLENERMAN, P., MULLINS, J., KORBER, B., COOVADIA, H. M., WALKER, B. D. & GOULDER, P. J. 2004. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. *Nature,* 432**,** 769-75.

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. H., NICKLE, D., ROUSSEAU, C., COOVADIA, H., MULLINS, J. I., HECKERMAN, D., WALKER, B. D. & GOULDER, P. 2007. CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. *Nat Med*, 13**,** 46-53.

KIERTIBURANAKUL, S., MAHASIRIMONGKOL, S., RAJATANAVIN, N., CHAROENYINGWATTANA, A., ROJANAWIWAT, A., WANGSOMBOONSIRI, W., MANOSUTHI, W., KANTIPONG, P., APISARNTHANARAK, A., SANGSIRINAKAKUL, W., WONGPRASIT, P., CHAIWARITH, R., TANTISIRIWAT, W., NAKAMURA, Y., MUSHIRODA, T., CHANTRATITA, W. & SUNGKANUPARPH, S. HLA-B*35:05 and CCHCR1 screening reduces nevirapine-associated cutaneous adverse reactions in Thailand: a prospective multicenter randomized controlled trial. 7th IAS conference on HIV pathogenesis, treatment and prevention., 2013 Kuala Lumpur, Malaysia.

KIKUCHI-MAKI, A., YUSA, S., CATINA, T. L. & CAMPBELL, K. S. 2003. KIR2DL4 is an IL-2-regulated NK cell receptor that exhibits limited expression in humans but triggers strong IFN-gamma production. *J Immunol*, 171**,** 3415-25.

KLEIN, J. & SATO, A. 2000. The HLA system. First of two parts. *N Engl J Med*, 343**,** 702-9.

KLOVERPRIS, H. N., HARNDAHL, M., LESLIE, A. J., CARLSON, J. M., ISMAIL, N., VAN DER STOK, M., HUANG, K. H., CHEN, F., RIDDELL, L., STEYN, D., GOEDHALS, D., VAN VUUREN, C., FRATER, J., WALKER, B. D., CARRINGTON, M., NDUNG'U, T., BUUS, S. & GOULDER, P. 2012a. HIV control through a single nucleotide on the HLA-B locus. *J Virol*, 86**,** 11493-500.

KLOVERPRIS, H. N., PAYNE, R. P., SACHA, J. B., RASAIYAAH, J. T., CHEN, F., TAKIGUCHI, M., YANG, O. O., TOWERS, G. J., GOULDER, P. & PRADO, J. G. 2013a. Early antigen presentation of protective HIV-1 KF11Gag and KK10Gag epitopes from incoming viral particles facilitates rapid recognition of infected cells by specific CD8+ T cells. *J Virol*, 87**,** 2628-38.

KLOVERPRIS, H. N., STRYHN, A., HARNDAHL, M., PAYNE, R., TOWERS, G. J., CHEN, F., RIDDELL, L., WALKER, B. D., NDUNG'U, T., LESLIE, A., BUUS, S. & GOULDER, P. 2013b. HLA-specific intracellular epitope processing shapes an immunodominance pattern for HLA-B*57 that is distinct from HLA-B*58:01. *J Virol*, 87**,** 10889-94.

KLOVERPRIS, H. N., STRYHN, A., HARNDAHL, M., VAN DER STOK, M., PAYNE, R. P., MATTHEWS, P. C., CHEN, F., RIDDELL, L., WALKER, B. D., NDUNG'U, T., BUUS, S. & GOULDER, P. 2012b. HLA-B*57 Micropolymorphism shapes HLA allele-specific epitope immunogenicity, selection pressure, and HIV immune control. *J Virol*, 86**,** 919-29.

KOEHLER, R. N., WALSH, A. M., SAATHOFF, E., TOVANABUTRA, S., ARROYO, M. A., CURRIER, J. R., MABOKO, L., HOELSCHER, M., ROBB, M. L., MICHAEL, N. L., MCCUTCHAN, F. E., KIM, J. H. & KIJAK, G. H. 2010. Class I HLA-A*7401 is associated with protection from HIV-1 acquisition and disease progression in Mbeya, Tanzania. *J Infect Dis,* 202**,** 1562-6.

KORBER, B., MULDOON, M., THEILER, J., GAO, F., GUPTA, R., LAPEDES, A., HAHN, B. H., WOLINSKY, S. & BHATTACHARYA, T. 2000. Timing the ancestor of the HIV-1 pandemic strains. *Science,* 288**,** 1789-96.

KOSMRLJ, A., READ, E. L., QI, Y., ALLEN, T. M., ALTFELD, M., DEEKS, S. G., PEREYRA, F., CARRINGTON, M., WALKER, B. D. & CHAKRABORTY, A. K. 2010. Effects of thymic selection of the T-cell repertoire on HLA class I-associated control of HIV infection. *Nature,* 465**,** 350-4.

KULKARNI, S., MARTIN, M. P. & CARRINGTON, M. 2008. The Yin and Yang of HLA and KIR in human disease. *Semin Immunol,* 20**,** 343-52.

KULKARNI, S., SAVAN, R., QI, Y., GAO, X., YUKI, Y., BASS, S. E., MARTIN, M. P., HUNT, P., DEEKS, S. G., TELENTI, A., PEREYRA, F., GOLDSTEIN, D., WOLINSKY, S., WALKER, B., YOUNG, H. A. & CARRINGTON, M. 2011. Differential microRNA regulation of HLA-C expression and its association with HIV control. *Nature,* 472**,** 495-8.

LANIER, L. L., CORLISS, B. C., WU, J., LEONG, C. & PHILLIPS, J. H. 1998. Immunoreceptor DAP12 bearing a tyrosine-based activation motif is involved in activating NK cells. *Nature,* 391**,** 703-7.

LAZARYAN, A., SONG, W., LOBASHEVSKY, E., TANG, J., SHRESTHA, S., ZHANG, K., MCNICHOLL, J. M., GARDNER, L. I., WILSON, C. M., KLEIN, R. S., ROMPALO, A., MAYER, K., SOBEL, J., KASLOW, R. A., GROUP, H. I. V. E. R. S., REACHING FOR EXCELLENCE IN ADOLESCENT, C. & HEALTH STUDY, G. 2011. The influence of human leukocyte antigen class I alleles and their population frequencies on human immunodeficiency virus type 1 control among African Americans. *Hum Immunol,* 72**,** 312-8.

LEE, S. K., POTEMPA, M. & SWANSTROM, R. 2012. The choreography of HIV-1 proteolytic processing and virion assembly. *J Biol Chem,* 287**,** 40867-74.

LEIHERER, A., LUDWIG, C. & WAGNER, R. 2009. Influence of extended mutations of the HIV-1 transframe protein p6 on Nef-dependent viral replication and infectivity in vitro. *Virology,* 387**,** 200-10.

LESLIE, A., MATTHEWS, P. C., LISTGARTEN, J., CARLSON, J. M., KADIE, C., NDUNG'U, T., BRANDER, C., COOVADIA, H., WALKER, B. D., HECKERMAN, D. & GOULDER, P. J. 2010. Additive contribution of HLA class I alleles in the immune control of HIV-1 infection. *J Virol,* 84**,** 9879-88.

LESLIE, A. J., PFAFFEROTT, K. J., CHETTY, P., DRAENERT, R., ADDO, M. M., FEENEY, M., TANG, Y., HOLMES, E. C., ALLEN, T., PRADO, J. G., ALTFELD, M., BRANDER, C., DIXON, C., RAMDUTH, D., JEENA, P., THOMAS, S. A., ST JOHN, A., ROACH, T. A., KUPFER, B., LUZZI, G., EDWARDS, A., TAYLOR, G., LYALL, H., TUDOR-WILLIAMS, G., NOVELLI, V., MARTINEZ-PICADO, J., KIEPIELA, P., WALKER, B. D. & GOULDER, P. J. 2004. HIV evolution: CTL escape mutation and reversion after transmission. *Nat Med,* 10**,** 282-9.

LEVY, J. A., HOFFMAN, A. D., KRAMER, S. M., LANDIS, J. A., SHIMABUKURO, J. M. & OSHIRO, L. S. 1984. Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science,* 225**,** 840-2.

LIANG, B., LUO, M., SCOTT-HERRIDGE, J., SEMENIUK, C., MENDOZA, M., CAPINA, R., SHEARDOWN, B., JI, H., KIMANI, J., BALL, B. T., VAN DOMSELAAR, G., GRAHAM, M., TYLER, S., JONES, S. J. & PLUMMER, F. A. 2011. A comparison of parallel pyrosequencing and sanger clone-based sequencing and its impact on the characterization of the genetic diversity of HIV-1. *PLoS One,* 6**,** e26745.

LIKANONSAKUL, S., RATTANATHAM, T., FEANGVAD, S., UTTAYAMAKUL, S., PRASITHSIRIKUL, W., TUNTHANATHIP, P., NAKAYAMA, E. E. & SHIODA, T. 2009. HLA-Cw*04 allele associated with nevirapine-induced rash in HIV-infected Thai patients. *AIDS Res Ther,* 6**,** 22.

LITTERA, R., CARCASSI, C., MASALA, A., PIANO, P., SERRA, P., ORTU, F., CORSO, N., CASULA, B., LA NASA, G., CONTU, L. & MANCONI, P. E. 2006. HLA-dependent hypersensitivity to nevirapine in Sardinian HIV patients. *AIDS,* 20**,** 1621-6.

LIU, R., PAXTON, W. A., CHOE, S., CERADINI, D., MARTIN, S. R., HORUK, R., MACDONALD, M. E., STUHLMANN, H., KOUP, R. A. & LANDAU, N. R. 1996. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell,* 86**,** 367-77.

LIU, S., KONG, C., WU, J., YING, H. & ZHU, H. 2012. Effect of CCR5-Delta32 heterozygosity on HIV-1 susceptibility: a meta-analysis. *PLoS One,* 7**,** e35020.

LOCATELLI, F., PENDE, D., MINGARI, M. C., BERTAINA, A., FALCO, M., MORETTA, A. & MORETTA, L. 2013. Cellular and molecular basis of haploidentical hematopoietic stem cell transplantation in the successful treatment of high-risk leukemias: role of alloreactive NK cells. *Front Immunol,* 4**,** 15.

LONG, E. O., KIM, H. S., LIU, D., PETERSON, M. E. & RAJAGOPALAN, S. 2013. Controlling natural killer cell responses: integration of signals for activation and inhibition. *Annu Rev Immunol,* 31**,** 227-58.

MAHNKE, L. & CLIFFORD, D. 2006. Cytotoxic T cell recognition of an HIV-1 reverse transcriptase variant peptide incorporating the K103N drug resistance mutation. *AIDS Res Ther,* 3**,** 21.

MALLAL, S., NOLAN, D., WITT, C., MASEL, G., MARTIN, A. M., MOORE, C., SAYER, D., CASTLEY, A., MAMOTTE, C., MAXWELL, D., JAMES, I. & CHRISTIANSEN, F. T. 2002. Association between presence of HLA-B*5701, HLA-DR7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir. *Lancet,* 359**,** 727-32.

MALLAL, S., PHILLIPS, E., CAROSI, G., MOLINA, J. M., WORKMAN, C., TOMAZIC, J., JAGEL-GUEDES, E., RUGINA, S., KOZYREV, O., CID, J. F., HAY, P., NOLAN, D., HUGHES, S., HUGHES, A., RYAN, S., FITCH, N., THORBORN, D., BENBOW, A. & TEAM, P.-S. 2008. HLA-B*5701 screening for hypersensitivity to abacavir. *N Engl J Med,* 358**,** 568-79.

MANDELBOIM, O., REYBURN, H. T., VALES-GOMEZ, M., PAZMANY, L., COLONNA, M., BORSELLINO, G. & STROMINGER, J. L. 1996. Protection from lysis by natural killer cells of group 1 and 2 specificity is mediated by residue 80 in human histocompatibility leukocyte antigen C alleles and also occurs with empty major histocompatibility complex molecules. *J Exp Med,* 184**,** 913-22.

MAO, Y., WANG, L., GU, C., HERSCHHORN, A., XIANG, S. H., HAIM, H., YANG, X. & SODROSKI, J. 2012. Subunit organization of the membrane-bound HIV-1 envelope glycoprotein trimer. *Nat Struct Mol Biol,* 19**,** 893-9.

MARTIN, A. M., KULSKI, J. K., GAUDIERI, S., WITT, C. S., FREITAS, E. M., TROWSDALE, J. & CHRISTIANSEN, F. T. 2004a. Comparative genomic analysis, diversity and evolution of two KIR haplotypes A and B. *Gene,* 335**,** 121-31.

MARTIN, A. M., NOLAN, D., GAUDIERI, S., ALMEIDA, C. A., NOLAN, R., JAMES, I., CARVALHO, F., PHILLIPS, E., CHRISTIANSEN, F. T., PURCELL, A. W., MCCLUSKEY, J. & MALLAL, S. 2004b. Predisposition to abacavir hypersensitivity conferred by HLA-B*5701 and a haplotypic Hsp70-Hom variant. *Proc Natl Acad Sci U S A,* 101**,** 4180-5.

MARTIN, A. M., NOLAN, D., JAMES, I., CAMERON, P., KELLER, J., MOORE, C., PHILLIPS, E., CHRISTIANSEN, F. T. & MALLAL, S. 2005. Predisposition to nevirapine hypersensitivity associated with HLA-DRB1*0101 and abrogated by low CD4 T-cell counts. *AIDS,* 19**,** 97-9.

MARTIN, M. P. & CARRINGTON, M. 2008. KIR locus polymorphisms: genotyping and disease association analysis. *Methods Mol Biol,* 415**,** 49-64.

MARTIN, M. P. & CARRINGTON, M. 2013. Immunogenetics of HIV disease. *Immunol Rev,* 254**,** 245-64.

MARTIN, M. P., GAO, X., LEE, J. H., NELSON, G. W., DETELS, R., GOEDERT, J. J., BUCHBINDER, S., HOOTS, K., VLAHOV, D., TROWSDALE, J., WILSON, M., O'BRIEN, S. J. & CARRINGTON, M. 2002. Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS. *Nat Genet,* 31**,** 429-34.

MARTIN, M. P., QI, Y., GAO, X., YAMADA, E., MARTIN, J. N., PEREYRA, F., COLOMBO, S., BROWN, E. E., SHUPERT, W. L., PHAIR, J., GOEDERT, J. J., BUCHBINDER, S., KIRK, G. D., TELENTI, A., CONNORS, M., O'BRIEN, S. J., WALKER, B. D., PARHAM, P., DEEKS, S. G., MCVICAR, D. W. & CARRINGTON, M. 2007. Innate partnership of HLA-B and KIR3DL1 subtypes against HIV-1. *Nat Genet,* 39**,** 733-40.

MARTINSON, J. J., CHAPMAN, N. H., REES, D. C., LIU, Y. T. & CLEGG, J. B. 1997. Global distribution of the CCR5 gene 32-basepair deletion. *Nat Genet,* 16**,** 100-3.

MARTINSON, N. A., GUPTE, N., MSANDIWA, R., MOULTON, L. H., BARNES, G. L., RAM, M., GRAY, G., HOFFMANN, C. & CHAISSON, R. E. 2014. CD4 and viral load dynamics in antiretroviral-naive HIV-infected adults from Soweto, South Africa: a prospective cohort. *PLoS One,* 9**,** e96369.

MASEMOLA, A., MASHISHI, T., KHOURY, G., MOHUBE, P., MOKGOTHO, P., VARDAS, E., COLVIN, M., ZIJENAH, L., KATZENSTEIN, D., MUSONDA, R., ALLEN, S., KUMWENDA, N., TAHA, T., GRAY, G., MCINTYRE, J., KARIM, S. A., SHEPPARD, H. W., GRAY, C. M. & TEAM, H. S. 2004a. Hierarchical targeting of subtype C human immunodeficiency virus type 1 proteins by CD8+ T cells: correlation with viral load. *J Virol,* 78**,** 3233-43.

MASEMOLA, A. M., MASHISHI, T. N., KHOURY, G., BREDELL, H., PAXIMADIS, M., MATHEBULA, T., BARKHAN, D., PUREN, A., VARDAS, E., COLVIN, M., ZIJENAH, L., KATZENSTEIN, D., MUSONDA, R., ALLEN, S., KUMWENDA, N., TAHA, T., GRAY, G., MCINTYRE, J., KARIM, S. A., SHEPPARD, H. W., GRAY, C. M. & TEAM, H. S. 2004b. Novel and promiscuous CTL epitopes in conserved regions of Gag targeted by individuals with early subtype C HIV type 1 infection from southern Africa. *J Immunol,* 173**,** 4607-17.

MATTHEWS, P. C., LISTGARTEN, J., CARLSON, J. M., PAYNE, R., HUANG, K. H., FRATER, J., GOEDHALS, D., STEYN, D., VAN VUUREN, C., PAIONI, P., JOOSTE, P., OGWU, A., SHAPIRO, R., MNCUBE, Z., NDUNG'U, T., WALKER, B. D., HECKERMAN, D. & GOULDER, P. J. 2012. Co-operative additive effects between HLA alleles in control of HIV-1. *PLoS One,* 7**,** e47799.

MAUGHAN-BROWN, B. 2013. Concurrent sexual partnerships among young adults in Cape Town, South Africa: how is concurrency changing? *Sex Health,* 10**,** 246-52.

MAXWELL, L. D., WALLACE, A., MIDDLETON, D. & CURRAN, M. D. 2002. A common KIR2DS4 deletion variant in the human that predicts a soluble KIR molecule analogous to the KIR1D molecule observed in the rhesus monkey. *Tissue Antigens,* 60**,** 254-8.

MENENDEZ-ARIAS, L. 2013. Molecular basis of human immunodeficiency virus type 1 drug resistance: overview and recent developments. *Antiviral Res,* 98**,** 93-120.

MERINO, A., MALHOTRA, R., MORTON, M., MULENGA, J., ALLEN, S., HUNTER, E., TANG, J. & KASLOW, R. A. 2011. Impact of a functional KIR2DS4 allele on heterosexual HIV-1 transmission among discordant Zambian couples. *J Infect Dis,* 203**,** 487-95.

MERINO, A. M., DUGAST, A. S., WILSON, C. M., GOEPFERT, P. A., ALTER, G., KASLOW, R. A. & TANG, J. 2014. KIR2DS4 promotes HIV-1 pathogenesis: new evidence from analyses of immunogenetic data and natural killer cell function. *PLoS One,* 9**,** e99353.

MERINO, A. M., SABBAJ, S., EASLICK, J., GOEPFERT, P., KASLOW, R. A. & TANG, J. 2013. Dimorphic HLA-B signal peptides differentially influence HLA-E- and natural killer cell-mediated cytolysis of HIV-1-infected target cells. *Clin Exp Immunol,* 174**,** 414-23.

MERINO, A. M., SONG, W., HE, D., MULENGA, J., ALLEN, S., HUNTER, E., TANG, J. & KASLOW, R. A. 2012. HLA-B signal peptide polymorphism influences the rate of HIV-1 acquisition but not viral load. *J Infect Dis,* 205**,** 1797-805.

MIDDLETON, D., GONZALEZ, A. & GILMORE, P. M. 2007. Studies on the expression of the deleted KIR2DS4*003 gene product and distribution of KIR2DS4 deleted and nondeleted versions in different populations. *Hum Immunol,* 68**,** 128-34.

MIDDLETON, D. & GONZELEZ, F. 2010. The extensive polymorphism of KIR genes. *Immunology,* 129**,** 8-19.

MIURA, T., BROCKMAN, M. A., SCHNEIDEWIND, A., LOBRITZ, M., PEREYRA, F., RATHOD, A., BLOCK, B. L., BRUMME, Z. L., BRUMME, C. J., BAKER, B., ROTHCHILD, A. C., LI, B., TROCHA, A., CUTRELL, E., FRAHM, N., BRANDER, C., TOTH, I., ARTS, E. J., ALLEN, T. M. & WALKER, B. D. 2009. HLA-B57/B*5801 human immunodeficiency virus type 1 elite controllers select for rare gag variants associated with reduced viral replication capacity and strong cytotoxic T-lymphocyte [corrected] recognition. *J Virol,* 83**,** 2743-55.

MKHWANAZI, N., THOBAKGALE, C. F., VAN DER STOK, M., REDDY, S., MNCUBE, Z., CHONCO, F., WALKER, B. D., ALTFELD, M., GOULDER, P. J. & NDUNG'U, T. 2010. Immunodominant HIV-1-specific HLA-B- and HLA-C-restricted CD8+ T cells do not differ in polyfunctionality. *Virology,* 405**,** 483-91.

MLISANA, K., WERNER, L., GARRETT, N. J., MCKINNON, L. R., VAN LOGGERENBERG, F., PASSMORE, J. A., GRAY, C. M., MORRIS, L., WILLIAMSON, C., ABDOOL KARIM, S. S. & FOR THE CENTRE FOR THE, A. P. O. R. I. S. A. S. T. 2014. Rapid Disease Progression in HIV-1 Subtype C-Infected South African Women. *Clin Infect Dis*.

MOESTA, A. K., GRAEF, T., ABI-RACHED, L., OLDER AGUILAR, A. M., GUETHLEIN, L. A. & PARHAM, P. 2010. Humans differ from other hominids in lacking an activating NK cell receptor that recognizes the C1 epitope of MHC class I. *J Immunol,* 185**,** 4233-7.

MOSS, J. A. 2013. HIV/AIDS Review. *Radiol Technol,* 84**,** 247-67; quiz p 268-70.

MOUREZ, T., SIMON, F. & PLANTIER, J. C. 2013. Non-M variants of human immunodeficiency virus type 1. *Clin Microbiol Rev,* 26**,** 448-61.

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. *J Virol,* 86**,** 10870-2.

NEOGI, U., PALCHAUDHURI, R. & SHET, A. 2013. High viremia in HIV-1 subtype C infection and spread of the epidemic. *J Infect Dis,* 208**,** 866-7.

NEOGI, U., RAO, S. D., BONTELL, I., VERHEYEN, J., RAO, V. R., GORE, S. C., SONI, N., SHET, A., SCHULTER, E., EKSTRAND, M. L., WONDWOSSEN, A., KAISER, R., MADHUSUDHAN, M. S., PRASAD, V. R. & SONNERBORG, A. 2014. Novel tetra-peptide insertion in Gag-p6 ALIX-binding motif in HIV-1 subtype C associated with protease inhibitor failure in Indian patients. *AIDS,* 28**,** 2319-22.

NGUMBELA, K. C., DAY, C. L., 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. G., ROUSSEAU, C.,

NICKLE, D., MULLINS, J., HECKERMAN, D., KORBER, B., COOVADIA, H., KIEPIELA, P., GOULDER, P. J. & WALKER, B. D. 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 Res Hum Retroviruses,* 24**,** 72-82.

NIEDERKORN, J. Y. & WANG, S. 2005. Immune privilege of the eye and fetus: parallel universes? *Transplantation,* 80**,** 1139-44.

NORCROSS, M. A., LUO, S., LU, L., BOYNE, M. T., GOMARTELI, M., RENNELS, A. D., WOODCOCK, J., MARGULIES, D. H., MCMURTREY, C., VERNON, S., HILDEBRAND, W. H. & BUCHLI, R. 2012. Abacavir induces loading of novel self-peptides into HLA-B*57: 01: an autoimmune model for HLA-associated drug hypersensitivity. *AIDS,* 26**,** F21-9.

NORMAN, P. J., ABI-RACHED, L., GENDZEKHADZE, K., KORBEL, D., GLEIMER, M., ROWLEY, D., BRUNO, D., CARRINGTON, C. V., CHANDANAYINGYONG, D., CHANG, Y. H., CRESPI, C., SARUHAN-DIRESKENELI, G., FRASER, P. A., HAMEED, K., KAMKAMIDZE, G., KORAM, K. A., LAYRISSE, Z., MATAMOROS, N., MILA, J., PARK, M. H., PITCHAPPAN, R. M., RAMDATH, D. D., SHIAU, M. Y., STEPHENS, H. A., STRUIK, S., VERITY, D. H., VAUGHAN, R. W., TYAN, D., DAVIS, R. W., RILEY, E. M., RONAGHI, M. & PARHAM, P. 2007. Unusual selection on the KIR3DL1/S1 natural killer cell receptor in Africans. *Nat Genet,* 39**,** 1092-9.

NOVITSKY, V., GILBERT, P., PETER, T., MCLANE, M. F., GAOLEKWE, S., RYBAK, N., THIOR, I., NDUNG'U, T., MARLINK, R., LEE, T. H. & ESSEX, M. 2003. Association between virus-specific T-cell responses and plasma viral load in human immunodeficiency virus type 1 subtype C infection. *J Virol,* 77**,** 882-90.

NOWAK, I., MAJORCZYK, E., PLOSKI, R., SENITZER, D., SUN, J. Y. & KUSNIERCZYK, P. 2011. Lack of KIR2DL4 gene in a fertile Caucasian woman. *Tissue Antigens,* 78**,** 115-9.

O'CONNOR, G. M., GUINAN, K. J., CUNNINGHAM, R. T., MIDDLETON, D., PARHAM, P. & GARDINER, C. M. 2007. Functional polymorphism of the KIR3DL1/S1 receptor on human NK cells. *J Immunol,* 178**,** 235-41.

OSTROV, D. A., GRANT, B. J., POMPEU, Y. A., SIDNEY, J., HARNDAHL, M., SOUTHWOOD, S., OSEROFF, C., LU, S., JAKONCIC, J., DE OLIVEIRA, C. A., YANG, L., MEI, H., SHI, L., SHABANOWITZ, J., ENGLISH, A. M., WRISTON, A., LUCAS, A., PHILLIPS, E., MALLAL, S., GREY, H. M., SETTE, A., HUNT, D. F., BUUS, S. & PETERS, B. 2012. Drug hypersensitivity caused by alteration of the MHC-presented self-peptide repertoire. *Proc Natl Acad Sci U S A,* 109**,** 9959-64.

OTT, D. E. 2008. Cellular proteins detected in HIV-1. *Rev Med Virol,* 18**,** 159-75.

PANDO, M. J., GARDINER, C. M., GLEIMER, M., MCQUEEN, K. L. & PARHAM, P. 2003. The protein made from a common allele of KIR3DL1 (3DL1*004) is poorly expressed at cell surfaces due to substitution at positions 86 in Ig domain 0 and 182 in Ig domain 1. *J Immunol,* 171**,** 6640-9.

PARHAM, P. & MOFFETT, A. 2013. Variable NK cell receptors and their MHC class I ligands in immunity, reproduction and human evolution. *Nat Rev Immunol,* 13**,** 133-44.

PARHAM, P., NORMAN, P. J., ABI-RACHED, L. & GUETHLEIN, L. A. 2012. Human-specific evolution of killer cell immunoglobulin-like receptor recognition of major histocompatibility complex class I molecules. *Philos Trans R Soc Lond B Biol Sci,* 367**,** 800-11.

PATTERSON, N., PETERSEN, D. C., VAN DER ROSS, R. E., SUDOYO, H., GLASHOFF, R. H., MARZUKI, S., REICH, D. & HAYES, V. M. 2010. Genetic structure of a

unique admixed population: implications for medical research. *Hum Mol Genet,* 19**,** 411-9.

PAULUS, C., LUDWIG, C. & WAGNER, R. 2004. Contribution of the Gag-Pol transframe domain p6* and its coding sequence to morphogenesis and replication of human immunodeficiency virus type 1. *Virology,* 330**,** 271-83.

PAXIMADIS, M., MATHEBULA, T. Y., GENTLE, N. L., VARDAS, E., COLVIN, M., GRAY, C. M., TIEMESSEN, C. T. & PUREN, A. 2012. Human leukocyte antigen class I (A, B, C) and II (DRB1) diversity in the black and Caucasian South African population. *Hum Immunol,* 73**,** 80-92.

PAYNE, R., MUENCHHOFF, M., MANN, J., ROBERTS, H. E., MATTHEWS, P., ADLAND, E., HEMPENSTALL, A., HUANG, K. H., BROCKMAN, M., BRUMME, Z., SINCLAIR, M., MIURA, T., FRATER, J., ESSEX, M., SHAPIRO, R., WALKER, B. D., NDUNG'U, T., MCLEAN, A. R., CARLSON, J. M. & GOULDER, P. J. 2014. Impact of HLA-driven HIV adaptation on virulence in populations of high HIV seroprevalence. *Proc Natl Acad Sci U S A,* 111**,** E5393-400.

PEETERS, M., JUNG, M. & AYOUBA, A. 2013. The origin and molecular epidemiology of HIV. *Expert Rev Anti Infect Ther,* 11**,** 885-96.

PELAK, K., NEED, A. C., FELLAY, J., SHIANNA, K. V., FENG, S., URBAN, T. J., GE, D., DE LUCA, A., MARTINEZ-PICADO, J., WOLINSKY, S. M., MARTINSON, J. J., JAMIESON, B. D., BREAM, J. H., MARTIN, M. P., BORROW, P., LETVIN, N. L., MCMICHAEL, A. J., HAYNES, B. F., TELENTI, A., CARRINGTON, M., GOLDSTEIN, D. B., ALTER, G. & IMMUNOLOGY, N. C. F. H. A. V. 2011. Copy number variation of KIR genes influences HIV-1 control. *PLoS Biol,* 9**,** e1001208.

PETERS, H. O., MENDOZA, M. G., CAPINA, R. E., LUO, M., MAO, X., GUBBINS, M., NAGELKERKE, N. J., MACARTHUR, I., SHEARDOWN, B. B., KIMANI, J., WACHIHI, C., THAVANESWARAN, S. & PLUMMER, F. A. 2008. An integrative bioinformatic approach for studying escape mutations in human immunodeficiency virus type 1 gag in the Pumwani Sex Worker Cohort. *J Virol,* 82**,** 1980-92.

PETERS, P. J., MARSTON, B. J., WEIDLE, P. J. & BROOKS, J. T. 2013. 27 - Human Immunodeficiency Virus Infection. *In:* RYAN, A. J. M. R. H. S. T. (ed.) *Hunter's Tropical Medicine and Emerging Infectious Disease (Ninth Edition).* London: W.B. Saunders.

PETERSON, T. A., KIMANI, J., WACHIHI, C., BIELAWNY, T., MENDOZA, L., THAVANESWARAN, S., NARAYANSINGH, M. J., KARIRI, T., LIANG, B., BALL, T. B., NGUGI, E. N., PLUMMER, F. A. & LUO, M. 2013. HLA class I associations with rates of HIV-1 seroconversion and disease progression in the Pumwani Sex Worker Cohort. *Tissue Antigens,* 81**,** 93-107.

PHILLIPS, E., BARTLETT, J. A., SANNE, I., LEDERMAN, M. M., HINKLE, J., ROUSSEAU, F., DUNN, D., PAVLOS, R., JAMES, I., MALLAL, S. A. & HAAS, D. W. 2013. Associations between HLA-DRB1*0102, HLA-B*5801, and hepatotoxicity during initiation of nevirapine-containing regimens in South Africa. *J Acquir Immune Defic Syndr,* 62**,** e55-7.

PHILLIPS, E. J., WONG, G. A., KAUL, R., SHAHABI, K., NOLAN, D. A., KNOWLES, S. R., MARTIN, A. M., MALLAL, S. A. & SHEAR, N. H. 2005. Clinical and immunogenetic correlates of abacavir hypersensitivity. *AIDS,* 19**,** 979-81.

PICTON, A. C., PAXIMADIS, M. & TIEMESSEN, C. T. 2010. Genetic variation within the gene encoding the HIV-1 CCR5 coreceptor in two South African populations. *Infect Genet Evol,* 10**,** 487-94.

POMPEU, Y. A., STEWART, J. D., MALLAL, S., PHILLIPS, E., PETERS, B. & OSTROV, D. A. 2012. The structural basis of HLA-associated drug hypersensitivity syndromes. *Immunol Rev,* 250**,** 158-66.

PYO, C. W., GUETHLEIN, L. A., VU, Q., WANG, R., ABI-RACHED, L., NORMAN, P. J., MARSH, S. G., MILLER, J. S., PARHAM, P. & GERAGHTY, D. E. 2010. Different patterns of evolution in the centromeric and telomeric regions of group A and B haplotypes of the human killer cell Ig-like receptor locus. *PLoS One,* 5**,** e15115.

QU, W., ROBINSON, M. & ZHANG, F. J. 2008. Factors influencing the natural history of HIV-1 infection. *Chin Med J (Engl),* 121**,** 2613-21.

QUINN, T. C., WAWER, M. J., SEWANKAMBO, N., SERWADDA, D., LI, C., WABWIRE-MANGEN, F., MEEHAN, M. O., LUTALO, T. & GRAY, R. H. 2000. Viral load and heterosexual transmission of human immunodeficiency virus type 1. Rakai Project Study Group. *N Engl J Med,* 342**,** 921-9.

RAGHAVAN, S., SELVARAJ, P., SWAMINATHAN, S. & NARENDRAN, G. 2009. Short communication: association of HLA-A*1101 with resistance and B*4006 with susceptibility to HIV and HIV-TB: an in silico analysis of promiscuous T cell epitopes. *AIDS Res Hum Retroviruses,* 25**,** 1023-8.

RAJAGOPALAN, S. 2010. Endosomal signaling and a novel pathway defined by the natural killer receptor KIR2DL4 (CD158d). *Traffic,* 11**,** 1381-90.

RAUCH, A., NOLAN, D., MARTIN, A., MCKINNON, E., ALMEIDA, C. & MALLAL, S. 2006. Prospective genetic screening decreases the incidence of abacavir hypersensitivity reactions in the Western Australian HIV cohort study. *Clin Infect Dis,* 43**,** 99-102.

ROBINSON, J., HALLIWELL, J. A., MCWILLIAM, H., LOPEZ, R. & MARSH, S. G. 2013a. IPD--the Immuno Polymorphism Database. *Nucleic Acids Res,* 41**,** D1234-40.

ROBINSON, J., HALLIWELL, J. A., MCWILLIAM, H., LOPEZ, R., PARHAM, P. & MARSH, S. G. 2013b. The IMGT/HLA database. *Nucleic Acids Res,* 41**,** D1222-7.

RODRIGUEZ-NOVOA, S., CUENCA, L., MORELLO, J., CORDOBA, M., BLANCO, F., JIMENEZ-NACHER, I. & SORIANO, V. 2010. Use of the HCP5 single nucleotide polymorphism to predict hypersensitivity reactions to abacavir: correlation with HLA-B*5701. *J Antimicrob Chemother,* 65**,** 1567-9.

SAATHOFF, E., PRITSCH, M., GELDMACHER, C., HOFFMANN, O., KOEHLER, R. N., MABOKO, L., MAGANGA, L., GEIS, S., MCCUTCHAN, F. E., KIJAK, G. H., KIM, J. H., ARROYO, M. A., GERHARDT, M., TOVANABUTRA, S., ROBB, M. L., WILLIAMSON, C., MICHAEL, N. L. & HOELSCHER, M. 2010. Viral and host factors associated with the HIV-1 viral load setpoint in adults from Mbeya Region, Tanzania. *J Acquir Immune Defic Syndr,* 54**,** 324-30.

SABIR, H. J., NEHLIN, J. O., QANIE, D., HARKNESS, L., PROKHOROVA, T. A., BLAGOEV, B., KASSEM, M., ISA, A. & BARINGTON, T. 2013. Separate developmental programs for HLA-A and -B cell surface expression during differentiation from embryonic stem cells to lymphocytes, adipocytes and osteoblasts. *PLoS One,* 8**,** e54366.

SACHA, J. B., CHUNG, C., RAKASZ, E. G., SPENCER, S. P., JONAS, A. K., BEAN, A. T., LEE, W., BURWITZ, B. J., STEPHANY, J. J., LOFFREDO, J. T., ALLISON, D. B., ADNAN, S., HOJI, A., WILSON, N. A., FRIEDRICH, T. C., LIFSON, J. D., YANG, O. O. & WATKINS, D. I. 2007. Gag-specific CD8+ T lymphocytes recognize infected cells before AIDS-virus integration and viral protein expression. *J Immunol,* 178**,** 2746-54.

SALIE, M., VAN DER MERWE, L., MOLLER, M., DAYA, M., VAN DER SPUY, G. D., VAN HELDEN, P. D., MARTIN, M. P., GAO, X. J., WARREN, R. M., CARRINGTON, M. & HOAL, E. G. 2014. Associations between human leukocyte antigen class I variants and the Mycobacterium tuberculosis subtypes causing disease. *J Infect Dis,* 209**,** 216-23.

SAMPATHKUMAR, R., PETERS, H. O., MENDOZA, L., BIELAWNY, T., NGUGI, E., KIMANI, J., WACHIHI, C., PLUMMER, F. A. & LUO, M. 2014. Influence of HLA class I haplotypes on HIV-1 seroconversion and disease progression in Pumwani sex worker cohort. *PLoS One,* 9**,** e101475.

SANTOS, S., OBUKHOV, Y., NEKHAI, S., BUKRINSKY, M. & IORDANSKIY, S. 2012. Virus-producing cells determine the host protein profiles of HIV-1 virion cores. *Retrovirology,* 9**,** 65.

SASSEVILLE, V. G., ROTTMAN, J. B., DU, Z., VEAZEY, R., KNIGHT, H. L., CAUNT, D., DESROSIERS, R. C. & LACKNER, A. A. 1998. Characterization of the cutaneous exanthem in macaques infected with a Nef gene variant of SIVmac239. *J Invest Dermatol,* 110**,** 894-901.

SCHMITZ, J. E., KURODA, M. J., SANTRA, S., SASSEVILLE, V. G., SIMON, M. A., LIFTON, M. A., RACZ, P., TENNER-RACZ, K., DALESANDRO, M., SCALLON, B. J., GHRAYEB, J., FORMAN, M. A., MONTEFIORI, D. C., RIEBER, E. P., LETVIN, N. L. & REIMANN, K. A. 1999. Control of viremia in simian immunodeficiency virus infection by CD8+ lymphocytes. *Science,* 283**,** 857-60.

SHAW, J. & KOLLNBERGER, S. 2012. New perspectives on the ligands and function of the killer cell immunoglobulin-like receptor KIR3DL2 in health and disease. *Front Immunol,* 3**,** 339.

SHEPHERD, B. L., FERRAND, R., MUNYATI, S., FOLKARD, S., BOYD, K., BANDASON, T., JALLOW, S., ROWLAND-JONES, S. L. & YINDOM, L. M. 2015. HLA Correlates of Long-Term Survival in Vertically Infected HIV-1-Positive Adolescents in Harare, Zimbabwe. *AIDS Res Hum Retroviruses*.

SHOSTAKOVICH-KORETSKAYA, L., CATANO, G., CHYKARENKO, Z. A., HE, W., GORNALUSSE, G., MUMMIDI, S., SANCHEZ, R., DOLAN, M. J., AHUJA, S. S., CLARK, R. A., KULKARNI, H. & AHUJA, S. K. 2009. Combinatorial content of CCL3L and CCL4L gene copy numbers influence HIV-AIDS susceptibility in Ukrainian children. *AIDS,* 23**,** 679-88.

SIDNEY, J., PETERS, B., FRAHM, N., BRANDER, C. & SETTE, A. 2008. HLA class I supertypes: a revised and updated classification. *BMC Immunol,* 9**,** 1.

SIMONITSCH, I., GEUSAU, A., CHOTT, A. & JURECKA, W. 2000. Cutaneous dendritic cells are main targets in acute HIV-1-infection. *Mod Pathol,* 13**,** 1232-7.

SONG, W., HE, D., BRILL, I., MALHOTRA, R., MULENGA, J., ALLEN, S., HUNTER, E., TANG, J. & KASLOW, R. A. 2011. Disparate associations of HLA class I markers with HIV-1 acquisition and control of viremia in an African population. *PLoS One,* 6**,** e23469.

SPIVAK, A. M., BRENNAN, T. P., O'CONNELL, K. A., SYDNOR, E., WILLIAMS, T. M., SILICIANO, R. F., GALLANT, J. E. & BLANKSON, J. N. 2010. A case of seronegative HIV-1 infection. *J Infect Dis,* 201**,** 341-5.

STERN, M., DE ANGELIS, C., URBANI, E., MANCUSI, A., AVERSA, F., VELARDI, A. & RUGGERI, L. 2010. Natural killer-cell KIR repertoire reconstitution after haploidentical SCT. *Bone Marrow Transplant,* 45**,** 1607-10.

STERN, M., RUGGERI, L., CAPANNI, M., MANCUSI, A. & VELARDI, A. 2008. Human leukocyte antigens A23, A24, and A32 but not A25 are ligands for KIR3DL1. *Blood,* 112**,** 708-10.

STOCCHI, L., CASCELLA, R., ZAMPATTI, S., PIRAZZOLI, A., NOVELLI, G. & GIARDINA, E. 2012. The Pharmacogenomic HLA Biomarker Associated to Adverse Abacavir Reactions: Comparative Analysis of Different Genotyping Methods. *Curr Genomics,* 13**,** 314-20.

SUNDQUIST, W. I. & KRAUSSLICH, H. G. 2012. HIV-1 assembly, budding, and maturation. *Cold Spring Harb Perspect Med,* 2**,** a006924.

TAMIYA, S., MARDY, S., KAVLICK, M. F., YOSHIMURA, K. & MISTUYA, H. 2004. Amino acid insertions near Gag cleavage sites restore the otherwise compromised replication of human immunodeficiency virus type 1 variants resistant to protease inhibitors. *J Virol,* 78**,** 12030-40.

TANG, J., MALHOTRA, R., SONG, W., BRILL, I., HU, L., FARMER, P. K., MULENGA, J., ALLEN, S., HUNTER, E. & KASLOW, R. A. 2010. Human leukocyte antigens and HIV type 1 viral load in early and chronic infection: predominance of evolving relationships. *PLoS One,* 5**,** e9629.

TANG, J., SHAO, W., YOO, Y. J., BRILL, I., MULENGA, J., ALLEN, S., HUNTER, E. & KASLOW, R. A. 2008. Human leukocyte antigen class I genotypes in relation to heterosexual HIV type 1 transmission within discordant couples. *J Immunol,* 181**,** 2626-35.

TANG, J., TANG, S., LOBASHEVSKY, E., MYRACLE, A. D., FIDELI, U., ALDROVANDI, G., ALLEN, S., MUSONDA, R., KASLOW, R. A. & ZAMBIA, U. A. B. H. I. V. R. P. 2002. Favorable and unfavorable HLA class I alleles and haplotypes in Zambians predominantly infected with clade C human immunodeficiency virus type 1. *J Virol,* 76**,** 8276-84.

TEBIT, D. M., NDEMBI, N., WEINBERG, A. & QUINONES-MATEU, M. E. 2012. Mucosal transmission of human immunodeficiency virus. *Curr HIV Res,* 10**,** 3-8.

TENZER, S., WEE, E., BURGEVIN, A., STEWART-JONES, G., FRIIS, L., LAMBERTH, K., CHANG, C. H., HARNDAHL, M., WEIMERSHAUS, M., GERSTOFT, J., AKKAD, N., KLENERMAN, P., FUGGER, L., JONES, E. Y., MCMICHAEL, A. J., BUUS, S., SCHILD, H., VAN ENDERT, P. & IVERSEN, A. K. 2009. Antigen processing influences HIV-specific cytotoxic T lymphocyte immunodominance. *Nat Immunol,* 10**,** 636-46.

TIEMESSEN, C. T., PAXIMADIS, M., MINEVICH, G., WINCHESTER, R., SHALEKOFF, S., GRAY, G. E., SHERMAN, G. G., COOVADIA, A. H. & KUHN, L. 2011. Natural killer cell responses to HIV-1 peptides are associated with more activating KIR genes and HLA-C genes of the C1 allotype. *J Acquir Immune Defic Syndr,* 57**,** 181-9.

UHRBERG, M. 2005. The KIR gene family: life in the fast lane of evolution. *Eur J Immunol,* 35**,** 10-5.

UMAPATHY, S., PAWAR, A. & GHOSH, K. 2007. Specific human leukocyte antigen alleles associated with HIV-1 infection in an Indian population. *J Acquir Immune Defic Syndr,* 44**,** 489-90.

VAN BERGEN, J., THOMPSON, A., HAASNOOT, G. W., ROODNAT, J. I., DE FIJTER, J. W., CLAAS, F. H., KONING, F. & DOXIADIS, II 2011. KIR-ligand mismatches are associated with reduced long-term graft survival in HLA-compatible kidney transplantation. *Am J Transplant,* 11**,** 1959-64.

VAN MANEN, D., GRAS, L., BOESER-NUNNINK, B. D., VAN SIGHEM, A. I., MAURER, I., MANGAS RUIZ, M. M., HARSKAMP, A. M., STEINGROVER, R., PRINS, J. M., DE WOLF, F., VAN 'T WOUT, A. B., SCHUITEMAKER, H. & DUTCH, H. I. V. M. F. H. I. V. H. G. S. 2011. Rising HIV-1 viral load set point at a population level coincides with a fading impact of host genetic factors on HIV-1 control. *AIDS,* 25**,** 2217-26.

VAN TEIJLINGEN, N. H., HOLZEMER, A., KORNER, C., GARCIA-BELTRAN, W. F., SCHAFER, J. L., FADDA, L., SUSCOVICH, T. J., BRANDER, C., CARRINGTON, M., EVANS, D. T., VAN BAARLE, D. & ALTFELD, M. 2014. Sequence variations

in HIV-1 p24 Gag-derived epitopes can alter binding of KIR2DL2 to HLA-C*03:04 and modulate primary natural killer cell function. *AIDS,* 28**,** 1399-408.

VENDELBOSCH, S., DE BOER, M., GOUW, R. A., HO, C. K., GEISSLER, J., SWELSEN, W. T., MOORHOUSE, M. J., LARDY, N. M., ROOS, D., VAN DEN BERG, T. K. & KUIJPERS, T. W. 2013. Extensive variation in gene copy number at the killer immunoglobulin-like receptor locus in humans. *PLoS One,* 8**,** e67619.

VERMA, R., GUPTA, R. B., SINGH, K., BHASIN, R., ANAND SHUKLA, A., CHAUHAN, S. S. & LUTHRA, K. 2007. Distribution of CCR5delta32, CCR2-64I and SDF1-3'A and plasma levels of SDF-1 in HIV-1 seronegative North Indians. *J Clin Virol,* 38**,** 198-203.

VERPLANK, L., BOUAMR, F., LAGRASSA, T. J., AGRESTA, B., KIKONYOGO, A., LEIS, J. & CARTER, C. A. 2001. Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55(Gag). *Proc Natl Acad Sci U S A,* 98**,** 7724-9.

VILCHES, C., CASTANO, J., GOMEZ-LOZANO, N. & ESTEFANIA, E. 2007. Facilitation of KIR genotyping by a PCR-SSP method that amplifies short DNA fragments. *Tissue Antigens,* 70**,** 415-22.

VILCHES, C., GARDINER, C. M. & PARHAM, P. 2000. Gene structure and promoter variation of expressed and nonexpressed variants of the KIR2DL5 gene. *J Immunol,* 165**,** 6416-21.

VITEZICA, Z. G., MILPIED, B., LONJOU, C., BOROT, N., LEDGER, T. N., LEFEBVRE, A. & HOVNANIAN, A. 2008. HLA-DRB1*01 associated with cutaneous hypersensitivity induced by nevirapine and efavirenz. *AIDS,* 22**,** 540-1.

WANG, Y. E., LI, B., CARLSON, J. M., STREECK, H., GLADDEN, A. D., GOODMAN, R., SCHNEIDEWIND, A., POWER, K. A., TOTH, I., FRAHM, N., ALTER, G., BRANDER, C., CARRINGTON, M., WALKER, B. D., ALTFELD, M., HECKERMAN, D. & ALLEN, T. M. 2009. Protective HLA class I alleles that restrict acute-phase CD8+ T-cell responses are associated with viral escape mutations located in highly conserved regions of human immunodeficiency virus type 1. *J Virol,* 83**,** 1845-55.

WATTS, J. M., DANG, K. K., GORELICK, R. J., LEONARD, C. W., BESS, J. W., JR., SWANSTROM, R., BURCH, C. L. & WEEKS, K. M. 2009. Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature,* 460**,** 711-6.

WILLIAMS, F., MAXWELL, L. D., HALFPENNY, I. A., MEENAGH, A., SLEATOR, C., CURRAN, M. D. & MIDDLETON, D. 2003. Multiple copies of KIR 3DL/S1 and KIR 2DL4 genes identified in a number of individuals. *Hum Immunol,* 64**,** 729-32.

WONG-BAEZA, I., RIDLEY, A., SHAW, J., HATANO, H., RYSNIK, O., MCHUGH, K., PIPER, C., BRACKENBRIDGE, S., FERNANDES, R., CHAN, A., BOWNESS, P. & KOLLNBERGER, S. 2013. KIR3DL2 binds to HLA-B27 dimers and free H chains more strongly than other HLA class I and promotes the expansion of T cells in ankylosing spondylitis. *J Immunol,* 190**,** 3216-24.

WOOD, R., ARASTEH, K., STELLBRINK, H. J., TEOFILO, E., RAFFI, F., POLLARD, R. B., ERON, J., YEO, J., MILLARD, J., WIRE, M. B. & NADERER, O. J. 2004. Six-week randomized controlled trial to compare the tolerabilities, pharmacokinetics, and antiviral activities of GW433908 and amprenavir in human immunodeficiency virus type 1-infected patients. *Antimicrob Agents Chemother,* 48**,** 116-23.

WOROBEY, M., GEMMEL, M., TEUWEN, D. E., HASELKORN, T., KUNSTMAN, K., BUNCE, M., MUYEMBE, J. J., KABONGO, J. M., KALENGAYI, R. M., VAN MARCK, E., GILBERT, M. T. & WOLINSKY, S. M. 2008. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature,* 455**,** 661-4.

YAMAGUCHI, J., VALLARI, A. S., SWANSON, P., BODELLE, P., KAPTUE, L., NGANSOP, C., ZEKENG, L., GURTLER, L. G., DEVARE, S. G. & BRENNAN, C. A. 2002. Evaluation of HIV type 1 group O isolates: identification of five phylogenetic clusters. *AIDS Res Hum Retroviruses,* 18**,** 269-82.

YUAN, J., GUO, S., HALL, D., CAMMETT, A. M., JAYADEV, S., DISTEL, M., STORFER, S., HUANG, Z., MOOTSIKAPUN, P., RUXRUNGTHAM, K., PODZAMCZER, D., HAAS, D. W. & NEVIRAPINE TOXICOGENOMICS STUDY, T. 2011. Toxicogenomics of nevirapine-associated cutaneous and hepatic adverse events among populations of African, Asian, and European descent. *AIDS,* 25**,** 1271-80.

YUEN, G. J., WELLER, S. & PAKES, G. E. 2008. A review of the pharmacokinetics of abacavir. *Clin Pharmacokinet,* 47**,** 351-71.

YUNIS, E. J., ROMERO, V., DIAZ-GIFFERO, F., ZUNIGA, J. & KOKA, P. 2007. Natural Killer Cell Receptor NKG2A/HLA-E Interaction Dependent Differential Thymopoiesis of Hematopoietic Progenitor Cells Influences the Outcome of HIV Infection. *J Stem Cells,* 2**,** 237-248.

ZHANG, W., WANG, L., HONG, K., LIU, Y., SU, B., XU, C., XU, J., RUAN, Y., SHAO, Y. & ZHENG, D. 2013. Frequency of HLA-A 03 associates with HIV-1 infection in a Chinese cohort. *Sci China Life Sci,* 56**,** 1014-9.

ZHAO, G., PERILLA, J. R., YUFENYUY, E. L., MENG, X., CHEN, B., NING, J., AHN, J., GRONENBORN, A. M., SCHULTEN, K., AIKEN, C. & ZHANG, P. 2013. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature,* 497**,** 643-6.

ZHAO, K., ISHIDA, Y., OLEKSYK, T. K., WINKLER, C. A. & ROCA, A. L. 2012. Evidence for selection at HIV host susceptibility genes in a West Central African human population. *BMC Evol Biol,* 12**,** 237.

ZHU, P., LIU, J., BESS, J., JR., CHERTOVA, E., LIFSON, J. D., GRISE, H., OFEK, G. A., TAYLOR, K. A. & ROUX, K. H. 2006. Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature,* 441**,** 847-52.

ZHU, T., KORBER, B. T., NAHMIAS, A. J., HOOPER, E., SHARP, P. M. & HO, D. D. 1998. An African HIV-1 sequence from 1959 and implications for the origin of the epidemic. *Nature,* 391**,** 594-7.

ZIPPERLEN, K., GALLANT, M., STAPLETON, S., HEATH, J., BARRETT, L. & GRANT, M. 2014. Protective genotypes in HIV infection reflect superior function of KIR3DS1 over KIR3DL1 CD8 T cells. *Immunol Cell Biol*.