ACQUIRED AND TRANSMITTED DRUG RESISTANCE IN HIV-1 SUBTYPE C: IMPLICATIONS OF NOVEL MUTATIONS ON REPLICATION CAPACITY, CLEAVAGE AND DRUG SUSCEPTIBILITY

BY: URISHA SINGH

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy in Virology

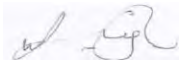
Nelson R. Mandela School of Medicine, College of Health Sciences,
University of KwaZulu-Natal, Durban, South Africa.

2015

Preface

The experimental work described in this thesis was conducted at the HIV Pathogenesis Programme Hasso Platner Research Laboratory and the Africa Centre laboratory, Doris Duke Medical Research Institute, Nelson R Mandela School of Medicine, University of KwaZulu-Natal, Durban, South Africa, from March 2013 to November 2015, under the supervision of Dr Michelle Lucille Gordon.

This work has not been submitted in any form for any degree or diploma to any tertiary institution, where use has been made of the work of others, it is duly acknowledged in the text.

U. Singh  Date 14 March 2016

Dr M.L. Gordon  Date 14 March 2016

As the candidates supervisor I agree to the submission of this thesis:

Dr Michelle Lucille Gordon  Date 14 March 2016

Declaration

I Urisha Singh declare that:

1. The research reported in this dissertation, except where otherwise indicated, is my original research.

2. This thesis has not been submitted for any degree or examination at any other university.

3. This data does not contain other person's data, graphs, pictures, or other information, unless specifically acknowledged as being sourced from other persons.

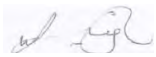
4. This dissertation does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

    a) Their words have been re-written but the general information attributed to them has been referenced.

    b) Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.

    c) This dissertation does not contain text, graphics or tables copied and pasted from the internet unless specifically acknowledged, and the source being detailed in the thesis and reference section.

5. Where I have reproduced a publication of which I am author, co-author or editor, I have indicated in detail which part of the publication was actually written by myself alone and have fully referenced such publications.

Signed  Date 14 March 2016

Publications and presentations

Peer reviewed publications:

Singh, U., Singh, A., Noguera-Julian, M., Jaggernath, M., Moodley, A., Reddy, T., Dong, K., Walker, B.D., Ndung'u, T., Gordon, M.L. 2015. Low frequency drug resistance mutations are common in HIV-1 subtype C acute infection. *AIDS*, submitted for review.

Conference Presentations

Singh, U., Singh, A., Moodley, K., Dong, K., Walker, B.D., Ndung'u, T., Gordon, M.L. Evolution of transmitted drug resistance associated mutations in hiv-1 subtype c infected females from an acute infection cohort, Durban, South Africa. 7th SA AIDS Conference, 9–12 June 2015, Durban, South Africa.

Singh, U., Singh, A., Moodley, A., Dong, K., Walker, B.D., Ndung'u T., Gordon, M.L. 2015 Transmitted drug resistance in an acute infection cohort from Durban, South Africa: a view of Sanger sequencing versus ultra- deep sequencing. Abstract 95. XXIV International Drug Resistance Workshop, 21-22 February 2015, Seattle, United States of America.

Singh, U., Pillay, K., Singh, A., Ndung'u, T., Gordon, M.L. Gag-Pro mutations associated with altered replication capacity in HIV-1 subtype C infected individuals failing a Protease inhibitor inclusive treatment regimen. College of Health Sciences Research Symposium, 11-12 September 2014, University of KwaZulu-Natal, Durban South Africa.

Singh, U., Singh, A., Moodley, A., Dong, K., Walker, B.D., Ndung'u T., Gordon, M.L. Transmitted Drug Resistance in HIV-1 Subtype C Hyperacute Infection. Conference on Retroviruses and Opportunistic Infections, 22-25 February 2016, Boston Massachusetts.

Statement

The following publication has been included as a chapter in this thesis (i.e. Chapter 5):

Singh, U., Singh, A., Noguera-Julian, M., Jaggernath, M., Moodley, A., Reddy, T., Dong, K., Walker, B.D., Ndung'u, T., Gordon, M.L. 2015. Low frequency drug resistance mutations are common in HIV-1 subtype C acute infection. AIDS, submitted for review.

The PhD candidate performed experimental work described in this publication, where others have made contributions it is duly acknowledged in the text. The candidate drafted this publication in full and it has been reviewed by co-authors.

U. Singh  Date 14 March 2016

Dr Michelle Lucille Gordon  Date 14 March 2016

This thesis is dedicated to my loving Mum, Dad and Husband.

Acknowledgements

I would like to thank:

Dr Michelle Gordon for her supervision, guidance, encouragement, support and great talks during my PhD studies. Thank you for sharing your knowledge with me. It has been a pleasure being mentored by you.

Professor Thumbi Ndung'u for his mentorship, encouragement and motivation to strive for the best.

Dr Jaclyn Mann for her unbelievable kindness, for teaching me laboratory techniques and always helping with my many questions. I am so grateful to you.

Dr Zaza Ndhlovu for his great friendship and belief in me and for sharing data effortlessly.

Saleha Omarjee, Dr Kavidha Reddy, Dr Ravesh Singh, Mpume Mkhwanazi and Bongiwe Ndhlovu for demonstrating laboratory techniques and great company in the laboratory.

Dr Santiago Avila, Dr Luke Swenson, Marissa Perez, Maribel Soto-Nava, Tania Escamilla Gomez and Thalia G. Tellez for their hospitality, demonstration of laboratory techniques and friendship.

Professor Mark Brockman and Dr Zabrina Brumme are thanked for GXR cells and for access to the codon by codon analysis tool.

Avashna Singh and Kelly Pillay for their kind assistance with laboratory work and the entertaining conversations.

The Africa Centre Laboratory for providing access to tissue culture facilities.

Students and staff at the HIV Pathogenesis Programme, Hasso Platner Research Laboratory and Africa Centre Laboratory for their kind assistance in the laboratory.

The National Research Foundation, The Canada sub-Saharan Africa HIV/AIDS Network, HIV Pathogenesis Programme and University of KwaZulu-Natal College of Health Sciences for financial support.

The PCS cohort, FRESH cohort and HPP Acute cohort study teams and participants for their invaluable contribution to this study.

My dearest parents for their unconditional love and support. Thank you both for the million phone calls and endless words of encouragement. With you guys in my life anything is possible.

My sisters for their love and encouragement.

My dear husband Clinton Reddy for sharing this journey with me. I am so grateful for your love, advice, support and kindness. Thank you for your patience and for taking care of me every step of the way. I am truly blessed to have you in my life.

Abstract

Introduction

Large scale roll-out of combination antiretroviral therapy (cART) has been successful in improving the quality of life of HIV-1 infected individuals in South Africa (SA). However the development and transmission of drug resistance threatens the future success and longevity of cART in the country. Studies have shown that resistance to Protease inhibitors (PI's), in the absence of mutations in Protease (PR), is increasing in SA. Whilst some studies attribute this to poor treatment adherence, others have shown that mutations in Gag contribute to PI resistance. The majority of these studies however have been conducted on HIV-1 subtype B, despite HIV-1 subtype C being the most prevalent subtype globally. Given that Gag is highly polymorphic between subtypes, studies focusing on HIV-1 subtype C are required. Despite the high rate of virologic failure of patients on PI inclusive treatment regimens, no transmitted drug resistance (TDR) studies have identified PI associated TDR mutations. This could be due to the high fitness cost associated with PR mutations which would result in rapid reversion or low frequency of mutations within the viral quasispecies. Most TDR studies in SA, as in other resource limited settings, have used recently infected cohorts to measure TDR. It is however unlikely that rapidly reverting mutations would be detected in recent infection. Furthermore, these studies have all used Sanger sequencing which only detects mutations at frequencies >15-20%. With recent studies showing that low frequency mutations present at frequencies as low as 1% impact treatment outcomes, the elucidation of these mutations using deep-sequencing techniques is necessary. For a true measure of TDR, studies employing acute infection cohorts and deep-sequencing techniques are required.

The current study aimed to identify mutations in Gag-Protease associated with PI resistance/exposure, and to determine their impact on replication capacity and drug susceptibility. The prevalence of low frequency TDR mutations in an HIV-1 subtype C acute infection cohort was also investigated.

Methods

A cohort of 80 HIV-1 subtype C infected participants failing a PI inclusive treatment regimen (i.e. PCS cohort) from 2009–2013 in Durban, South Africa was used to assess the role of Gag in PI resistance. Gag mutations were divided into three groups: PI exposure associated Gag

mutations; resistance associated Gag mutations (rGag) and novel Gag mutations (nGag). Frequencies of each of these mutations were compared amongst: 80 PCS cohort sequences, 2,481 HIV-1 subtype B treatment naïve sequences, 954 HIV-1 subtype C treatment naïve sequences and 54 HIV-1 subtype C sequences from acutely infected individuals, in order to identify PI associated mutations and natural polymorphisms. Next, recombinant viruses for all 80 participants were generated by co-transfection of a CEM derived T-cell line (i.e. GXR cells) with an NL43-deleted-*gag-protease* (NL43Δ*gag-protease*) backbone and patient derived Gag-Protease amplicons. Thereafter, the replication capacity of each virus was assessed using a replication assay that employed a green fluorescent protein reporter cell line and flow cytometry. Associations between replication capacity and Gag-Protease mutations were established. Eighteen viruses with mutations of interest were then selected for use in drug susceptibility assays, where the impact of mutations on susceptibility to lopinavir (LPV) and darunavir (DRV) was assessed in a luciferase based assay. Lastly, the impact of novel Gag mutations on replication capacity and drug susceptibility was validated by generating site-directed mutant viruses with mutations of interest and using these mutant viruses in replication capacity and drug susceptibility assays. Furthermore the cleavage profile of each site-directed mutant virus was established by western blotting.

Samples available from 47 HIV-1 subtype C acutely infected individuals collected from 2007-2014 in Durban, South Africa, was used to assess low frequency TDR mutations in HIV-1 subtype C acute infection. Firstly the RT and PR region of each virus was genotyped using the Viroseq HIV-1 genotyping system in order to identify the prevalence of TDR in the cohort. Thereafter 14 participant samples were selected, based on the availability of plasma at one week after onset of plasma viremia (OPV), for sequencing by ultra-deep pyrosequencing (UDPS). This served to identify low frequency mutations. Comparisons in TDR prevalence was made between Sanger sequencing and UDPS. Thereafter, the impact of low frequency TDR mutations on treatment outcomes was assessed by comparing time to virologic suppression for two participants with low frequency mutations to that of four participants without low frequency mutations.

Results

Protease resistance associated mutations (RAMs) occurred in 34/80 (42.5%) participants, whilst Gag mutations associated with PI resistance in subtype B were detected in 67/80 (84%) participants. Overall, 12 Gag mutations associated with PI exposure (i.e. E12K, V35I, G62R, V370A/M, S373P/Q/T, A374P, T375N, I376V, G381S, I389T, I401T and H219Q), eight rGag mutations (i.e. R76K, Y79F, V128I, A431V, K436R, L449F, R452K and P453L) and four nGag mutations (i.e. Q69K, S111C/I, T239A/S and I256V) were identified in the PCS cohort. The

E12K, V370A/M, T375N, G381S, R76K and Y79F mutations all occurred as natural polymorphism in HIV-1 subtype C. The A431V, K436R, L449F, R452K, P453L, Q69K, S111C/I, T239A/S and I256V mutations were all associated with PI resistance/exposure. Interestingly all viruses with PR RAMs harboured rGag and nGag mutations, however rGag and nGag mutations were also found to occur without PR RAMs.

Protease RAMs were associated with significantly reduced replication capacity. The K335R and A431V mutations were the only Gag mutations associated with significantly reduced replication capacity.

Viruses with PR RAMs were associated with significantly reduced susceptibility to LPV (>15 FC in IC_{50}) and DRV (>6 FC in IC_{50}). Furthermore, the following combinations of rGag and nGag mutations were found to confer reduced susceptibility to LPV and DRV in the absence of PR RAMs: R76K+Y79F+K436R+L449P+I256V (5.2 fold increase in IC_{50} for DRV), R76K+R453L (23.88 fold increase in IC_{50} for LPV and a 6.73 fold increase in IC_{50} for DRV) and R76K+K436R+Q69K+S111C (7.40 fold increase in IC_{50} for LPV).

Analysis of recombinant viruses showed that the Q69K nGag mutation rescued replication capacity of all viruses harbouring A431V+PR RAMs. This was validated by SDM, where Q69K rescued the replication capacity of site-directed mutant viruses harbouring A431V+V82A. The Q69K mutation was also associated with increasing polyprotein cleavage when found in conjunction with A431V+V82A.

With regards to TDR, we demonstrated a prevalence of 57% of TDR mutations with UDPS and 2.2% with Sanger sequencing. Sanger sequencing identified the K103N non-nucleoside reverse transcriptase inhibitor (NNRTI)-associated TDR mutation. In addition to K103N (frequency: >99%), the following low frequency mutations were detected by UDPS: the K65R (1-1.5%) and D67N (3.88%) nucleotide reverse transcriptase inhibitor (NRTI)-associated TDR mutations, the F53L (17.6%) and M46L (6.3%) Protease inhibitor (PI)-associated TDR mutations, and the T97A (2.90%) integrase strand transfer inhibitor (InSTI)-associated TDR mutations. Participants with low frequency TDR mutations took 40 days longer to achieve viral suppression than participants without low frequency TDR mutations, when placed on fixed dose combination antiretroviral therapy.

Conclusion

Most participants experiencing PI failure did not harbour PR RAMs. The majority however did harbour Gag mutations which we show can confer resistance to PI's in the absence or presence of PR RAMs. These Gag mutations can function either as primary resistance mutations, causing resistance to PI's in the absence of PR RAMs, or as compensatory mutations where they enhance polyprotein cleavage which manifests as improved replication capacity. The presence of rGag mutations in the absence of PR RAMs suggests that the development of Gag mutations may precede the development of PR RAMs and could play a role in PR RAM development. Gag mutations could therefore be a precursor to indicate PI resistance and should thus be included in PI resistance algorithms. The significant variations in Gag between HIV-1 subtype B and HIV-1 subtype C highlights that research on HIV-1 subtype B cannot always be translated to HIV-1 subtype C. With regards to TDR, low frequency mutations to PI's, NRTI's, NNRTI's and InSTI's are common in HIV-1 subtype C acute infection and can impact treatment outcomes. Their identification however is dependent upon the use of deep sequencing technologies, highlighting the need for cost effective deep sequencing technologies for use in low income countries such as SA.

Table of contents

List of figures

List of tables

Abbreviations

µg	Microgram
µl	Microliter
µM	Micromolar
ml	Milliliter
ºC	Degrees Celsius
3TC	Lamivudine
aa	Amino acid
ABC	Abacavir
AI	Acute infection
AIDS	Acquired immunodeficiency syndrome
ALIX	ALG2 interacting protein
ANRS	French AIDS research agency
APOBEC3G	Apolipoprotein B mRNA-editing enzyme-catalyticpolypeptide-like 3G
ANOVA	Analysis of Variance
ART	Antiretroviral therapy
ARV	Antiretroviral
ATZ	Atazanavir
AZT	Azidothymidine, Zidovudine
bp	Base pairs
CA	CApsid
cART	Combination antiretroviral therapy
CCR5	C-C chemokine receptor 5
CD4	Cluster of differentiation 4
CDC	Center for disease control
cPPT	Central polypurine tract
CPSF-6	cleavage and polyadenylation specificity factor 6
CRF	Circulating recombinant form
CS	Cleavage site
CSMs	Cleavage site mutations
CTS	Central termination signal
CYP45034A	Cytochrome P450 34A
cypA	Cyclophillin A
CXCR4	CXC chemokine receptor 4
DEAE Dextran	Diethylaminoethyl-dextran hydrochloride
DEPC	Diethyl pyrocarbonate
DFOPV	Days following onset of plasma viremia
d4T	Stavudine
DRM	Drug resistance mutation
DRV	Darunavir
ddi	Didanosine
ddntp	Dideoxynucleotide triphosphate
dntp	Deoxynucleotide triphosphate
dsDNA	Double stranded deoxyribonucleic acid
DTG	Dolutegravir
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethylsulfoxide
ECL2	Extracellular loop 2
EDTA	Ethylenediaminetetraacetic acid

EFV	Efavirenz
Env	Envelope
ETR	Etravirine
EVG	Elvitegravir
FBS	Foetal bovine serum
FC	Fold change
FRESH	Females Rising through Education, Support and Health
FPV	Fosamprenavir
FTC	Emitricitabine
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme linked immunosorbent assay
emPCR	Emulsion polymerase chain reaction
ESCRT	Endosomal sorting complex required for transport
FDA	Food and drug administration
FDC	Fixed dose combination
FI	Fusion inhibitor
Gag	Group specific antigen
GALT	Gut-associated lymphoid tissue
GFP	Green fluorescent protein
Gp(X)	Glycoprotein of (X) kDA
HEPES	N-2-hydroxyethylpiperazine-N'-2ethanesulfonic acid
HIV	Human Immunodeficiency Virus
HIV-1	Human Immunodeficiency virus type 1
HIV-2	Human Immunodeficiency virus type 2
HIVdb	HIV database
HLA	Human leucocyte antigen
HPP	HIV Pathogenesis Programme
HR-1	Heptad region 1
HR-2	Heptad region 2
HR-N	Amino terminal helical region
HRP	Horse radish peroxidase
HTLV-III	Human T-lymphotropic virus type III
IAS-USA	International AIDS society – United States of America
IDV	Indinavir
IL-1	Interleukin 1
IL-6	Interleukin 6
INT	Integrase
InSTI	Integrase strand transfer inhibitor
IPTG	Isopropyl-1-thio-β-D-galctopyranoside
KS	Kaposi's Sarcoma
KZN	Kwa-Zulu Natal
LAV	Lymphadenopathy associated virus
LB	Luria Bertani
L-domains	Late assembly domains
LPS	Lipopolysaccharide
LPV	Lopinavir
LTR	Long terminal repeat
MA	Matrix
MID	Multiplex identifiers
MIP-1	Macrophage inflammatory protein-1
MHC-I	Major histocompatibility complex I
MHC-II	Major histocompatibility complex II
mRNA	Messenger ribonucleic acid

MSM	Men who have sex with men
MVC	Maraviroc
MTT	3-(4, 5-dimethylthiazol)—2, 5 –diphenyltetrazolium bromide
MOI	Multiplicity of Infection
NBP	NNRTI binding pocket
NC	Nucleocapsid
Non-CS	Non-cleavage site
NDOH	National department of health
Nef	Negative regulation factor
NFV	Nelfinavir
NKC	Natural killer cell
nGag mutations	Novel Gag mutations
NNRTI	Non-nucleoside reverse transcriptase inhibitor
NPC	Nuclear pore complex
NRTI	Nucleoside reverse transcriptase inhibitor
NL43Δgag-protease	NL43 deleted gag protease
NUP153	Nucleoporin 153
NUP358	Nucleoporin 358
NVP	Nevirapine
OPV	Onset of plasma viremia
P1	Protein of 1 kDA
P2	Protein of 2 kDA
P6	Protein of 6kDA
PBMC	Peripheral blood mononuclear cells
pbs	Primer binding site
PBS	Phosphate buffered saline
PCP	*Pneumocystis Carinii* Pneumonia
PCR	Polymerase chain reaction
PCS	Protein Cleavage Study
PFA	Paraformaldehyde
PI	Protease inhibitor
PIC	Pre-integration complex
PTdinS(4,5)P2	phosphoinositide phosphatidylinositol-4,5-bisphosphate
pTEF-b	positive transcription elongation factor b
PPT	Polypurine tract
Pol	Polymerase
PR	Protease
$Pr55^{gag}$	Precursor of 55 kDA
PVDF	Polyvinylide fluoride
QCMD	Quality control for molecular diagnostics
RAL	Raltegravir
RAM	Resistance associated mutation
RCA	Replication capacity assay
REF1	restriction factor 1
RegaDB	Rega database
Rev	Regulator of virion
rGag mutations	Resistance gag associated mutations
RLU	Relative light unit
RPV	Rilpivirine
RRE	Rev response element
RT	Reverse transcriptase
RTC	Reverse transcription complex

RT-PCR	Reverse transcription polymerase chain reaction
RTV	Ritonavir
RS	Reduced susceptibility
S	Susceptible
SA	South Africa
SDM	Site-directed mutagenesis
SDS	Sodium dodecyl sulphate
SGA	Single genome amplification
SIV	Simian immunodeficiency virus
SIV*cpz*	Simian immunodeficiency virus in chimpanzees
SIV*gor*	Simian immunodeficiency virus in gorillas
SIV*smm*	Simian immunodeficiency virus in sooty mangabeys
SOC	Super optimal broth with catabolite repression
SQV	Saquinavir
ssDNA	Strong stop deoxyribonucleic acid
ssRNA	Single stranded ribonucleic acid
T-20	Enfuvirtide
TAR	transactivator response element
TAM	Thymidine analogue mutation
Tat	Transactivator of transcription
TBS	Tris buffered saline
$TCID_{50}$	Tissue culture infective dose
TDR	Transmitted drug resistance
TDF	Tenofovir
TFR	Transframe region
TPV	Tipranavir
T_{reg}	Regulatory T-lymphocytes
Trim 5 α	Tripartite motif-containing protein 5 alpha
tRNA	Transfer ribonucleic acid
TNPO3	Transportin 3
TNF	Tumor necrosis factor
TSG101	Tumor susceptibility gene 101
UDPS	Ultra-deep pyrosequencing
UTR	Untranslated region
V1	Variable loop 1
V2	Variable loop 2
V3	Variable loop 3
VESPA	Viral Epidemiology Signature Pattern Analysis
Vif	Viral infectivity factor
Vpr	Viral protein R
Vpu	Viral protein U
VPS4	Vacuolar protein sorting 4
WHO	World health organisation
WT	Wild-type
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galctopyranoside

CHAPTER 1

Introduction & Literature Review

1. CHAPTER 1 INTRODUCTION AND LITERATURE REVIEW

1.1 Introduction

Human immunodeficiency virus (HIV) is classified as a retrovirus from the *Lentivirus* genus and the *Retroviridae* family of viruses (1). The human immunodeficiency virus infects immune cells expressing the cluster of differentiation (CD)4 protein, causing progressive deterioration in immune function that results in a syndrome of neoplastic diseases and opportunistic infections, referred to as acquired immune deficiency syndrome (AIDS) (2-5). HIV is transmitted in blood and body fluids with most infections acquired via: heterosexual or homosexual intercourse, mother-to-child transmission during pregnancy, birth or breastfeeding and injection of blood or blood-derived products (6).

The human immunodeficiency virus originated over three decades ago. More than 78 million people have since been infected globally and approximately 39 million people have died from AIDS (7). In 2014, 36.9 million people (range: 34.3 to 41.1 million) have been reported to be living with HIV, with roughly 2 million new infections and a total of 1.2 million AIDS related deaths in 2014 alone (8). Sub-Saharan Africa remains the most severely affected region globally, with almost one in every 20 adults estimated to be HIV positive, accounting for over 70% of the global burden of HIV (Figure 1.1) (7).

Within sub-Saharan Africa, South Africa (SA) has the highest burden of HIV (9), with an astounding 6.8 million (range: 6.5 to 7.5 million) people infected, equating to a prevalence of approximately 18.9% (range: 17.9 to 19.9%). An estimated 40% of these infected individuals reside in the province of Kwa-Zulu Natal (KZN), the site of this study (10). Despite having the world's largest antiretroviral therapy (ART) program, reaching almost 60% of all ART eligible patients in the country, the success and longevity of ART in SA may be compromised by the acquisition and transmission of drug resistance as a result of poor adherence and issues with access to treatment (11, 12).

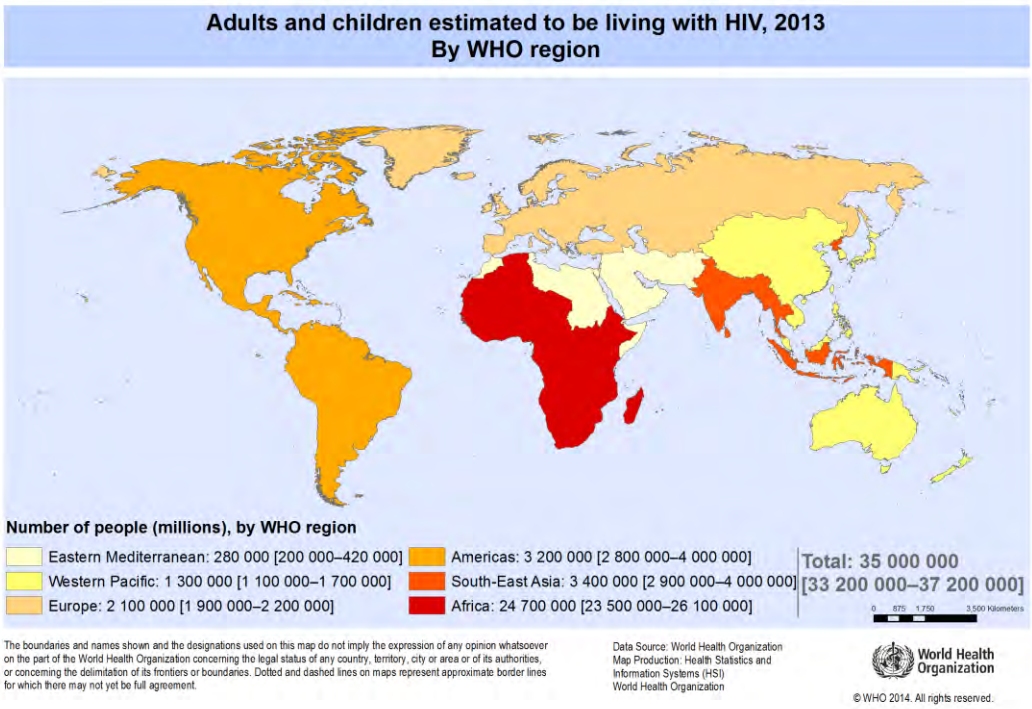


Figure 1-1 Global distribution of HIV. Sub-Saharan Africa accounts for the most infections globally (approximately 24.7 million HIV infections), this is followed by south-east Asia (3.4 million HIV infections) and the Americas (3.2 million HIV infections). Taken from World Health Organisation 2014 (13)

The present study investigated the role of Gag-Protease mutations in Protease inhibitor (PI) drug resistance, with particular interest in identifying resistance associated mutations (RAMs) in Gag and protease (PR) and determining their impact on viral replication, drug susceptibility and cleavage. Additionally, the prevalence of transmitted drug resistance (TDR) mutations in acutely infected individuals was investigated and low frequency (i.e. minority variant) mutations were identified.

In this chapter the history, origin and epidemiology of HIV are discussed. Additionally the structure of HIV-1, the protein constituents of the virus and their function and the viral replication cycle are detailed. Lastly the pathogenesis, treatment and drug resistance mechanisms of HIV-1 are also discussed.

1.2 The history of HIV

AIDS was first recognized in 1981 following an increase in incidence of Kaposi's Sarcoma (KS)[1], and *Pneumocystis carinii* pneumonia[2] (PCP), in previously healthy young homosexual men in Los Angeles and New York (2, 14, 15). The main symptoms of the illness amongst these men included severely depressed cell-mediated immunity accompanied by rare malignancies and opportunistic infections (15). The predominance of the illness in homosexuals drove the belief that the new entity of diseases only affected gay men. This however was disproved in 1982 when it was established that injecting drug users, recipients of blood transfusions and hemophiliacs were also affected. At this time, the Centre for Disease Control and prevention (CDC) named the group of disease entities AIDS (16).

The first mother-to-child transmission of AIDS and the first AIDS-related death of an infant (aged 20 months) in 1982 indicated that AIDS was not caused by lifestyle[3] issues, as originally believed, but rather was caused by an infectious agent transmissible in blood, blood products and body fluids (14, 17).

The causative agent of AIDS was first described in 1983 when Françoise Barré-Sinoussi and Luc Montagnier from the Pasteur Institute in France announced that they had isolated a new retrovirus, named lymphadenopathy associated virus (LAV), and suggested that it could be the cause of AIDS. However no proof of causality was established at the time (18). One year later, Robert Gallo's group from the national cancer institute announced that they had isolated the virus which causes AIDS and provided sufficient evidence to convince scientific and medical communities. This virus was named Human T-lymphotropic virus type III (HTLV-III) (19). Gallo's findings were confirmed three months later when Jay Levy's group isolated an AIDS-associated retrovirus from 22 patients infected with AIDS revealing the close association of the virus with AIDS (20).

By March 1985 whole genome sequencing of LAV, HTLV-III and the AIDS-associated retrovirus confirmed that they were variants of the same virus (21, 22), and the causative

[1] Karposi's sarcoma is a rare form of cancer which presents as purple lesions on the skin. It is usually rare in young individuals and results in the young looking older.
[2] PCP is a rare lung infection which is treatable with a 10 day course of pentamine. A healthy individual should easily fend off the infection.
[3] Lifestyle issues thought to cause AIDS encompass: immune overload from multiple infections, a reaction to semen, infection by an unknown fungus and multiple sex partners.

agent of AIDS was renamed HIV (23). In 1986 scientists isolated a second less virulent and prevalent retrovirus from West African patients presenting with clinical symptoms of AIDS, which was named HIV-type 2 (HIV-2). HIV was subsequently renamed to HIV-type 1 (HIV-1) (24).

1.3 Classification and origins of HIV

HIV is classified into two major types: HIV-1 and HIV-2. HIV-1 is morphologically similar to HIV-2; it is more virulent and accounts for almost 95% of HIV infections globally (25, 26). HIV-1 is subdivided into four groups *namely*; M (Major), N (non-M, non-O) O (outlier) and P (26). The group M-strain, responsible for the pandemic spread of HIV-1, is further subdivided into nine subtypes (A-D, F-H, J and K), six sub-subtypes (A1-A4 and F1-F2) and approximately 54 circulating recombinant forms (CRFs) (26-30). HIV-2 is subdivided into eight groups (A-H), with groups A and B being the most predominant (26). It has been estimated that the genetic diversity within a subtype is approximately 8–17% whilst the genetic diversity between subtypes is approximately 17–35% (31).

Both HIV-1 and HIV-2 originated from independent zoonotic transfer of Simian immunodeficiency virus (SIV), from primates to humans in West central Africa, where primates were butchered for bush meat (Figure 1.2) (26, 32). Simian immunodeficiency virus is a lentivirus which naturally infects old world monkeys (25). The cross-species transfer of SIV between: old world monkeys and chimpanzees, and chimpanzees and gorillas have been responsible for the varieties of SIV from which HIV-1 originated (Figure 1.2) (25).

For HIV-1, strains M and N are related to SIV from the subspecies of chimpanzee (SIV*cpz*) *Pan troglodytes troglodytes* (25, 33, 34) whilst strain P and O, is related to SIV transfer from western gorillas (SIV*gor*) (26, 28, 32). HIV-2 is related to SIV from the sooty mangabey (SIV*smm*) subspecies *Cerocebus atys atys* (33, 35) (Figure 1.2).

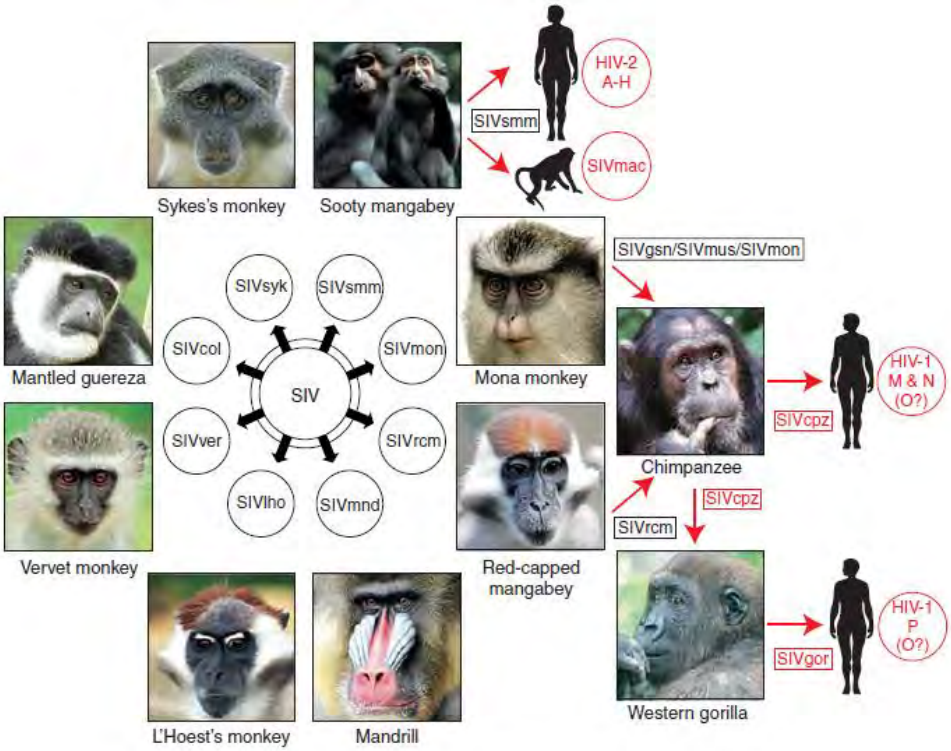


Figure 1-2 Origins of HIV-1 showing SIV in several old world monkeys which crossed the species barrier to infect chimpanzees, gorillas and humans. (Taken from Sharp et al., 2011) (25).

Using the earliest available HIV-1 positive sequences (1959 – 1960) and statistical models to evaluate the rate of evolution, the estimated date of origin for HIV-1 group M, N and O were estimated to be 1908 (range: 1884 – 1924), 1963 (range: 1948 – 1977) and 1920 (range: 1890 – 1940) respectively (25, 26, 31, 36). The estimated date of origin for HIV-2 group A and B was 1932 (range: 1906 - 1955) and 1935 (range: 1907 - 1961) respectively (25, 32).

Faria et al., (2014) showed that group M, which represents the oldest lineage of HIV-1, originated in Kinshasa and spread through central Africa, via the extensive use of railway networks by infected individuals at the time (37).

Several studies suggest that the growth of cities in Africa, urbanization, migration and the improved ease of travel by the mid twentieth century contributed to the exponential growth of the HIV epidemic and the uneven global distribution of subtypes and CRFs (26, 32, 37) (Figure 1.3). For the purposes of this dissertation, the following sections will focus on aspects of HIV-1 infection.

1.4 Global distribution of HIV-1

HIV-1 subtype B predominates in Europe, Australia and the Americas, CRF-01 (a combination of HIV-1 subtype A and E) predominates in most of Asia (Figure 1.3). Central Africa, the origin of HIV, has the greatest diversity of HIV subtypes and CRFs. In Africa the most prevalent subtypes of HIV-1 include, subtype C, A and D. HIV-1 subtype C predominates in southern Africa and India and is increasing in frequency in China and eastern Africa (Figure 1.3) (9, 26). HIV-1 subtype C, the focus of the present study, is responsible for over 48% of global HIV infections (9, 26, 38, 39). Its rapid spread has been attributed to the reduced replicative fitness associated with HIV-1 subtype C compared to other HIV-1 group M subtypes (38).

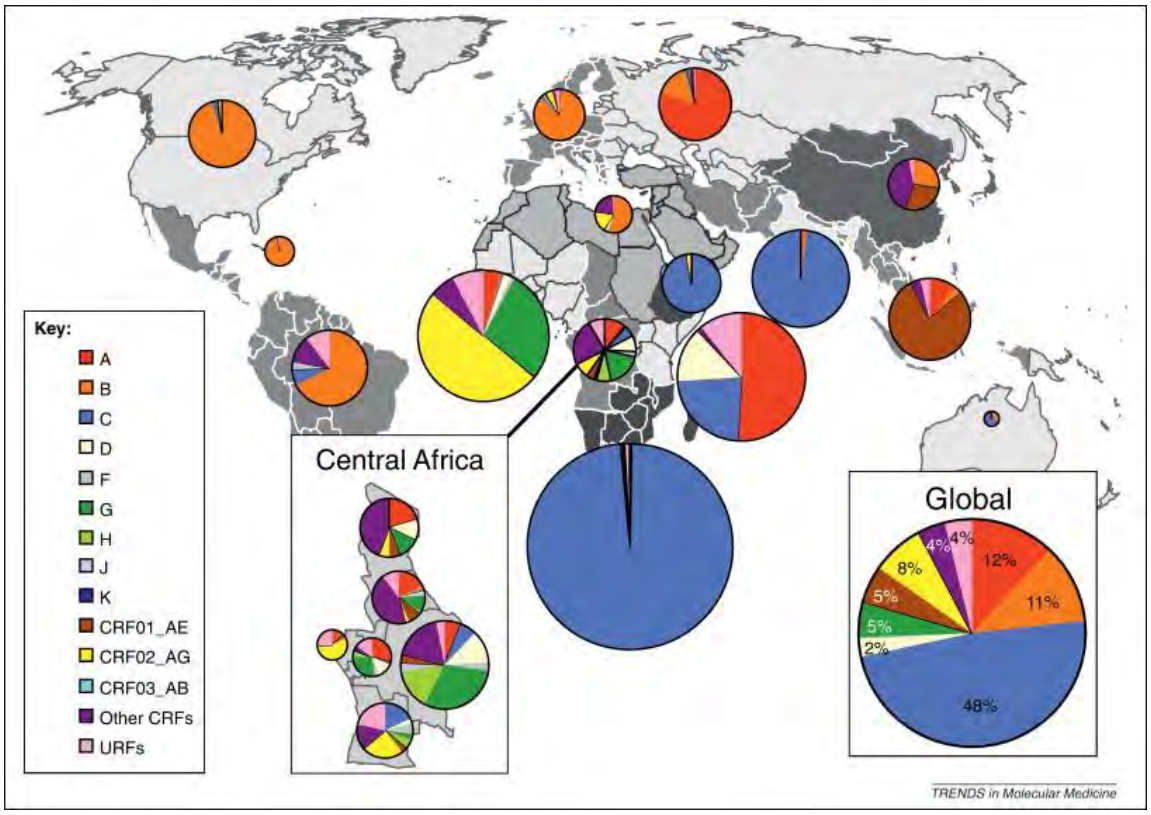


Figure 1-3 Global distribution of HIV-1. The surface area of the pie chart is representative of the number of people living with HIV-1 in a particular region. HIV-1 subtype C (blue) is the most predominant subtype and is responsible for over 48% of all HIV infections. Taken from Hemelaar et al., 2012 (39, 40). CRF – circulating recombinant form; URF – unique recombinant form.

1.5 Structure of HIV-1

The virions of HIV-1 are spherical in morphology and are between 100-120 nm in diameter (41). The following section describes the: genomic organization of HIV-1, HIV-1 proteins and their function and the structure of HIV-1.

1.5.1 The genomic organization of HIV-1

The HIV-1 genome consists of approximately 9,700 base pairs (bp) divided into nine overlapping genes which are flanked on both the 5' and 3' ends by identical long terminal repeats (LTR, 634 bp) (42) (41). The 5' LTR contains the HIV-1 promoter sequence located in the U3 region which initiates viral transcription whilst the 3' LTR contains the polyadenylation signal required for dimerization and genome packaging (Figure 1.4a) (43, 44).

The nine genes of HIV-1 encode for three polyproteins and six smaller accessory/regulatory proteins (Figure 1.4a):

- group specific antigen (g*ag*, 1,503bp) which encodes for structural components of the virion,
- polymerase (*pol*, 3,012 bp), responsible for encoding the PR, reverse transcriptase (RT) and integrase (INT) enzymes which are encapsulated within the virion,
- envelope (*Env*, 2,571bp), which encodes for viral envelope glycoproteins (gp) expressed in the outer membrane envelope of the virion,
- transactivator of transcription factor (*Tat*, 306 bp) and regulator of virion protein (*Rev*, 351 bp) which code for regulatory proteins and
- viral infectivity factor (*Vif*, 579 bp), viral protein R (*Vpr*, 292 bp), viral protein U (*Vpu*, 249 bp) and the negative regulation factor (*Nef*, 621 bp) which encode for accessory proteins.

1.5.2 HIV-1 proteins and their function

Gag, Pol and Env, the three proteins shared by all retroviruses, are initially synthesized as polyprotein precursors. These polyprotein precursors require enzymatic cleavage in order to produce mature proteins (Figure 1.4a) (42).

1.5.2.1 Group specific antigen (Gag)

The Gag polyprotein precursor named precursor of 55 kDa ($Pr55^{Gag}$) is cleaved by viral Protease during maturation into p17 matrix (132 amino acids [aa]), p24 capsid (231 aa), p7 nucleocapsid (55 aa), protein of 6 kDA (52 aa, p6) protein of 1 kDA (14 aa, p1) and protein of 2 kDA (16 aa, p2) (Figure 1.4) (42, 44-47). The p17 protein is responsible for: directing the Gag and Gag-Pol polyprotein precursors to the plasma membrane for virion assembly and for the incorporation of envelope glycoproteins into developing virions (48). The p24 protein forms a cone-shaped structure which protects and encapsulates the viral genetic material and facilitates its delivery into the nucleus of the host cell (49) whilst p7 encapsulates unspliced genetic material and has a nucleic acid chaperoning function in which it facilitates the structural rearrangement of genomic material during replication by RT (50). The p6 protein mediates the virus ESCRT (endosomal sorting complex required for transport) dependent budding via two late assembly domains (L-domains) and enables the integration of Vpr into virions (51).

1.5.2.2 Polymerase (Pol)

The Pol enzymes namely PR (99 aa), RT (560 aa) and INT (288 aa) are cleaved from the 160 kDa Gag-pol polyprotein precursor by viral PR. The homodimeric PR is responsible for initiating viral maturation (52, 53). The heterodimer RT converts viral single stranded ribonucleic acid (ssRNA) into double stranded deoxyribose nucleic acid (dsDNA) after viral entry into a cell and has RNAse H activity which facilitates specific degradation of viral RNA from DNA-RNA duplexes (54). Integrase facilitates the incorporation of viral DNA into the host chromosomal DNA (55).

1.5.2.3 Envelope (Env)

The Env polyprotein precursor is cleaved by furin-like host cellular Protease into transmembrane glycoprotein 41 (gp41; 345 aa) and surface glycoprotein 120 (gp120; 511 aa) subunits (56). The gp41 traverses the lipid bilayer and is non-covalently bound to gp120. The gp120 mediates virus attachment to host cells whilst gp41 facilitates entry of the virus into the host cell via fusion of host and viral cellular membranes (Figure 1.4) (57)

1.5.2.4 Transactivator of transcription factor (Tat)

Transactivator of transcription factor is a RNA binding protein which recognizes and binds to a transactivator response element (TAR) sequence from the HIV-1 RNA molecule and activates transcription from the viral LTR promoter (58-60).

1.5.2.5 Regulator of virion protein (Rev)

Regulator of virion protein is a regulatory protein which binds to the Rev response element (RRE), a viral RNA element present on individual unspliced or partially spliced viral RNA molecules and initiates the transport of viral RNA transcripts (both spliced and unspliced) out of the nucleus and into the cytoplasm where they serve as transcripts for translation (61, 62).

1.5.2.6 Viral infectivity factor (Vif)

Viral infectivity factor is an accessory protein with a key role in increasing pathogenicity of HIV-1 virions. Its principal target is Apolipoprotein B mRNA-editing enzyme-catalyticpolypeptide-like 3G (APOBEC3G), a member of the APOBEC family of deoxycytidine deaminases which function to suppress viral replication by inducing G to A hyper-mutations within newly synthesized viral DNA thereby inactivating the virus. Viral infectivity factor binds to APOBEC3G and coordinates its proteosomal degradation thereby inhibiting the packaging of APOBEC3G into budding virions (63, 64).

1.5.2.7 Viral protein R (Vpr)

Viral protein R is a second accessory protein which functions primarily to improve viral pathogenicity via: enhancing LTR transcription within infected cells, orchestrating the import of the reverse transcription complex (RTC)[4] into the host cell nucleus and inducing cell cycle arrest. It is also involved in inducing apoptosis of T-lymphocytes (65-67).

[4] Reverse transcription complex is a term used to describe the infectious viral unit, within which viral ribonucleoprotein enters the host cell and begins reverse transcription of the viral RNA genome (98).

1.5.2.8 Viral protein U (Vpu) and Negative regulation factor (Nef)

Viral protein U and Nef are two additional accessory proteins which also contribute towards improving viral pathogenicity. Viral protein U serves to: promote the release of viral progeny from infected cells, downregulate $CD4^+$ cells via the ubiquitin proteasome pathway and downregulate CD155 and natural killer cell receptors in order to evade natural killer cell (NKC) mediated immune responses (68, 69). The main function of the Nef protein is to reduce CD4 and major histocompatibility complex I and II (MHCI and MHCII) cell surface receptors thus facilitating immune evasion (70). Additionally, Nef has a role in the inhibition of apoptosis of infected cells thereby maintaining the longevity of infected cells and contributing to viral propagation and survival (70, 71).

1.5.3 Structure of HIV-1

Each HIV-1 viral particle is surrounded by a lipoprotein rich membrane (i.e. lipid bilayer) with heterodimer complexes comprised of trimers of surface gp120 (which protrudes from the lipid bilayer to the external region of the virion) and transmembrane gp41(which spans the interior of the lipid bilayer) bound together (41, 72). In general, virions have between 14-74 trimers (73). The matrix (MA) protein is attached to the inner surface of the viral lipoprotein membrane. The inner core of the virus is enclosed by the capsid (CA) protein arranged in a fullerene conical structure (49, 74, 75). This structure has approximately 250 CA hexameric rings arranged in a lattice and 12 pentamer rings (5 at the top end and 7 at the bottom end) which provide its conical fullerene structure (76, 77). It encases the: viral genetic material (i.e. a positive sense ssRNA) which is in contact with the nucleocapsid (NC), viral proteins (i.e. p6, Vpr, Vif and Nef), viral enzymes (i.e. PR, RT and INT) and cellular proteins (i.e. APOBEC3G and cyclophilin A [cypA]) (Figure 1.4b) (78-80).

1.6 Replication cycle of HIV-1

The replication cycle of HIV-1 takes approximately 24 hours and comprises of several sequential steps (depicted in figure 1.4c) including: (1) virus entry (binding to respective receptors and subsequent fusion with the host cell membrane); (2) reverse transcription of the single stranded viral RNA genome to dsDNA; (3), uncoating of the viral capsid to release viral genetic material (i.e. RNA) and proteins; (4) nuclear entry of the RTC; (5) integration of viral DNA into host DNA; (6) transcription and nuclear export of new viral

RNA; (7) translation of proteins and its subsequent translocation to the cell surface; (8) viral assembly; (9) viral release/budding and (10) maturation (Figure 1.4c) (81).

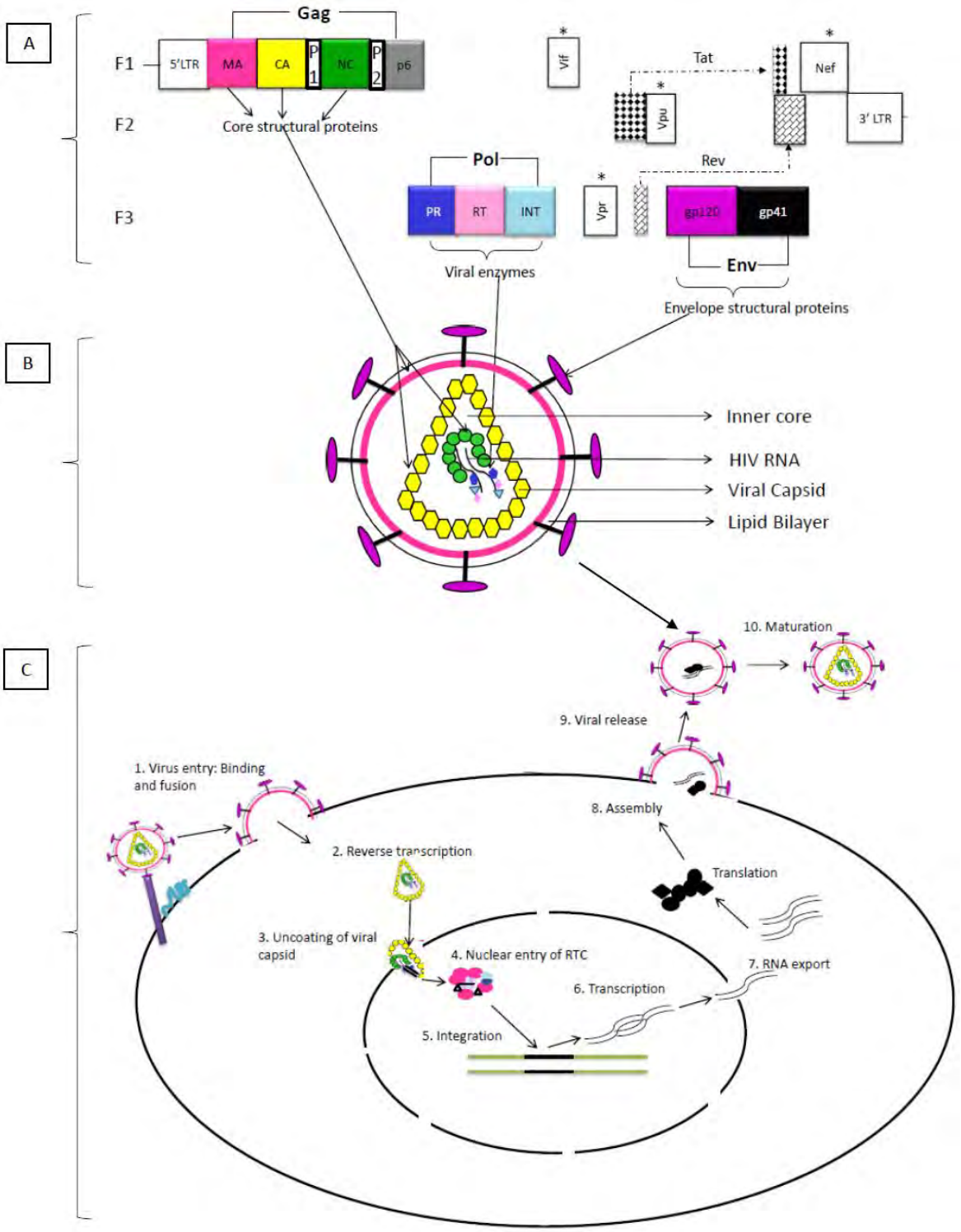


Figure 1-4 Genomic organization, virion structure and replication cycle of HIV-1. (a) Nine genes of HIV-1, flanked on either end by an LTR, within their respective reading frames. Regions

which code for the major proteins (i.e. gag, pol and env) are depicted in color. Regions which code for accessory proteins (i.e. vif, vpu, vpr and nef) are presented in white boxes and are denoted by an asterisk. Regions which code for regulatory proteins (i.e.tat and rev) are represented by patterned boxes. Adapted from Fanales-Belasio et al., 2010 (41) and Los Alamos National Laboratory database 2015 (82). (b) Structure of a mature infectious virion. Structural and enzymatic components are color coded to match the corresponding gene as depicted in Figure 1.4a above. Adapted from Robinson et al., 2002 (84) and Fanales-Belasio et al., 2010 (41, 83). (c) The ten steps of the HIV-1 replication cycle. Adapted from Fanales-Belasio et al., 2010 (42) and Engelman et al., 2012 (41, 84). Abbreviations: HIV-1 – human immunodeficiency virus 1; F1 – Frame 1; F2 – Frame 2; F3 – Frame 3; Gag – group specific antigen; LTR – Long terminal repeat; MA – Matrix; CA – Capsid, p2 – protein of 2 kDA, NC – nucleocapsid; P1 – protein of 1 kDA; P6 – protein of 6 kDA; pol – polymerase; PR – Protease, RT – reverse transcriptase; INT – integrase; Vif – viral infectivity factor; Vpr – viral protein R; Tat – transactivator of transcription factor; Rev – regulator of virion protein; Vpu – viral protein U; Env – envelope; gp120 – glycoprotein of 120 kDA, gp41 – glycoprotein of 41 kDA; Nef – negative regulation factor; RNA – ribonucleic acid; RTC – reverse transcription complex; NRTI – nucleoside reverse transcriptase inhibitor; NNRTI – non-nucleoside reverse transcriptase inhibitor; InSTI – integrase strand transfer inhibitor.

1.6.1 Virus entry

The first step in viral entry is the adsorption of the virion to the host cell. This is mediated by cell attachment factors[5] and either viral envelope proteins or proteins from the host cell membrane incorporated into the virion envelope (72, 85-89). This interaction brings the virion into close proximity to the host receptors and is associated with improving the efficiency of infection. However, attachment factors are not essential for virus entry and are not always employed (72, 85).

The second step in viral entry involves the binding of HIV-1 gp120 to $CD4^+$ receptors (expressed on the surface of T-helper cells, T-regulatory cells, macrophages, dendritic cells, Langerhans cells, microglial cells and monocytes) (75, 85). This binding triggers several structural changes in gp120 required for co-receptor binding. Firstly, the variable loop 1 and 2 (V1/V2) region of gp120, located at its surface, moves from a central axis of symmetry toward the lateral aspect of the gp120 trimer subsequently moving the variable loop 3 (V3) stem toward the distal region of the trimer (72). This exposes the V3 region and brings it into direct contact with the host cell membrane. Secondly, the bridging sheet (a second co-receptor binding site) is formed via assembly of four antiparallel beta sheets (2 from the outer and 2 from the inner domains of gp120) (72). Thirdly, rearrangements within gp120 result in an outward rotation of the gp120 monomer which partially exposes the stalk of gp41. Collectively these structural rearrangements make the two co-receptor binding sites (i.e. the V3 region and the bridging sheet) more accessible to co-receptors.

Co-receptor binding forms the third step of viral entry and serves as a trigger for activation of virion and host membrane fusion. Two co-receptors are predominantly used in this step: C-C chemokine receptor 5 (CCR5)[6] or CXC chemokine receptor 4 (CXCR4)[7] (90). These co-receptors can be used either individually or in combinations. Viruses which use CCR5 only are termed R5, those which use CXCR4 only are termed X4 and those which use both are termed R5X4 (91). During co-receptor binding, the base of the V3 loop and bridging sheet is engaged by the N-terminus of the co-receptor. The extracellular loop 2 (ECL2) of the co-receptor binds to the tip of the V3 loop resulting in the exposure of a

[5] Cell attachment factors include: heparin sulfate proteoglycans, α4β7 integrin or dendritic cell–specific intercellular adhesion molecular 3-grabbing non-integrin (DC-SIGN)

[6] CCR5 receptors are predominantly found on the surface of: memory $CD4^+$ T lymphocytes, macrophages and microglial cells.

[7] CXCR4 receptors are predominantly found on naïve $CD4^+$ T lymphocytes.

hydrophobic fusion peptide, located in the amino terminal of gp41, which subsequently inserts into the host cell membrane triggering the last step of viral entry, membrane fusion (72, 85, 92-94).

During membrane fusion, the amino terminal helical regions (HR-N) of each gp41 subunit arranges to form a triple stranded coiled structure. The three carboxy terminal helical regions of gp41 folds and packs in an antiparallel fashion into grooves at the interface of the three HR-N domains of gp41 resulting in the development of a six helix bundle (95, 96). The six helix bundle brings the viral and host cell lipid bilayers into close proximity resulting in the development of a fusion pore via which the viral capsid is injected into the host cell cytoplasm (Figure 1.4c) (72, 94-96).

1.6.2 Reverse transcription

Following entry into the cytoplasm, the viral capsid remains intact and migrates towards the nuclear membrane using the microtubule network of the host cell (97). Reverse transcription[8] occurs within the intact capsid, as part of a RTC (Figure 1.4c). The viral capsid has an important role in protecting the viral genome from host factors, immune identification and in maintaining a high stoichiometry of the RT enzyme thereby preventing its dissociation from its template (74, 75, 98, 99).

Reverse transcription is initiated when the 3' end of host transfer (t)RNA^{lys3} anneals to a primer binding site (pbs)[9] located at the 5' end of the single stranded viral RNA, facilitated by the NC protein (98, 100). The pbs is complementary to 18 nucleotides of the 3' terminal of host tRNA^{lys3}. Reverse transcriptase recognizes the tRNA-RNA initiation complex and initiates extension of the 3' end of the primer using the RNA template to guide the synthesis of a minus-strand strong stop DNA segment ([-] ssDNA). The RNA from the newly synthesized RNA-DNA duplex is concomitantly degraded by RNase H (cleaved from RT) during DNA synthesis (98, 101).

[8] Reverse transcription is defined as the process by which viral single stranded RNA is converted into linear double stranded DNA.
[9] The primer binding site comprises of approximately 180 nucleotides that are complementary to the tRNAlys3.

The resulting (-) ssDNA segment is released from the degraded RNA-DNA duplex and acts as a primer which binds to a complementary sequence at the 3' end of the viral RNA. Hybridization of the (-) ssDNA segment and the 3' end of the viral RNA, facilitated by complementary direct repeat (R) sequences on either segment, allows for DNA synthesis to resume (this is known as first strand transfer). Reverse transcriptase now facilitates the synthesis of a full length (-) ssDNA segment which serves as a template for (+) strand DNA synthesis. RNase H again degrades RNA from the RNA-DNA duplex, however two short purine rich sequences referred to as the polypurine tract (PPT), located at the 3' end of the viral RNA and in the centre of the viral RNA (named 3'PPT and cPPT respectively), are resistant to RNase H and remain un-cleaved (88, 98, 101, 102).

These un-cleaved segments serve as primers for plus (+) strand DNA synthesis. The complementary sequence of the PPT region adjacent to the 3' end of the viral genome anneals to the newly synthesized nascent (-) ssDNA segment (i.e. template), this triggers RT action and synthesis of a (+) strand DNA segment. As part of (+) strand DNA synthesis the U3, R and U5 sequences of viral RNA and the first 18 nucleotides of the $tRNA^{lys3}$ are also copied. Thereafter, the tRNA is partially degraded by RNase H. This exposes the pbs at the 3' end of the (+) strand DNA which binds to the complementary pbs sequence at the 3' end of the (-) strand DNA leading to the second strand transfer (98, 101, 102).

Reverse transcriptase facilitates the elongation of the (+) and (-) strands of DNA resulting in the synthesis of completed double stranded linear viral DNA. Termination of the (+) strand DNA synthesis occurs once the central termination signal (CTS) (i.e. a sequence at the end of the (-) strand DNA) is encountered (98, 101).

The final steps of DNA syntheses is marked by the displacement of several nucleotides of (+) strand DNA, driven by the cPPT upstream of the CTS, to form the central DNA flap. The function of which remains unclear, however some studies claim that it may trigger uncoating of the viral capsid (103, 104) or have a role in translocation of viral DNA into the host cell's nucleus (105).

1.6.3 Uncoating

Uncoating is defined as the process by which the viral CA is disassembled to reduce its size in order to facilitate nuclear import of viral genetic material (97). Historically it was

believed that uncoating occurred in the cytoplasm and that the RTC was imported into the nucleus in isolation of the CA (106, 107). This however is a controversial topic with recent models suggesting various locations for uncoating including: (1) rapid disassembly of the core in the cytoplasm (2) partial uncoating of the core in the cytoplasm followed by further uncoating in the nucleus (3) partial uncoating during reverse transcription immediately outside the nuclear pore complex (NPC)[10] and final uncoating within the nucleus (108-111). The latter two models being most favored (97, 112).

Uncoating of the CA is mediated by a combination of: viral, cellular and host factors, and interactions/events within the capsid core. Several factors could account for viral uncoating, some of which are discussed below.

Polymerized nascent viral DNA is thought to exert pressure on the capsid core since the flexible single stranded viral RNA is converted into rigid dsDNA, resulting in core remodelling to accommodate the genetic material. Additionally, RT and INT have been shown to provide stability to the CA core. It has been suggested that the onset of reverse transcription may result in the dissociation of RT from its stabilizing reactions with INT thereby resulting in some core disassembly (97).

Cylophilin A[11] is also associated with uncoating. It has been shown to interact with a conserved proline rich loop on the CA where it catalyses the cis-trans isomerization of the Gly89-Pro90 CA peptide bond. This results in conformational changes within the amino terminal domain terminal of CA, which could play a role in destabilization/disassembly of the CA (97, 113, 114). In addition to its role in uncoating, cypA is also associated with protecting HIV-1 from host restriction factors such as Tripartite motif-containing protein 5 alpha (TRIM5α) and restriction factor 1 (REF1), which bind directly to CA and promotes premature uncoating resulting in abortion of viral replication. It is also thought to contribute towards the ability of HIV-1 to infect non-dividing cells (115).

Cytoplasmic trafficking using dynein and kinesin-1 have also been identified to play a role in viral uncoating. It has been suggested that dynein facilitates the transport of the CA to the NPC where kinesin-1 mediates the CA uncoating (116, 117).

[10] The nuclear pore complex is a multiprotein channel located on the nuclear envelop. It allows for selective trafficking of macromolecules between the cytoplasm and nucleus (81).
[11] Cylophin A is a host peptidyl prolyl isomerase which plays in role in viral uncoating and has been shown to protect HIV-1 from host restriction factors (98).

Lastly, nucleoporin 358 (NUP358), a NPC channel protein with a cypA homology-domain (118, 119) has been shown to interact with the viral CA triggering CA isomerization which suggest that NUP358 has a role in CA uncoating (115, 120).

The process of uncoating, whilst unclear, remains an obligatory step in viral replication and is integral for the import of viral genetic material into the host nucleus.

1.6.4 Nuclear import

As part of nuclear import constituents of the RTC (i.e. viral DNA, Gag MA, Gag NC, Gag CA, RT, INT, Vpr and host cellular proteins) are translocated from the cytoplasm into the nucleus (97, 103, 112).

Nuclear import is thought to be mediated by interactions between CA and several host cellular factors including (121-123): cleavage and polyadenylation specificity factor 6 (CPSF6), transportin 3 (TNPO3), NUP358 (also referred to as RanBP2) and nucleoporin 153 (NUP 153) (124-126) as described below.

Cleavage and polyadenylation specificity factor 6, a host messenger (m)RNA processing protein, shuttles between the cytoplasm and nucleus and has been shown to attach to binding pockets on the assembled CA to facilitate nuclear import of the RTC (97, 109, 127, 128).

Transportin 3, a member of the importin β family of proteins responsible for nuclear localization of serine-arginine rich proteins (129), has either a direct or indirect role in nuclear import. For the direct route and similarly to CPSF6, TNPO3 binds to the capsid binding pocket and facilitates nuclear import (97). As part of the indirect route, it has been suggested that TNPO3 mediates the nucleoplasmic localization of proteins required for nuclear import of the RTC, such as CPSF6 (97, 129). Additionally, it has been suggested that TNPO3 may also play a role in viral genome integration whereby it removes reminant CA from the RTC for integration (111).

Nucleoporin 358 and NUP153 are two cellular factors associated with HIV-1 infection and nuclear import of the RTC. Both cellular factors are located within the NPC with NUP358 projecting towards the cytoplasm and NUP153 projecting toward the nucleus (118, 119).

Whilst the mechanism of nuclear import associated with NUP358 remains unclear, research suggests that it interacts with the CA via a CA-cypA domain, on the cytoplasmic side of the NPC, to facilitate uncoating and stimulate perinuclear localization of the CA (97, 112). The nucleoplasmic NUP153 protein appears to be involved in a finalization step which renders the RTC competent for integration, whilst this is an essential step for viral integration the exact mechanisms remains elusive (97, 130, 131).

Even though the complete mechanism of nuclear import remains unclear, it is definite that a portion of the intact CA is included in the RTC where it plays an essential role in binding host factors, most likely in a sequential manner, to mediate nuclear import. The large array of host factors may either bind directly to CA or may function as signals to attract other host factors to a specific region.

1.6.5 Integration

Integration is the process by which viral DNA is irreversibly integrated into host DNA via steps catalyzed by retroviral INT. Retroviral INT catalyses two reactions: 3' processing of the ends of viral DNA and a strand transfer reaction during which processed ends of viral DNA are integrated into host DNA (84, 132).

In the first reaction, INT engages the LTR ends of viral DNA, shortly after its synthesis, and processes each end to yield a 3' hydroxyl group. This point marks the transition from a RTC to a pre-integration complex (PIC)[12] (133). Integrase then cleaves host chromosomal DNA, using the 3' hydroxyls, in a staggered fashion and subsequently joins the 3' ends of the viral DNA to the 5' phosphates of the host target DNA (84, 132, 134). Thereafter host cell enzymes fill in the gaps between host and viral DNA and subsequently ligate the 5' ends of un-joined viral DNA to host DNA yielding an integrated provirus (84, 132).

1.6.6 Transcription and nuclear export

Following integration into the host genome, the HIV-1 provirus acts as a template for transcription.

[12] Pre-integration complex (PIC) describes a unit comprising of viral DNA and host and viral proteins which can integrate into host DNA. For this to occur, reverse transcribed DNA from the RTC undergoes integrase mediated endonuclease priming of both the 3' and 5' ends (98).

Transcription of HIV-1 is mediated by an individual promoter located in the 5' LTR of the provirus. The LTR consists of 3 sub-regions (U3, R and U5) involved in transcription. The U3 region comprises of several cis-acting DNA elements which function as binding sites for cellular transcription factors (e.g. NF-kappa B) (135). The first nucleotide at the R region marks the point at which transcription begins, whilst the last nucleotide in the R region marks the point at which polyadenylation occurs. The U5 region comprises of the TAR region to which Tat binds to (136-138).

The NF-kappa B family of cellular transcription factors, which enter the nucleus when the cell is activated, binds to the U3 region of the LTR and initiates transcription (135). Several short transcripts are formed in conjunction with some complete transcripts which enable the generation of Tat protein (139). The Tat protein then binds to the TAR element in the U5 region and significantly increases the transcription of viral RNA by recruiting positive transcription elongation factor b (PTEF-b) to the TAR element in viral transcripts (138). This induces phosphorylation of residues within RNA polymerase II stimulating elongation and increased transcription (84, 137, 138, 140-142).

1.6.7 Assembly

Virion assembly occurs at nucleation sites[13] once Gag is synthesized and translocated to the inner leaflet of the host plasma membrane (i.e. the site of viral assembly), either as: a 55 kDa Gag polyprotein, a 160 kDa Gag-Pol polyprotein or as a Gag-viral RNA duplex (143).

Trafficking of Gag to the host plasma membrane is mediated by direct electrostatic interactions between residues 15-31 of Gag MA and phosphoinositide phosphatidylinositol-4,5-bisphosphate (PtdIns(4,5)P2), a phospholipid present in large quantities within the inner leaflet of the plasma membrane (143, 144). The PtdIns(4,5)P2 displaces viral RNA bound to some Gag polyproteins resulting in the exposure of the N-terminal of MA myristate which subsequently anchors to the host plasma membrane (143, 145).

Molecules of Gag are aligned and oriented in a radial manner with the MA region bound to the inner leaflet of the plasma membrane and the carboxy region of Gag packed towards

[13] Nucelation sites refer to plasma membrane lipid rafts which are enriched in cholesterol, sphingomyelin and plasmalogen-PE and have a higher saturated fatty acid concentration in comparison to other regions of the plasma membrane.

the centre of the virus particle (143). The resultant immature viral particle comprises of a continuous Gag lattice with a few gaps that are void of Gag (143, 146-148).

The incorporation of two strands of viral RNA into the nascent viral particle is mediated by the Gag NC (143). The exact mechanism remains elusive, however cross linking immunoprecipitation sequencing showed that Gag interacts with RNA in the cytosol and binds to several distinct sites at the 5' untranslated region (UTR) and the RRE of viral RNA (143, 149, 150). This directs RNA dimers, via the endosomal pathway, to the host plasma membrane where the zinc finger like domains of the Gag NC domain (particularly at residues 390 to 423) binds to a packaging signal near the 5'UTR of viral RNA (143, 151, 152). This enables the NC to assemble around the RNA dimer resulting in immobilization of a usually dynamically moving RNA dimer (150). The binding of viral RNA to the zinc finger region triggers Gag multimerization which concentrates Gag monomers to the plasma membrane and provides the RNA framework for assembly (153). Interactions between Gag CA domains results in the formation of CA hexamers that join to form a spherical structure which encases the viral RNA and viral and cellular proteins (154).

The arrival of Gag at the plasma membrane also triggers the recruitment and coalescence of lipid rafts, which are possibly involved in the incorporation of Env glycoproteins into the host plasma membrane (143). The Env glycoprotein is transported towards the host plasma membrane via the secretory pathway (155). During this process the host cellular enzyme furin cleaves the Env glycoprotein into gp120 and gp41 subunits. Complexes of gp120-gp41 segregate toward lipid rafts at the host cell plasma membrane (156). Here an unexplained interaction between the cytoplasmic tail of gp41 and MA (particularly residues 6 to 17) occurs to facilitate the incorporation of the envelope proteins into the immature virus particle (157). Even though the exact mechanism of Env incorporation into virions is unclear, it is well established that several MA residues are essential for this process. These include residues: 8, 9, 13, 16, 17, 18, 31 and 35 (158-160).

Interactions between the Gag p6 carboxy terminal and viral Vpr, Vif and Nef mediates the incorporation of these viral proteins into the nascent viral particle (161). Several host proteins (such as cypA) are also incorporated into the nascent virion, however the exact mechanism of recruitment remains elusive.

1.6.8 Virion release

Membrane scission, defined as the separation of the host plasma membrane and virion membrane, of HIV-1 is mediated by interactions between the host cellular ESCRT pathway and two late domain motifs located in Gag p6 (i.e. PTAP and YPXL) (143, 162).

These late domain motifs bind and recruit factors that act early in the ESCRT pathway. The PTAP domain binds tumor susceptibility gene 101 (TSG101), a host cellular factor which is part of the ESCRT-1 complex whilst the YPXL domain binds ALG2-interacting protein (ALIX), an ESCRT-III factor (143, 162, 163). This binding recruits ESCRT-III proteins to the nascent viral particle (143, 162), where they assemble into circular spirals within the head of the budding virion. This is thought to assist in constriction of the membrane, at the neck of the bud, which drives membrane scission (164, 165). The ESCRT-III proteins also recruits host AAA ATPase vacuolar protein sorting 4 (VPS4) to complete fission via the hydrolysis of ATP (143).

1.6.9 Maturation

Virion maturation is driven by viral PR, an aspartyl PR comprising of two identical subunits with its active site situated in a cleft at the dimer interface (166-170). The PR active site comprises of two aspartic acid residues (i.e. one from each subunit) which are covered by two identical flexible flaps that open and close to regulate entry and exit of substrates and products (Figure 1.5a) (169-174). The substrate binding pocket is made up of residues 25-32, 47-53, 76 and 80-84 (175).

The mechanism by which PR recognizes its substrates is unclear with several studies suggesting that substrate recognition, in HIV-1, is based on shape/conformational structure of the substrate rather than identification of aa sequences (176-178). This is largely driven by variability of aa sequences in the Gag and Gag-pol polyprotein cleavage sites (CS's), which will make these sites impossible to cleave if they were recognized solely by aa sequences (174, 177). Three to four aa residues on either side of the substrate cleavable peptide bond initiate substrate binding to the PR cleft. Following substrate binding, PR utilizes two aspartic acid side chains within an Asp-Thr-Gly motif for activation of a water molecule which catalyses peptide bond hydrolysis (169, 173, 174).

For maturation of HIV-1 virions, PR performs a total of 12 cleavages resulting in the production of structural proteins and viral enzymes which transforms the nascent viral particle into an infectious virion (169).

Partially active PR in the nascent viral particle initiates autocatalysis of the Gag-pol p2↓p7 junction, transframe region (TFR) and $p6^{pol}$ junction (TFR↓$p6^{pol}$) and the $p6^{pol}$↓PR junction sequentially (179, 180). This releases fully functional PR which is used in subsequent cleavage steps. HIV-1 pol is cleaved at three sites to yield RT, INT and PR enzymes, whilst it is known that PR is cleaved first the order of cleavage for INT and RT is unclear (Figure 1.5c) (169).

The Gag polyprotein is cleaved at 5 sites to yield structural proteins (including: MA, CA and NC) in a highly conserved order (Figure 1.5b) (181). There are a total of three stages of Gag polyprotein cleavage. Primary processing occurs at the p2↓p7 site and yields p43 (an intermediate comprising of p17-p24-p2) and p14 (intermediate for p7-p1-p6). Secondary processing of the p1↓p6 site and the p17p24 site occurs respectively and produces p17, p6, p8 (intermediate of p7 and p1) and p25 (intermediate of p24 and p2). Tertiary processing of p7↓p1 and p24↓p2 yields p1, p7, p24 and p2 (182).

Cleavage of the Gag and Gag-pol polyproteins results in stabilization and condensation of the viral genome by the cleaved NC (154). Additionally, morphological changes within the virion occur with the MA domain of Gag remaining attached to the plasma membrane, whilst the CA domain rearranges to form a conical fullerene structure (143, 154). The process of maturation is essential for formation of infectious viruses (183).

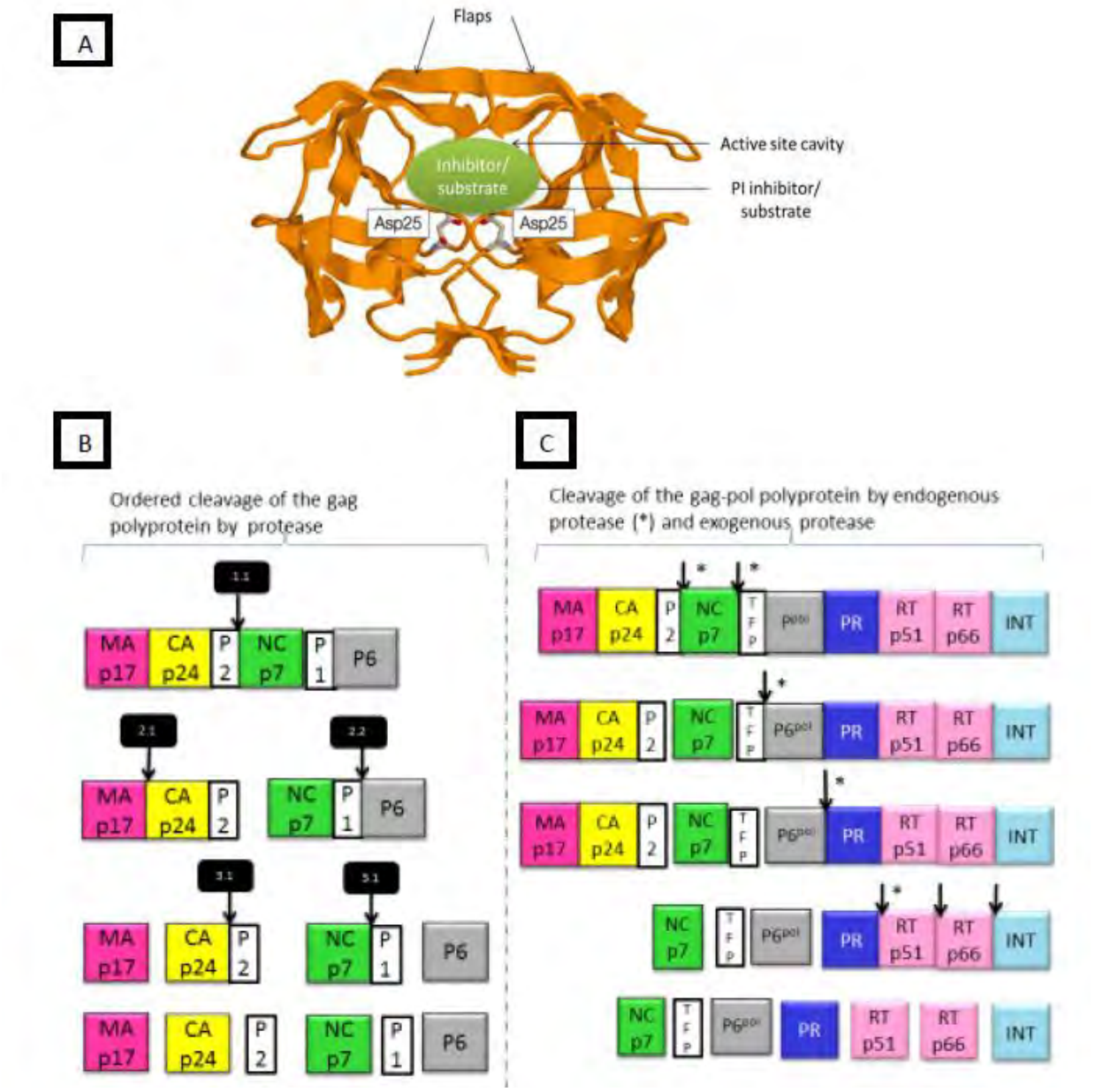


Figure 1-5 Crystal structure of Protease and cleavage sites in the Gag and Gag-pol polyprotein precursor required to undergo cleavage for viral maturation. (a) The two identical subunits (i.e. homodimer) of Protease, showing the two flexible flaps which enable entry of the polyprotein substrate/ inhibitor into the active site. The substrate/ inhibitor binds to the active site of Protease. Two aspartate residues at position 25 (i.e. Asp 25) provide hydrolytic activity to the enzyme. Modified from Blundell et al., 2002 (184).(b) The ordered cleavage of the Gag polyprotein, at five sites, by Protease is illustrated. The primary cleavage site is at the p2↓p7 junction. The secondary cleavage sites include the p1↓p6 site and the p17↓p24 junctions. Tertiary processing occurs at the p7↓p1 and p24↓p2 junction. Cleavage of Gag at these five sites results in the

production of core structural proteins including: matrix, capsid, p2, nucleocapsid, p1 and p6. Adapted from Pettit et al., 2004 (180), Fun et al., 2012 (183) and de Olivera et al., 2003 (179, 182, 185). (c) The cleavage of the Gag-pol polyprotein precursor. This process involves the use of endogenous Protease represented by an asterisk (*), to initially cleave the Gag-pol polyprotein at a few sites to release viral PR. These sites include: the p2↓p7 junction, the transframe and $p6^{pol}$ junction (TFR↓$p6^{pol}$) and the $p6^{pol}$↓PR junction sequentially (179, 180). Thereafter Protease which is released is used to cleave the pol polyprotein at three sites PR↓RTp51, RTp51↓RTp6, RTp66↓INT. The final products from Gag-pol cleavage include, nucleocapsid, a transframe protein, $p6^{pol}$, Protease, reverse transcriptase RNase H and integrase Adapted from Pettit et al., 2004 (180), Fun et al., 2012 (183) and de Olivera et al., 2003 (179, 182, 185).Abbreviations: Asp25 – Aspartate at position 25 of Protease; PI – Protease inhibitor; MA – matrix (p17), CA – capsid (p24); P2 – protein of 2 kDA; NC – nucleocapsid (p7); p1 – protein of 1 kDA; p6 – protein of 6 kDA; TFP – transframe protein; PR – Protease; RT (p51) – reverse transcriptase (protein of 51 kDA); RT (p66) – reverse transcriptase (protein of 66 kDA).

The viral replication cycle takes approximately 24 hours, with a possible 10^{10} to 10^{11} virions produced daily in chronically infected treatment naïve individuals (183). Several genetic variations of HIV-1 are produced in each replication cycle as a result of the low fidelity error prone viral RT used in reverse transcription and the subsequent lack of an error correcting/ proof reading mechanism (25). Some of these mutations enable HIV-1 immune escape thereby contributing to variations in HIV-1 pathogenesis amongst infected individuals.

1.7 HIV-1 pathogenesis

Transmission of HIV-1 is usually established by a single founder virus and most commonly occurs via exposure of a mucosal membrane; however transmission can also be percutaneous, in-utero, or intravenous (9, 186, 187).

Entry of HIV-1 into the genital or rectal submucosa is facilitated by: breaches in the epithelium caused by sexual intercourse, transcytosis through the mucosal epithelium, virus movement through epithelial intercellular spaces or interaction with intraepithelial dendritic cells (188). Signaling from mucosal epithelial cells recruits dendritic cells to the site of transmission. These dendritic cells secrete cytokines that attract activated $CD4^+$ T-lymphocytes, which are susceptible to HIV-1 infection, to the epithelium. Here, HIV-1 preferentially infects $CD4^+$ T-lymphocytes which co-express CCR5 receptors (189). The viral replication cycle begins once the viral gp120 protein binds to the cellular $CD4^+$ receptor of a target cell (9).

Irrespective of entry/transmission route, host and viral markers appear in an orderly pattern and help distinguish between stages of HIV-1 infection. In total there are three stages of HIV-1 infection: acute infection (i.e. primary infection), chronic infection (i.e. clinical latency) and advanced HIV-1 infection (also referred to as AIDS) (190), each of which is described below (Figure 1.6).

1.7.1 Acute infection

Acute HIV-1 infection occurs approximately 2 weeks after exposure to HIV-1 and ends once antibodies to HIV-1 are produced (9, 187, 188). It lasts approximately three to four weeks and is divided into an eclipse phase and five Fiebig stages which are characterized

by the progressive appearance of viral markers and antibodies in the blood (Figure 1.6b) (187, 188).

The eclipse phase represents the first 10 days after HIV-1 transmission, during which time the virus begins establishing itself in the local tissue at the site of exposure. At this time HIV-1 RNA is undetectable (187, 188). During this early phase, reservoirs of latently infected cells are established within $CD4^+$ memory T-lymphocytes and macrophages (191-193). These latently infected cells are able to carry HIV-1 without expressing antigens on their surface thereby enabling escape of host immune recognition and resistance to virus induced cytopathic effects (193). These viral reservoirs persist in the presence of ART and can be activated by cellular factors to produce infectious viruses (191, 193).

By the tenth day, cell free virus and infected cells reach the draining lymph node, were they encounter additional $CD4^+$ cells for infection (188). Dendritic cells internalize some viral particles and present them to activated $CD4^+$ T-lymphocytes further augmenting infection (88). Increased interaction between HIV-1 and cells expressing $CD4^+$ receptors results in increased cellular infection and a subsequent increased viral spread. This allows for dissemination of HIV-1 into the blood and lymphoid tissues, particularly the gut associated lymphoid tissue (GALT) where a significant fraction of $CD4^+$ T-lymphocytes reside (194, 195).

Replication of HIV-1 in the GALT and other lymphoid tissues results in an exponential increase in plasma viremia, which reaches a peak (i.e. > 1 million copies of virus per ml) between 21-28 days after infection (Figure 1.6) (188).

In response to peak viremia, the acutely infected individual mounts an intense inflammatory immune response which is characterized by high cytokine and chemokine levels, often referred to as a “cytokine storm” (196). Both adaptive and innate immune responses are activated and collectively they contribute towards a decrease in viral load. As part of the adaptive immune response, $CD8^+$ T-lymphocytes begin killing productively infected $CD4^+$ cells shortly after infection. Some viruses develop mutations in varying epitopes, as a result of error prone *de novo* viral replication (see section 1.6.9), and are able to escape immune selection (197). The innate immune response, driven mainly by NKCs also plays a role in control of viral load. Viruses however can also develop mutations which restrict the antiviral effects of NKCs (198).

A combination of the immune response and programmed cell death in response to viral infection is thought to contribute towards a decline in $CD4^+$ T-lymphocyte count and a decrease in viral load. Studies have shown that almost 80% of $CD4^+$ T-lymphocytes in the GALT are depleted within the first three weeks of HIV-1 infection (199). Once $CD4^+$ T-lymphocyte decline occurs infected individuals may become symptomatic (i.e. fever, headache, muscle ache, rash, lymphadenopathy, oral candidiasis, esophageal ulceration or anal ulceration and pharyngitis) (Figure 1.6) (200, 201).

Viral load continues to decrease until a point of stabilization is reached; this is referred to as the viral set-point and usually marks the onset of chronic HIV-1 infection (188). The viral set-point varies between individuals with a higher set-point associated with faster disease progression and lower viral set-point associated with slower disease progression. The maintenance of the viral set-point is facilitated by a balance between viral replication and host immune responses, with immune escape mutations being major contributors toward higher viral set-points (188).

Fiebig et al., 2003, divided acute viremia and early seroconversion into six stages, referred to as Fiebig stages, based on the detection of host and viral markers (188, 202). During Fiebig stage I (i.e. between days 10–15) viral RNA becomes detectable in blood by polymerase chain reaction (PCR). As part of Fiebig stage II (i.e. days 15–20), the p24[14] antigen becomes detectable by enzyme linked immunosorbent assays (ELISA). During Fiebig stage III (i.e. days 20–25), the first HIV-1 antibodies, which are unable to block virus entry into cells, become detectable by sensitive immunosorbent assays (i.e. seroconversion occurs). Fiebig stage IV (days 25 – 30) marks the beginning of viremic decline, thought to be driven by $CD8^+$ T-lymphocyte destruction of infected cells (9, 187, 188). This stage is characterized by an indeterminate western blot result for p31. Fiebig stage V (i.e. days 30 – 100) is marked by a positive western blot for p31. Fiebig stage V and VI (i.e. day 100 onwards) are associated with a plateau in viremia (i.e. viral set-point) with Fiebig stage VI marking the beginning of chronic HIV-1 infection (Figure 1.6b) (188, 202).

[14] The p24 antigen, a viral core protein, appears in blood once viral load reaches approximately 10 000 copies/ ml.

1.7.2 Chronic HIV-1 infection

Chronic HIV-1 infection is usually asymptomatic with a variable duration. It is characterized by a gradual progressive decline in $CD4^+$ T-lymphocyte count from the blood, mucosal tissue and lymphoid organs (203). This is mostly driven by: depletion of $CD4^+$ and $CCR5^+$ T-lymphocytes in the mucosa and lymphoid organs via direct or indirect viral cytopathicity (204, 205) and $CD8^+$ T-lymphocyte mediated destruction of infected $CD4^+$ T-lymphocytes and an eventual state of chronic immune activation (206). Viral replication continues during chronic infection, albeit at a steady rate mediated largely by the activated immune responses.

Several factors contribute to the state of chronic immune activation including: continuous immune stimulation (207), systemic inflammation (207), depletion of virally infected regulatory T cells (T_{reg}), direct activation of macrophages and T-lymphocytes by viral gp120, Nef and Tat, and allogeneic non-specific T-lymphocyte activation caused by molecular mimicry of the human leukocyte antigen (HLA) (188, 195, 203, 208-210). Continuous immune activation and systemic inflammation are explained further below.

The persistent replication of HIV-1 and subsequent high antigen loads result in continuous stimulation of immune cells that directly recognize components of HIV-1. This eventually results in functional and clonal exhaustion of immune cells (207) .

Translocation of microbial constituents (e.g. lipopolysaccharides [LPS], flagellin and cPGDNA) across the gut mucosa (driven by $CD4^+$ T-lymphocyte depletion in the GALT) results in the release of pro-inflammatory mediators (namely; interleukin-6 [IL-6], interleukin-1 [IL-1], tumor necrosis factor [TNF] and macrophage inflammatory protein-1 [MIP-1]) (203, 207, 211). These activate several subsets of immune cells including; $CD4^+$ T-lymphocytes, $CD8^+$ T-lymphocytes, NKC's, polynuclear neutrophils, monocytes and B-cells, which release inflammatory mediators and reactive oxygen species that subsequently drive systemic immune activation (203, 207).

Creating an environment of continued immune activation in conjunction with $CD4^+$ T-lymphocyte depletion has been shown to cause reactivation of latent viruses including: hepatitis B, hepatitis C, cytomegalovirus and Epstein-Barr virus. These viruses are commonly found in HIV-1 positive individuals, with their reactivation accounting for viral co-

infection seen in chronic HIV infection (203, 207). Additionally, latent HIV-1 reservoirs are also activated resulting in further viral replication and stimulation of uninfected $CD4^+$ T-lymphocytes (207). In the majority of cases, this leads to rapid apoptosis and has been considered as the biggest contributor to $CD4^+$ T-lymphocyte depletion (208).

The continuous HIV-1 infection of and replication in $CD4^+$ T-lymphocytes results in prolonged stimulation of $CD8^+$ T-lymphocytes which become exhausted and lose their cytokine secreting and cytolytic activity resulting in their depletion (198, 200). The control of viral propagation is ultimately impaired and even though $CD4^+$ T-lymphocytes are mostly depleted, viral replication continues to thrive in macrophages (212).

The state of chronic immune activation eventually culminates in immune exhaustion leaving an infected individual susceptible to opportunistic infections. Additionally, infected individuals may present with non-AIDS comorbidities including: cardiovascular disease, atherosclerosis, neurocognitive impairment and liver disease during this phase of HIV-1 infection (203). This marks the onset of advanced HIV-1 disease (i.e. AIDS) (203).

1.7.3 Advanced HIV-1 disease (AIDS)

A $CD4^+$ T-lymphocyte level of <200 cells/ mm^3 in conjunction with one or more AIDS defining illnesses marks the onset of AIDS. Acquired immunodeficiency syndrome defining illnesses include: oral candidiasis, tuberculosis, Karposi's sarcoma (caused by herpes simplex virus-8), lymphomas (caused by Epstein-Barr virus) and/or pneumococcal infections (207, 213, 214). Progression to AIDS usually takes between 8–10 years; this however is dependent on host and viral interactions (215). The progression to AIDS is divided into three categories: (1) fast, (2) intermediate or typical and (3) slow or long-term non-progression (213). Between 70-80% of infected individuals experience intermediate/typical progression to AIDS, whereby AIDS related illness occurs 6-10 years after HIV acquisition. In contrast 15% of infected individuals experience fast progression, whereby AIDS related illness occurs in <6 years. Furthermore, less than 5% of individuals are described as long term non-progressors, these individuals have been shown to develop AIDS related illness >10 years after HIV-1 acquisition (213).

Irrespective of the type of progression, once an individual reaches the AIDS stage of infection, the immune system is unable to cope and death ensues.

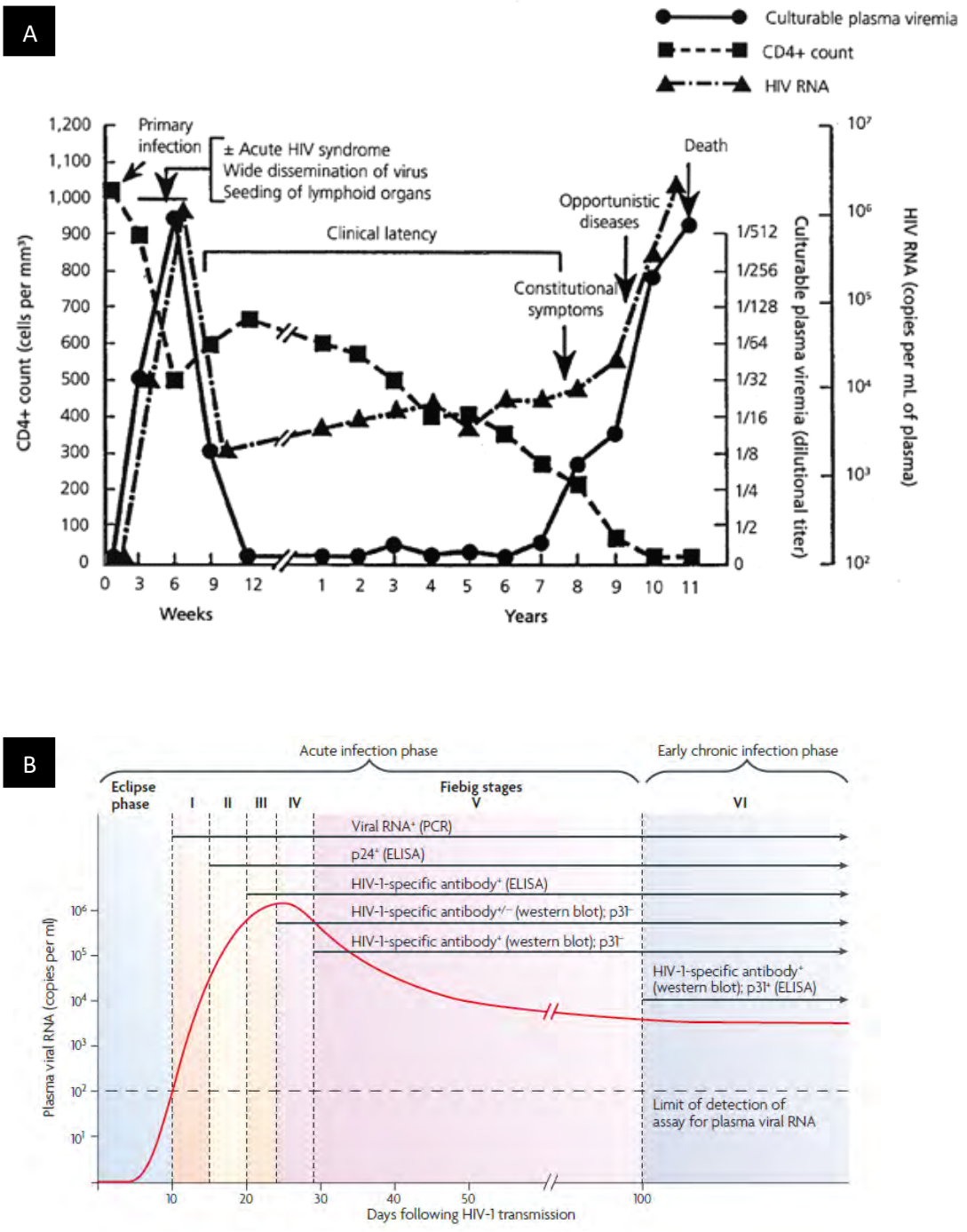


Figure 1-6 Illustration of the three stages of HIV-1 infection and diagram highlighting the Fiebig stages of infection. (a) Variations in viral load (HIV-1 RNA levels) and CD4⁺ T-lymphocyte counts during the progression of HIV-1 to AIDS. The stages of HIV-1 infection are divided into: primary infection (i.e. acute infection); clinical latency (i.e. chronic infection) and AIDS which is marked by the development of constitutional symptoms and opportunistic infections that culminates in death. Taken from Perlmutter et al., 1999 (216). (b) Overview of changes in viral load during

various stages of acute infection (i.e. eclipse phase and Fiebig stage I, II, III, IV, V) and chronic infection (i.e. Fiebig stage VI). The five stages of acute infection are illustrated by roman numerals (i.e. Fiebig stage I, II, III, IV, V), with the sixth Fiebig stage representing the onset of chronic HIV-1 infection. Diagnostic assays used to identify HIV-1 antigens and HIV-1 specific antibodies during the various stages of acute and chronic infection is provided in brackets. Taken from McMichael et al., 2010 (188). Abbreviations: HIV-1 – human immunodeficiency virus type 1; RNA – ribonucleic acid; CD4 – cluster of differentiation 4; PCR – polymerase chain reaction; ELISA – enzyme linked immunosorbent assay; p31 – protein of 31 kDA; ml – milliliters.

1.8 HIV pathogenesis and antiretroviral therapy

Treatment of HIV-1 infected individuals with ART has however transformed HIV-1 infection from a fatal to a chronic illness by prolonging the progression of HIV to AIDS (217). Patients with $CD4^+$ T-lymphocyte counts below 500 are treated with varying combinations of ART to decrease their viral load (i.e. to <50 copies/ml) (218). This enables significant immune reconstitution (measured as increase in $CD4^+$ T-lymphocyte count) and thereby prolongs progression to AIDS (219, 220). The long-term virologic suppression of HIV-1 is however hindered by the *in vivo* viral diversity of HIV-1 and the subsequent development of drug resistance. An overview of ART, drug resistance, mechanism of action of ARV's and their associated resistance pathways are discussed in the following section.

1.9 Antiretroviral therapy

There are currently 26 Food and Drug Administration (FDA) approved HIV-1 antiviral agents belonging to six drug classes namely: (1) NRTI's; (2) NNRTI's; (3) PI's; (4) integrase inhibitors (INI's); (5) fusion inhibitors (FI's) and (6) chemokine receptor antagonists (CCR5 inhibitors) (170). Each of which either targets an enzyme or a particular step in the lifecycle of HIV-1 to inhibit/prevent viral replication thereby reducing viral load (Figure 1.4c) (175).

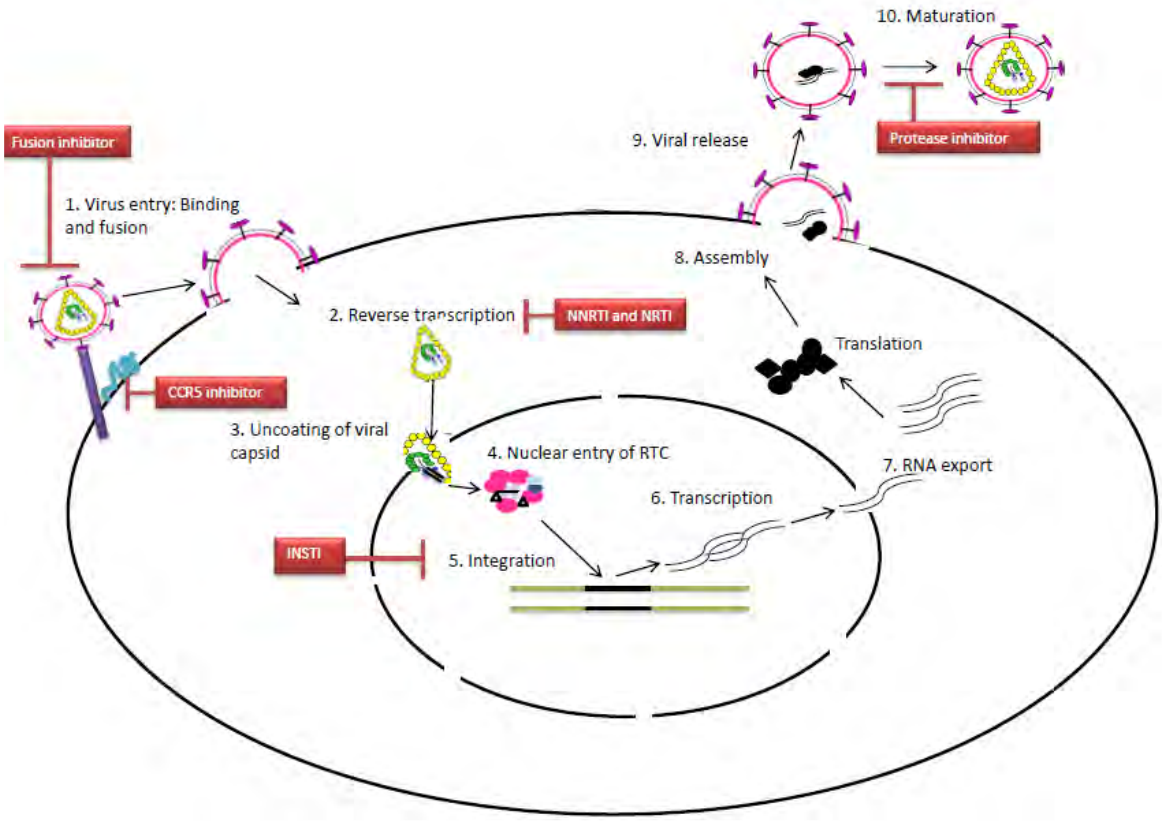


Figure 1-7 Site of action of antiretroviral drugs. The target site of: fusion inhibitors, CCR5 inhibitors, integrase strand transfer inhibitors (InSTI's), non-nucleoside reverse transcriptase inhibitors (NNRTIs), nucleotide transcriptase inhibitors (NRTIs) and protease inhibitors (PIs) are depicted by red boxes. Adapted from Fanales-Belasio et al., 2010 (42) and Engelman et al., 2012 (41, 84).

The first ARV approved by the FDA was Azidothymidine (AZT, Zidovudine), an NRTI, in 1987 (221, 222). It was administered as monotherapy and found to select for drug resistance (see section 1.9) thereby rendering its long-term use ineffective (223-225). This prompted the development of additional NRTI's in the early 1990's, the first PI in 1995 and the first NNRTI in 1996 (226). The availability of several drug classes acting on different stages in the lifecycle of HIV-1, led to the decision to utilize a combination of ARVs, referred to as combination ART (cART), to treat HIV-1. As part of cART, a combination of three drugs from at least two drug classes is administered on a daily basis, in order to target variants of HIV-1 that may display reduced susceptibility to any drug in the regimen (227, 228).

Currently high-income countries have access to all six classes of drugs, with the combination of drugs administered to a patient based upon their genotypic test results (229).

In contrast to high-income countries, most low income countries cannot afford to genotype patients prior to treatment initiation. Such countries usually only have access to three classes of HIV-1 drugs (namely: NRTI's, NNRTI's and PI's) that are provided as standardized treatment regimens (229). Options for salvage regimens are thus limited.

South Africa, a low-income country and the site of the present study, has access to four classes of HIV-1 drugs including: NRTI's, NNRTI's, PI's and InSTI's (218). These are used in three varying regimens. The first regimen comprises of two NRTI's and one NNRTI, provided as a fixed dose combination (FDC)[15] pill. The second regimen comprises of two NRTI's and a PI and the third treatment regimen comprises of an InSTI a PI and an NNRTI (Table 1.1) (218). If a patient reacts adversely to any drug in the first two regimens, a clinician can change the drug based on remaining options and the South African National Department of Health (NDOH) guidelines (218).

Interestingly, access to the third line regimen is limited and is managed centrally by the South African NDOH. Eligibility for the third line regimen stipulates that patients on a PI inclusive treatment for one year who have not achieved viral suppression must receive a genotype test. If PI resistance is detected, a full treatment history must be submitted to the NDOH where a committee reaches a consensus on the patient's future treatment. If agreed that a patient should receive third line treatment, the combination of drugs are sent to the respective facility on a named patient basis (218). The purpose of this strategy is to control the use of certain drugs in order to prevent widespread drug resistance (described in section 1.9) and a situation in which all treatment options are exhausted. Details of available treatment regimens in SA are described in Table 1.1 below.

[15] The fixed dose combination pill comprises of: Tenofovir, Efavirenz and either Lamivudine or Emitricitabine.

Table 1-1 Overview of available treatment options for first-line, second-line and third-line treatment regimens in SA.

First-Line ARV regimen (2 NRTI's + 1 NNRTI)		Second-Line ARV regimens (2 NRTI's + 1 Boosted PI)		Third-line ARV regimen (1 NNRTI + 1 PI + 1 InSTI)		
NRTI's	**NNRTI's**	**NRTI's**	**PI with Booster**	**NNRTI**	**PI**	**InSTI's**
Tenofovir (TDF)	Nevirapine (NVP)	Zidovudine (AZT)	Lopinavir (LPV)/ Ritonavir (r) (Kaletra)*	Etravirine (ETR)	Darunavir (DRV/r)	Raltegravir (RAL)
Emitricitabine (FTC)	Efavirenz (EFV)	Lamivudine (3TC)	Atazanvir (ATV)/r **			
Lamivudine (3TC)		Tenofovir (TDF)				
Abacavir (ABC)		Emitricitabine (FTC)				
		Abacavir (ABC)				
NB* TDF + FTC/3TC + EFV are provided as a fixed dose combination pill and comprise the standard first-line treatment for HIV-1 infection. If patients do not respond to the FDC pill they can be placed on a combination of any 3 ARVs available in the first line regimen list.				NB* combinations of other NNRTI's and NRTI's can be used in third line treatment depending on the patients resistance profile and ART history.		

* Protease inhibitors in regimen 2 are administered with a sub-therapeutic dose of Ritonavir to increase half-life.
** Is only used when a patient presents with dyslipidaemia or intractable diarrhoea associated with LPR/r

1.10 Drug resistance

HIV-1 drug resistance is defined as the ability of HIV-1 to mutate and reproduce itself in the presence of ARV drugs (230). This occurs as a result of: (a) poor adherence to treatment; (b) inadequate potency of ARVs; (c) suboptimal drug levels and (d) pre-existing resistance.

There are two categories of ARV drug resistance: acquired drug resistance and TDR. Acquired drug resistance occurs as a result of the development of resistance mutations, within the HIV-1 genome, due to drug-selection pressure in individuals receiving ARV

treatment. Transmitted drug resistance occurs when individuals who were not previously infected with HIV-1 become infected with a drug-resistant strain (231).

The most important viral factors responsible for the development of drug resistance are: the error prone *de novo* replication of HIV-1 caused by the use of its low fidelity RT during reverse transcription, the absence of an error correcting/proofreading mechanism during viral replication, rapid rate of viral replication, collection of archived proviral reservoirs during infection and genetic recombination when viruses with two different sequences infect the same cell (232-234).

It has been estimated that on each replication cycle, up to five incorrect nucleotides could be incorporated into the HIV-1 genome copied, with the likelihood of every single-base mutation possible in the HIV-1 genome occurring on a daily basis (232, 233). Since HIV-1 lacks error detection and correction mechanisms, these erroneous nucleotides cannot be corrected (235). The result is the production of HIV-1 virions which differ from wild-type (WT) HIV-1 (i.e. the quasispecies) in a matter of months after primary infection (232).

In the instance where mutations in sequences which code for viral enzymes occur, the result is the production of enzymes which differ slightly from the wild-type (i.e. production of mutant variants). If viral enzymes are altered/include mutations, the action of ARV's on these enzymes could be compromised thereby resulting in ARV resistance and prolonged viral replication (182).

In the absence of ARV drugs, the dominating population of HIV-1 is the WT virus. Mutant viruses still continues to replicate but are found in low levels. Commencement of drug therapy exerts a selective pressure on the viral population in which the WT virus is prevented from replicating and the viral load decreases. If there is a mutant variant of HIV-1 present, which displays resistance to the ARV treatment, replication will continue and the viral load of mutant virus will gradually increase. This can lead to failure of a regimen as well as the possible transmission of drug resistant variants (236, 237).

The location and pattern of mutations determines the type of resistance conferred. In some cases only one mutation is required to cause drug resistance (low genetic barrier); whilst in other cases more than one mutation is required to cause drug resistance (high genetic barrier) (237-239). In either instance, drug resistance impacts the concentration of an ARV

required to reduce viral replication by 50% (i.e. IC_{50}), with IC_{50} being greater in mutant viruses in comparison to the WT virus.

The use of combination therapy to treat HIV-1 infection is employed to target as many viral variants as possible, even those which may be resistant to an ARV in the regimen. In doing so viral replication is controlled more efficiently and the chances of replication of mutant viruses and transmission of drug resistant strains of HIV-1 are less likely. Successful drug therapy however, is dependent upon high levels of adherence to treatment. Poor drug adherence leads to suboptimal concentrations of drugs which results in viral rebound (240). The subsequent ongoing viral replication increases the probability of RAMs developing and this in turn increases the risk of transmitting drug resistant viral variants (226, 232, 237).

In order to monitor the emergence and transmission of drug resistance, the World Health Organization (WHO) implemented the global HIV drug resistance surveillance network in 2004. This network aimed to monitor the prevalence of acquired and TDR and to use such information to inform decisions on treatment and management strategies for HIV in low and high income countries whilst access to ARV's was scaled up (231).

Data from the WHO has shown that, over time high-income countries have experienced an increase in the number of HIV-1 positive treated patients who have achieved full viral suppression. This has in turn reduced emergence and subsequent transmission of drug resistance within these countries (231). In contrast, some low-income countries, such as SA, are experiencing difficulty in achieving full viral suppression of HIV-1 positive patients on ARVs. As a result of this, acquired drug resistance continues to persist and TDR is increasing, particularly in KZN, the site of the current study and the epicentre of the HIV epidemic (231, 241).

According to the WHO, TDR related to NNRTI's and NRTI's, in KZN, has increased from a low threshold level (5%) to a moderate threshold level (5-15%) between 2007 and 2012 (231). Two surveillance studies conducted in KZN in 2005 and 2009 identified TDR to be within the low threshold level (<5%), with the initial survey identifying no TDR and the follow up study showing one patient with an NNRTI associated TDR mutation (i.e. the K103N mutation) (231, 242). At the time of both surveys, <30% of HIV-1 infected individuals in SA were on cART and those who were receiving cART only did so for <5

years. Both surveys were thus in line with statistical models which suggested that the 5% (low) TDR threshold level may only be exceeded when >30% of HIV positive eligible individuals were on ARV treatment or 10 years after large scale rollout of ARV's (243). The 2012 WHO report, which showed an increase in TDR above 5%, was also in line with this model since SA had an ARV coverage of between 40-60% of HIV positive eligible individuals at the time of development of this survey (218).

The majority of drug resistance associated mutations found in KZN thus far have been related to NNRTI's and NRTI's, with very few PI resistance mutations being identified.

1.11 Overview of antiretroviral drugs

The mechanism of action of each drug class, their target and resistance pathways are described below.

1.11.1.1 Chemokine receptor antagonists (CCR5 inhibitors)

There is currently only one CCR5 inhibitor approved by the FDA (i.e. Maraviroc, MVC). Maraviroc is used to treat HIV-1 infection in R5 tropic individuals. In contrast to all other classes of HIV-1 drugs, which utilise viral targets, MVC targets the host protein CCR5 (244). It binds to an allosteric hydrophobic pocket produced by the transmembrane helices of host CCR5 and alters the conformation of the CCR5 co-receptor thereby rendering it unrecognizable to viral gp120 and preventing viral entry (245, 246).

Two diverse resistance mechanisms against MVC therapy have been reported as part of *in vitro* and *in vivo* studies. Moore et al., (2009) reported that HIV-1 can implement a tropism switch from CCR5 co-receptor use to CXCR4 co-receptor use in the presence of MVC, if CXCR4 co-receptors are available (247). This facilitates viral entry and enables the continuation of viral proliferation in the presence of bound MVC (247). The main path of resistance to MVC however involves the development of several mutations mostly in the V3 loop region of gp120 (248-250). These mutations enable the virus to utilize the CCR5 co-receptors with MVC bound to it (248, 251). There is currently no consensus on mutations which contribute to MVC resistance.

1.11.1.2 Fusion inhibitors

Fusion inhibitors target gp41, a transmembrane protein which anchors the viral envelop to the host cell membrane and undergoes conformational changes to facilitate the fusion of the host and viral lipid bilayers (252). As part of these conformational changes, the heptad repeat region 1 (HR-1) of gp41 folds onto its heptad repeat region 2 (HR-2). This shortens gp41 and is essential in the formation of a stable six helix bundle required for virus entry (described in section 1.6.1) (175).

Enfuvirtide (T-20), the only FDA approved FI, functions by binding to the HR-1 region of gp41 (175). In doing so it prevents the interaction between HR-1 and HR-2 required for fusion thereby resulting in inhibition of viral entry (described in section 1.6.1) (253). Mutations within the HR-1 region of gp41, specifically at positions G36, I37 and V38 have been shown to be associated with reduced susceptibility to T-20 and reduced viral replicative capacity, since they prevent binding of T-20 to the HR-1 region (254). Several combinations of mutations in the HR-1 region have also been associated with T-20 drug resistance including: I37M/N43D, Q41R/N43D, V38E/N42S and G36V/N42D (255-257) (258). A list of mutations associated with T-20 drug resistance is provided in Table 1.2.

1.11.1.3 Integrase strand transfer inhibitors

There are currently three InSTI's approved for use by the FDA namely: raltegravir (RAL), dolutegravir (DTG) and elvitegravir (EVG). Each of these InSTI's are referred to as integrase strand transfer inhibitors since they bind to the catalytic core of integrase (aa 50 to 212) and inhibit the strand transfer step of the viral integration process (described in section 1.6.5) (259).

Resistance to InSTI's is almost always caused by mutations within the integrase active site which is responsible for coordinating magnesium cofactors required in strand transfer (260, 261). Clinical studies on RAL have shown three independent sets of primary mutations/ pathways associated with RAL drug resistance: Q148, N155 and the Y143 pathway (262, 263). These primary mutations/pathways are accompanied by a variety of accessory mutations which either increase resistance to the drug or compensate for loss of activity caused by primary mutations (i.e. improves viral replication).

Position 148 of integrase is a critical component of the integrase active site. Mutations at this position (i.e. Q148K/H/R) cause reduced susceptibility to RAL and impair enzyme function resulting in a replication deficit. This deficit has been shown to be rescued by the G140S and/or the E138K INT compensatory/accessory mutations. Additional accessory mutations associated with the Q148 pathway include the: L74M, E92Q, T97A, G136R, or the V151I mutations (262, 263).

Position 155 of INT is located at the base of the HIV-1 integrase catalytic site and is involved in metal binding. Mutations at this position (namely, N155H) cause reduced RAL susceptibility. Accessory mutations commonly found with the N155H primary mutation include the: L74M, T97A, E157Q or G163R/K mutations (264). Collectively, primary and accessory mutations in this pathway confer high level resistance to RAL with no compensation offered to viral replicative ability (262).

The Y143 primary mutation/ pathway occurs less frequently than the N155 and Q148 pathways. It is commonly accompanied with the: L74A/I, T97A or the G163K/R accessory mutation (265, 266). Mutations in this pathway work to collectively increase RAL resistance.

The Q148 and N155 pathway also causes resistance to EVG (267, 268). Mutations at position 66 (T66A/I/K) and 92 (E92Q) have also been shown to cause direct resistance to EVG. Dolutegravir is thought to have a higher genetic barrier than RAL and EVG since it does not share their pathways to resistance (267). A list of InSTI-associated RAMs is provided in Table 1.2.

1.11.1.4 Nucleoside reverse transcriptase inhibitors

The term NRTI refers to both nucleoside and nucleotide reverse transcriptase inhibitors. This class of drug targets the action of RT, a multifunctional enzyme with both polymerase (i.e. DNA- and RNA dependent) and endonuclease (i.e. RNase H) activity (269). Reverse transcriptase is a heterodimer comprising of a 560 aa subunit (p66) and a 440 aa subunit (p51). Both subunits have the same aa sequences, the major difference being that p51 lacks endonuclease activity and a nucleic acid binding cleft, and thus has a shorter aa sequence (270). The p66 subunit comprises of four subdomains, which serve to join the polymerase and RNase H domains, they are named: fingers (aa 1 to 85 and 118 to 185),

palm (aa 86 to 117 and aa 156 to 236), thumb (aa 237 to 318) and connection (aa 319 to 426) (270) (Figure 1.7a). The polymerase active site, (i.e. the target for NRTI's) lies in the palm (Figure 1.7a, 1.7b).

Nucleoside reverse transcriptase inhibitors are administered as inactive prodrugs which enter an HIV-1 infected cell either via passive diffusion or nucleoside transporters (271). Once in the cell, they require metabolic transformation (i.e. phosphorylation) by cellular kinases into its active 5'-triphosphate form in order to elicit an antiviral effect (272). In its active form NRTI's mimic natural deoxynucleotide triphosphates (dNTP's) and compete with them for incorporation into the 3' end of the nascent viral DNA chain. The drugs' incorporation into viral DNA is as a monophosphate which lacks a 3' hydroxyl group at the sugar moeity. This prevents binding of further nucleotides culminating in the termination of chain elongation (175). The termination of chain elongation can occur during RNA or DNA dependent DNA synthesis and thus can inhibit the production of either the (-) or (+) strand of proviral DNA (273).

The backbone of most treatment regimens generally comprise of variations of NRTI's. Currently used NRTIs include: Lamivudine (3TC), Emiticitabine (FTC), Abacavir (ABC), Tenofovir (TDF) and Zidovudine (AZT). Stavudine (d4T) and Didanosine (ddI) are additional NRTI's, which were used in the past, they are however no longer recommended in treatment of HIV-1 due to their associated toxicities and side-effects (175, 218, 274).

Drug resistance to NRTI's can occur via two mechanisms. In the first instance, discriminatory mutations in viral DNA occur. These weaken the binding affinity of NRTI's whilst retaining the binding affinity of dNTP's. The result is reduced incorporation of NRTI's into viral DNA and continued incorporation of dNTPs into viral DNA resulting in ongoing viral proliferation (175, 272). The K65R, L74V, Q151M and M184V/I mutations are examples of discriminatory mutations which inhibit various NRTI's (Table 1) (175).

The second mechanism of NRTI drug resistance involves nucleotide excision by ATP dependent pyrophosphorolysis whereby RT utilizes co-substrates (i.e. either ATP or inorganic pyrophosphate) to remove the bound NRTI analogue monophosphate from the 3' end of the chain terminated viral DNA (275). This allows dNTP binding and chain elongation to resume, and thus viral proliferation continues in the presence of drug (276, 277). Thymidine analogue mutations (TAM's) are usually associated with

pyrophosphorolysis. There are two TAM pathways shown to cause drug resistance to NRTI's. The M41L, T210W, T215Y, and sometimes the D67N mutations constitute the first pathway (i.e. TAM1), whilst the second pathway (TAM2) involves the D67N, K70R, T215F, and 219E/Q mutations (278-282). A list of NRTI-associated DRMs is provided in Table 1.2.

1.11.1.5 Non-nucleoside reverse transcriptase inhibitors

Non-nucleoside reverse transcriptase inhibitors also targets RT for inhibition of viral replication (Figure 1.7b and 1.7c).

Non-nucleoside reverse transcriptase inhibitors bind to an allosteric hydrophobic site located at a distance of approximately 10 Å away from the RT polymerase active site (referred to as the NNRTI binding pocket [NBP]) (Figure 1.7b) (283). This induces conformational changes, whereby a hydrophobic pocket is formed proximal to the polymerase active site (284). The result is reduced polymerase activity and prevention of substrate alignment required for the formation of phosphodiester bonds (285, 286).

The NBP only exists in the presence of NNRTI's and consists of hydrophilic residues (including: K101, K103, S105, D192 and E224), hydrophobic residues (including: Y181, Y188, F227, W229 and Y232), residue E138 of the p55 domain and E224 of the p61 domain (287, 288). Amino acid substitutions at various positions of RT including position 100, 101, 103, 138, 179, 181 and 188 results in the incorrect formation of the NBP, this prevents NNRTI's from binding and as such confers resistance to the drug (289). The three most common NNRTI mutations include the K103N, Y181C and G190A mutations. Each of which confers high level NNRTI resistance and eventually results in clinical failure (290). A list of NNRTI-associated DRMs is provided in Table 1.2.

There are currently four NNRTI's used in the treatment of HIV-1 including: Efavirenz (EFV), Nevirapine (NVP), Rilpivirine (RPV) and Etravirine (ETR) (175). Lower toxicity levels and fewer side-effects associated with NNRTI's, in comparison to NRTI's, have encouraged their broad use in cART. A pitfall to the use of NNRTI's however is its low genetic barrier and ease with which drug resistance to NNRTI's can develop. Drug resistance mutations associated with NNRTI's generally have minimal effect on replicative ability of a mutant virus and have long reversion times in the absence of drug, making them easily transmissible (228).

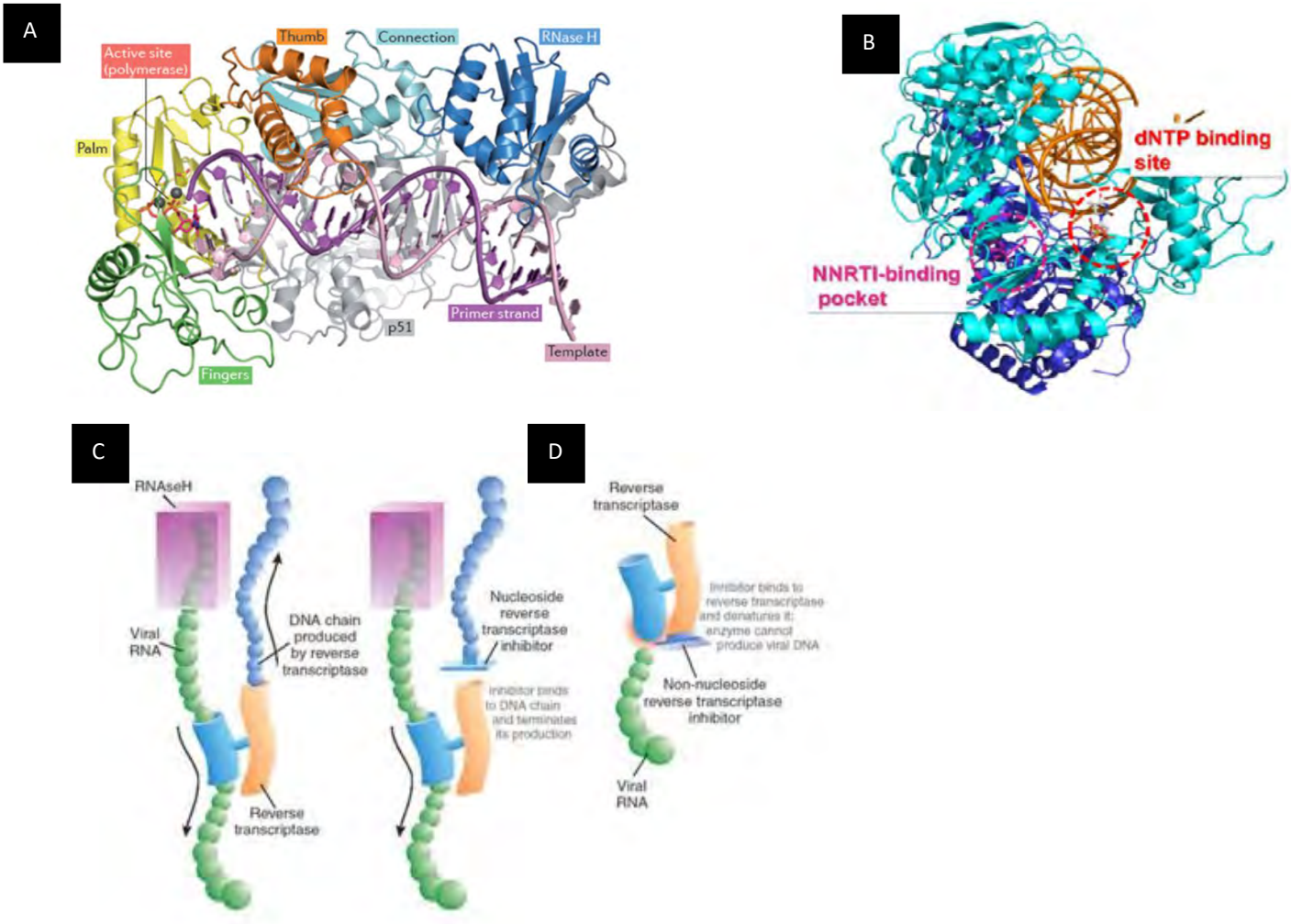


Figure 1-8 Crystal structure of reverse transcriptase with an illustration on the mechanism of action of NRTI's and NNRTI's. (a) The crystal structure of reverse transcriptase is shown. Grey ribbons represent the p51 domain. The p66 domain comprises of four subdomains: fingers (green), palm (yellow), thumb (orange) and connection (light blue) which collectively joins the polymerase active site to RNase H (dark blue). The template and primer strand are also represented in shades of purple. Taken from Engelman et al., 2012 (84). (b) This figure shows binding sites for NRTI's (red) and NNRTI's (magenta). Dark blue ribbons represent the p51 domain whilst light blue ribbons represent the p66 domain. Taken from Iyidogan et al., 2014 (175). (c) A schematic representation of a chain terminated viral DNA product, as a result of NRTI binding at the dNTP binding site, is depicted. Taken from Pomerantz et al., 2003 (291).(d) Binding of an NNRTI to RT which results in RT conformational changes that prevent action of RT. Taken from Pomerantz et al., 2003 (291). Abbreviations: NNRTI – non-nucleoside reverse transcriptase inhibitor; dNTP – deoxynucleotide triphosphates; RNA – ribonucleic acid; DNA – deoxyribonucleic acid.

1.11.1.6 Protease inhibitors

There are currently nine FDA approved PI's including: Lopinavir (LPV), Nelfinavir (NFV), Darunavir (DRV), Saquinavir (SQV), Atazanavir (ATZ), Indinavir (IDV), Tipranavir (TPV), Fosamprenavir (FPV) and Ritonavir (RTV) (175). All PI's, except NFV, are boosted/co-administered with RTV, which inhibits the metabolism of PI's by cytochrome P450 34A (CYP45034A), thereby increasing the bioavailability and half-life of PI's (292, 293).

Protease inhibitors target the action of HIV-1 PR, an enzyme required for cleaving the Gag and Gag-pol polyprotein precursors, in order to produce mature infectious viral particles (described in section 1.6.9, depicted in Figure 1.5) (170). All PI's, except TPV, bind to the active site (i.e. substrate binding pocket) of the Protease homodimer and competitively inhibits binding of the Gag and Gag-pol polyprotein substrates for cleavage (175). This prevents cleavage of the polyprotein precursors and as such virions remain immature and non-infectious. In contrast, TPV contains a dihydropyrone ring as its central scaffold that interacts directly with the flaps of HIV-1 PR, impacting on substrate entry and indirectly preventing substrate cleavage (294, 295). Together with DRV, TPV also inhibits dimerization of PR the impact of which is not fully elucidated but is thought to provide these drugs with a higher genetic barrier (182, 296).

Resistance to PI's occurs even though they are considered as high genetic barrier drugs, with more than two mutations generally required to confer resistance. Polymorphisms have been detected in 49 out of 99 aa sites in PR, with over 20 aa substitutions in PR found to be associated with PI resistance (297).

It has been suggested that PI associated drug resistance occurs in a stepwise manner with the initial development of primary mutations in PR (also referred to as major PI mutations) followed by the development of secondary mutations in PR and/or in Gag (also referred to as compensatory mutations) (175, 298).

Initially primary mutations (including D30N, I50V, V82A and I84V) occur at or near the PR active site (i.e., substrate binding site or catalytic site) causing altered electrostatic and hydrophobic interactions between the PI and the amino acids in the active site. This results in a reduced affinity of the binding site for PI's and a slightly reduced binding affinity of the binding site for the Gag and Gag-pol substrates which culminates in a deficit in replicative

ability of the virus (175, 299-301). Secondary mutations then arise in residues adjacent to the PR active site to either further inhibit PI resistance or to restore viral replication. For the latter, secondary mutations improve PR stability and activity to consequently restore viral replication (302-304).Viral replicative ability is also restored by the development of compensatory mutations in Gag and/or the evolution of Gag and Gag-pol CS's and non-cleavage sites (non-CS's) (305-309). Certain mutations at CS's can alter the Gag-pol frameshift resulting in the increased expression of pol products (i.e. more PR production, hence improved viral cleavage) (306). Mutations in non-CS's can improve the access of PR to cleavage sites thereby improving cleavage and viral replication (310-312).

The most common compensatory mutations in Gag have been shown to occur in the NC/p1 and p1/p6 cleavage sites, as these are the most variable sites in Gag after the p2/NC site (182, 311, 313). According to studies conducted on Subtype B, the L449F mutation in the p1/p6 cleavage site can only cause PI resistance in combination with the following PR mutations; D30N/ N88D, I50V and I84V (314). Similarly the p1/p6 cleavage site mutation P452K only causes PI resistance in the presence of the I84V/ L90M mutation in PR (315)

Mutations in Gag CS's and non-CS's can also occur as primary mutations to cause resistance to PI's in the absence of mutations in PR (177, 316-318). Amongst the identified primary drug resistant cleavage site mutations (CSM's) are A431V, I437V and the double mutation K436R and I437T, all located in the NC/p1 cleavage site of Gag (312, 315, 316, 319, 320).

A list of major and minor PI-associated RAMs in PR is given in Table 1.2, whilst mutations in Gag associated with PI resistance/exposure is summarized in Table 1.3 respectively.

Table 1-2 Overview of mutations associated with resistance to non-nucleoside reverse transcriptase inhibitors (NNRTI's), nucleoside reverse transcriptase inhibitors, protease inhibitors (major and primary mutations) integrase strand transfer inhibitors and fusion inhibitors. Adapted from the Stanford HIV drug resistance database (321) and the International AIDS society – USA (IAS-USA) database (322).

NON-NUCLEOSIDE REVERSE TRANSCRIPTASE ASSOCIATED RESISTANCE MUTATIONS																
CONSENSUS	V90	A98	L100	K101	K103	V106	V108	E138	V179	Y181	Y188	G190	H221	P225	F227	M230
Nevirapine			I	PEH	NS	AM	I		DEF	CIV	LCH	ASEQ			LC	L
Efavirenz			I	PEH	NS	AM	I		DEF	C	LCH	ASEQ		H	LC	L
Etravirine	I	G	I	PEH		I		KAGQ	DFT	CIV		EQ			C	L
Rilpivirine			I	PEH				KAGQR	DEFL	CIV	L	EQ	Y		C	LI

NUCLEOSIDE REVERSE TRANSCRIPTASE ASSOCIATED RESISTANCE MUTATIONS														
	Non-TAMs			TAMS								MDR		
CONSENSUS	M184	K65	K70	L74	Y115	M41	D67	K70	L210	T215	K219	T69	Q151	
Drug														
Lamivudine	VI	REN	E									Ins	M	
Emitricitabine	VI	REN	E									Ins	M	
Abacavir	VI	REN	E	VI	F	L			W	YF		Ins	M	
Didanosine	VI	REN	E	VI		L			W	YF		Ins	M	
Tenofovir		REN	E		F	L		R	W	YF		Ins	M	
Stavudine		REN	E			L	N	R	W	YF	QE	Ins	M	
Zidovudine						L	N	R	W	YF	QE	Ins	M	

MAJOR PROTEASE INHIBITOR ASSOCIATED MUTATIONS IN PROTEASE																		
CONSENSUS	L30	D32	V33	M46	I47	G48	I50	I54	Q58	T74	L76	V82	N83	I84	N88	L90		
Atazanavir/r		I	F	IL	V	VM	L	VTAM				ATSF		V	S	M		
Darunavir/r		I	F		V/A		V	LM			V	F		V				
Fosamprenavir/r		I	F	IL	V/A		V	VTALM			V	ATSF		V		M		
Indinavir/r		I		IL	V/A			VTALM			V	ATSF		V	S	M		
Lopinavir/r		I	F	IL	V/A	VM	V	VTALM			V	ATSF		V		M		
Nelfinavir	N		F	IL	V/A	VM		VTALM				ATSF		V	DS	M		
Saqinavir/r						VM		VTAM				ATS		V	S	M		
Tipranavir/r		I	f	IL	V/A			VTAM	E	P		TSL	D	V				

MINOR PROTEASE INHIBITOR ASSOCIATED MUTATIONS IN PROTEASE																		
CONSENSUS	L10	K20	L24	L33	E34	M36	K43	F53	D60	I62	L63	I64	H69	A71	G73	V77	I85	L89	I93
Drug																			
ATV/r	IFVC	RMITV	I	IFV	Q	ILV		LY	E	V		LMV		VITL	CSTA		V		LM
DRV/r				F														V	
FPV/r	FIRV														S				
IDV/r	IRV	MR	I			I								VT	SA	I			
LPV/r	FIRV	MR	I	F				L			P			VT	S				
NFV	FI					I								VT		I			
SQV/r	IRV		I							V				VT	S	I			
TPV/r	V			F		ILV	T						KR					IMV	

Abbreviations: ATV/r – Atazanavir boosted with ritonavir; DRV/r – Darunavir boosted with ritonavir; FPV/r = Fosamprenavir boosted with ritonavir; IDV – Indinavir boosted with ritonavir; LPV/r Lopinavir boosted with ritonavir; NFV – Nelfinavir; SQV/r – Saquinavir bosoted with ritonavir; TPV/r – Tipranavir boosted with ritonavir.

INTEGRASE STRAND TRANSFER INHIBITOR ASSOCIATED RESISTANCE MUTATIONS											
CONSENSUS	T66	L74	E92	T97	F121	E138	G140	Y143	S147	Q148	N155
Raltegravir	A	M	Q	A	Y	KA	SAC	CRH		HRK	H
Elvitegravir	IAK		QG	A	Y	KA	SAC		G	HRK	H
Dolutegravir			Q		Y	KA	SAC			HRK	

FUSION INHIBITOR ASSOCIATED RESISTANCE MUTATIONS IN GP41							
CONSENSUS	G36	I37	V38	Q39	Q40	N42	N43
Enfuvirtide	DS	V	AME	R	H	T	D

All amino acid substitutions are expressed relative to HIV-1 subtype B consensus HXB2 (GenBank accession number: K03455).

Table 1-3 Protease inhibitor resistance/exposure associated mutations in Gag. Adapted from Fun et al., 2012 (182) and Li et al., 2014 (323).

Gag Protein	Gag amino acid substitutions and their location		Reference
	Associated with PI resistance	Associated with PI exposure	
Matrix		E12K	(310, 324)
Matrix		G62R	(325, 326)
Matrix		L75R	(310, 324)
Matrix	R76K		(327, 328)
Matrix	Y79F		(327, 328)
Matrix	T81A		(327, 328)
Matrix		S125K	(326)
Matrix	V128I/T/A/del		(323, 329, 330)
Matrix	Y132F		(323, 326, 331, 332)
Capsid		H219Q/P	(310, 324)
Capsid		A360V	(331)
Capsid		V362I	(333)
Capsid		L363M/F/C/N/Y	(325)
P2		S368C/N	(320, 334)
P2		Q369H	(334)
P2		V370A/M/I/del	(14,(335)
P2		T371del	(312)
P2		S373P/Q/T	(320, 334)
P2		A374P/S/G/N/T	(336-339)
P2		T375N/S	(331, 336, 339)
P2		I376V	(320, 331)
Nucleocapsid		G381S	(331)
Nucleocapsid		I389T	(312)
Nucleocapsid		V390A/D	(310, 324)
Nucleocapsid		I401T/V	(312, 334)
Nucleocapsid		R409K	(310, 325)
Nucleocapsid		K415R	(323)
Nucleocapsid		Q430R	(323, 340)
Nucleocapsid	A431V		(305, 312, 315, 316, 320, 323, 331, 336, 341-345)
P1		G435E/R	(319, 346, 347)
P1	K436E/R/N		(315, 316, 319, 338, 346, 347)
P1	I437T/V/A		(315, 316, 319, 323, 325, 329, 338, 345, 346)
P1		W438R	(319)
P1		S440C	(346)
P6	L449F/P/V/Q		(312, 313, 315, 318, 323, 336, 338, 343, 345, 346, 348, 349)
P6		S451T/G/R/I/N	(315, 323, 350)
P6	R452K/S		(314, 315, 323, 343, 348)

Gag Protein	Gag amino acid substitutions and their location		Reference
	Associated with PI resistance	Associated with PI exposure	
P6	P453A/L/**T**		(312, 314, 315, 323, 341, 343, 344, 351, 352)
P6		E468K	(310)
P6		Q474L	(16)
P6		V484G/I/P/S	(338)
P6		A487S	(16)
P6		P497L	(16)

All amino acid substitutions are expressed relative to HIV-1 subtype B consensus HXB2 (GenBank accession number: K03455).

1.12 Drug resistance testing

Drug resistance testing is an essential component in the management of ARV treated individuals. It assists in identifying the appropriate treatment regimen for patients who are failing treatment as a result of drug resistant virus. In high-income settings, drug resistance testing is part of the standard of care of HIV positive individuals whereby each patient is prescribed a regimen based on the results from their drug resistance tests.

Drug resistance tests serve to identify the genotype (actual DNA sequence of virus) or phenotype (behaviour or physical traits expressed by the genotype) of the predominant viral strain (353).

Tests that are currently used to detect HIV-1 drug resistance include: viral genotyping (ie. Sanger sequencing, allele specific PCR, single genome amplification and ultra-deep pyrosequencing) and *in vitro* phenotypic drug resistance assays. Genotypic tests identify particular mutations/nucleotide substitutions associated with drug resistance in genes of interest whilst phenotypic tests measure the response (i.e. phenotype) of a virus to a particular ARV. Viral genotyping is generally the preferred method to monitor/test drug resistance due to its associated low cost, ease of use, availability and faster turnaround time. Various options are available for both genotypic and phenotypic testing as described below (304, 354).

1.12.1 Genotypic drug resistance testing

Most laboratories utilize either in-house developed methods or commercially available kits for standard genotyping of the HIV-1 PR and RT genes. There are currently three commercially available genotyping kits including: ViroSeq (Applied Biosystems, California), GeneSeq (Virologic, South San Fanscisco, CA) and TruGene (Bayer, Pittsburgh, PA) used for genotyping the PR and RT genes. All three of these kits require: viral RNA isolation from plasma of an infected individual, reverse transcription of RNA to cDNA and amplification of cDNA by PCR in order to produce adequate quantities of DNA for Dideoxynucleotide sequencing (Sanger sequencing) (297, 353)

As part of Sanger sequencing, amplicons are combined with a mixture of dNTPs, dideoxynucleotide triphosphates (ddNTPs), DNA polymerase and region specific primers. This mixture is subject to several thermal cycling steps to facilitate amplicon denaturation, primer annealing and dNTP incorporation into the growing strand. The incorporation of a ddNTP, in the place of a dNTP, results in chain termination and production of several strands of DNA of varying lengths. These strands are then sequenced using an automated sequencer based on a fluorometric method dependent upon labelling of either the primer or ddNTPs (355).

Sanger sequencing generates a consensus nucleotide sequence based on the most prevalent viral strain within a patient sample. This sequence is translated into a corresponding aa sequence and aligned to a reference sequence (i.e. WT strain, most commonly HXB2) (355). The input of the aligned sequence into one of several interpretation databases including: the Stanford HIV Drug Resistance Database (HIVdb), the Rega database (RegaDB) and the French AIDS research agency database (ANRS) produces a list of RAMs present in each sample and a score which predicts the level of drug susceptibility based on mutational patterns (356).

Whilst Sanger sequencing is widely used, it lacks the ability to detect DRMs at frequencies below 15 – 20% (357, 358). Several studies have shown that mutations occurring at frequencies as low as 1% can impact the clinical outcomes of a patient, as such various sequencing approaches have been developed to allow the detection of low frequency mutations (i.e. minority variants) (359-367). These include: point mutation assays (i.e.

allele specific PCR and oligonucleotide ligation assays) (364, 368), single genome amplification (SGA) assays (358) and next generation sequencing (369).

Point mutation assays generally employ specific probes/primers and/or labelled oligonucleotides to detect specific mutations of interest. Whilst the assay is very sensitive and can detect mutations present at frequencies between 0.1-1%, its major limit is only a few RAMs are detected in each run, rendering this method impractical for clinical settings where simultaneous detection of several RAMs is required per run. Additionally, the assay can be compromised by insensitivity of the template to primers and the large number of sequences which need to be analysed to detect low frequency mutations can be cumbersome (368, 370).

Single genome amplification (also known as limiting dilution PCR) involves the generation of cDNA from a patient sample, dilution of cDNA to one copy, amplification of that single template by PCR and sequencing of amplicons representing only one viral strain (369). This method avoids the preferential amplification of the dominant viral strain and reduces polymerase induced recombination artefacts commonly seen in bulk sequencing. The pitfalls however are that SGA is expensive, labour-intensive and several SGA's have to be conducted in order to achieve the same depth as that of multiple bulk PCRs (371).

Ultra-deep pyrosequencing (UDPS) is a parallel sequencing approach which allows for the sequencing of a mixed sample at a coverage of more than 1,000 reads per base. As part of UDPS viral RNA is extracted from plasma, purified and quantified. This is followed by generation of cDNA and the amplification of cDNA which is tagged with multiplex identifiers (MIDs). Tagged amplicons are purified, quantified, normalized and pooled. The sample pool is clonally amplified and sequenced. The three most commonly used deep sequencing platforms include: Roche 454 (GS Junior and GS-FLX), Ion Torrent PGM and Illumina Miseq (372). Deep sequencing offers extremely high throughput and requires expertise in bioinformatics to manage and interpret the data. Platforms such as the Roche 454 GS Junior and FLX are accompanied with built in bioinformatics tools which allows for user friendly data management and interpretation. A major benefit of UDPS is that all mutations in a region of interest can be identified and quantified in a single run making it suitable for use in clinical settings. Its major limitation however, particularly in resource limited settings, is its high cost (369, 370, 372).

1.12.2 Phenotypic drug resistance testing

Phenotyping assays measure the susceptibility of a clinical HIV-1 isolate to ARVs of interest by comparing the concentration of ARV required to inhibit the clinical sample to that of an HIV-1 WT/reference strain.

Similarly to conventional sequencing, phenotypic assays utilize viral RNA extracted from patient plasma for PCR in order to generate amplicons of the gene of interest. Amplicons are then used to generate recombinant viruses in a recombinant virus construct which has the analogous sequence deleted. A standard inoculum of recombinant virus is then used to infect a relevant cell line in the presence of varying concentrations of ARVs. The proliferation of the recombinant construct in the presence/absence of ARVs can be measured using either a single cycle phenotypic assay or a multiple cycle phenotypic assay. Results are obtained between –10 days and are reported as a fold change (FC) in drug susceptibility of the test sample in comparison to the reference strain (373).

As the name suggests, a single round infectivity assay is based on a single round of viral infection. A replication defective resistance test vector (RTV) is formed by cloning the region of interest from a patient sample into an HIV-1 expression vector which lacks the analogous region. Thereafter a cell line is co-transfected with a total of three plasmids: the RTV (comprising of the patient derived sequence), a vesicular stomatitis G protein expressing vector (this provides the Env region to the virus) and a reporter vector (this vector expresses luciferase which is used as a marker of viral replication and it also contains the HIV packaging sequence). The cell-line in conjunction with the three plasmids are exposed to varying concentrations of ARVs. The measurement of luciferase production in test samples versus the reference strain and drug control yields insight into viral replicative capacity and drug susceptibility.

For a multiple cycle assay, replication competent virus is produced via homologous recombination in cell culture to incorporate a patient derived sequence into a molecular HIV clone (typically NL43 which lacks the analogous region). A standardized inoculum of virus is then added to an appropriate cell line and the drug susceptibility is measured via luciferase production or expression of a reporter gene such as 3-(4, 5-dimethylthiazol)—2, 5 –diphenyltetrazolium bromide (MTT).

Whilst single cycle assays are performed in a shorter time and offer the benefit of accurate representation of the original virus, multiple cycle assays mimic in vivo conditions more closely and thus provide more accurate results.

Both the single and multiple cycle assays measure the change in IC50 of an ARV required to inhibit 50% of viral growth and report variations in drug susceptibility as FC. The FC is calculated by dividing the IC50 of the test sample by the IC50 of the reference strain.

There are currently two commercially available phenotyping kits: Antivirogram (Tibotec-Virco, Mechelen, Belgium) and Phenosense (Virologic, South San Francisco, California) that are used to measure variations in drug susceptibility of the PR, RT and a portion of the Gag gene. Antivirogram is a multiple cycle assay, whilst Phenosense is a single cycle assay (373).

1.13 Replication capacity and viral fitness

The main aim of ART is to suppress viral replication in order to stimulate immune reconstitution and reduce AIDS associated morbidity and mortality. In patients with a history of ARV failure, the selection of RAMs may prevent complete suppression of HIV-1 replication. Deeks et al., (2000) showed that when drug resistant viruses are selected under drug selection pressure, immunological stability can still be maintained despite ongoing slow viral replication (374). This has been attributed to mutations in genes such as PR and RT which induce structural changes that affect substrate binding and catalytic activity thereby impacting on the rate of viral replication (375-378). The effect of RAMs on viral replicative ability varies widely, with studies suggesting that clinical benefits can be derived from reduced viral replication (i.e. less fit viruses) particularly in resource limited settings where treatment options are limited (379).

Viral fitness is a measure of the capacity of HIV-1 to produce infectious progeny in a given environment. Viral fitness in HIV-1 is not a fixed attribute. This is largely due to the inherent genetic variation of HIV-1 whereby alterations at certain nucleotides cause disparities in replicative ability between viral variants within a given environment (380). This is portrayed by variations in replicative capacity between viral strains in treated and untreated HIV-1 infected individuals. In ARV treated individuals, the resistant viral strain is

more fit and hence replicates more than the WT viral strain, the converse however applies to treatment naïve individuals infected with HIV-1 (381, 382).

In most instances the evolution of viral fitness whilst on ARVs can be divided into two phases: (1) the development of primary mutations which drives reduced susceptibility and impacts negatively on replication capacity and (2) the development of secondary mutations (also referred to as accessory or compensatory mutations) which work in conjunction with primary mutations to improve the replication capacity of the virus (332).

1.13.1 Viruses used in replication capacity assays

Replication capacity assays make use of viruses in any one of the following forms: (1) site-directed mutants; (2) recombinant viruses or (3) whole viral isolates. Each of which is discussed below.

1.13.1.1 Site-directed mutants

The impact of a particular mutation or a group of mutations on viral fitness can be measured by introducing a mutation or group of mutations into a laboratory adapted strain of HIV-1 by site-directed mutagenesis (SDM) and comparing replication capacity of the mutant virus to that of a WT strain (383-385). Engineering the laboratory strain with a reporter gene such as; jellyfish green fluorescent protein (GFP, detected by flow cytometry) or firefly luciferase (detected by luminescence) enables easy detection of viral proliferation (386-388). However viral replication can also be measured by direct detection of gene products such as p24 or by measuring RT-activity. A limitation to using site-directed mutants is the oversight of other mutations which may contribute to viral kinetics. Whilst several mutations can be engineered into the laboratory strain at one time, there is still a chance that some mutations which work together are excluded (380). This holds particularly true for mutations in genomic regions omitted from the site-directed mutant.

1.13.1.2 Recombinant viruses

Recombinant viruses are generated by inserting an entire genomic region of interest into a standard viral backbone. This allows for links to be drawn between the viral region of interest and viral fitness (380).

A single clone or amplified pools of virus from a clinical sample can be used to generate recombinant viruses. Whilst single clones offer the advantage of using a precise known sequence, the use of amplified virus pools provides a sample that is more representative of *in vivo* viral diversity (380).

Recombinant viruses can be constructed by using one of four methods: (1) yeast recombination systems; (2) restriction enzymes; (3) homologous recombination of a vector and virus genomic region of interest in a cell line or (4) gene complementation which produces pseudovirions (387, 389-391). Each of these methods has its pitfalls.

Most yeast recombination systems usually require sub-cloning for viruses to be completely infective and are thus time-consuming and laborious (390, 392) . Restriction enzyme systems are limited by the variability in HIV-1 whereby restriction sites may not be available or may be unsuitable for use (390, 393). Homologous recombination systems are time-consuming and can result in poor recombination efficiency especially for eukaryotic samples (390). Gene complementation can result in the introduction of foreign genetic complements into the pseudovirus and generally produces viruses that can only be used in single cycle replication assays (see section 1.12.3) (390).

A limitation of using recombinant viruses is that the incorporated genomic region is not in its natural context and interactions with other genes are not accounted for. For example interactions between Gag and PR would not be accounted for if only PR was included in a recombinant virus. The most reliable fitness results are thus obtained from using whole viral isolates (380, 386).

1.13.1.3 Whole viral isolates

Whole HIV-1 viral isolates can be extracted from patient plasma or peripheral blood mononuclear cells (PBMCs) (394). This however can be costly, time-consuming and extraction in certain strains can be difficult (380).

Whole viral isolates can be applied to either PBMCs or cell lines to measure replication capacity. In PBMCs, replication capacity is measured by quantifying p24 (using ELISA) or by measuring RT-activity (386). If cell lines engineered to express reporter genes are used, measurement of replication capacity can be via flow cytometry (in the case of GFP)

or detection of luminescence (in the case of luciferase). The use of reporter gene engineered cell lines is less costly and simpler than using PBMCs (183, 395).

1.13.2 The use of primary cells versus T-cell lines

Replication capacity assays can utilize either primary cells (i.e. cells derived directly from human subjects such as PBMCs) or established T-cell lines (such as CEM-GXR cells). Studies have shown that results between PBMCs and cell-lines can differ (386, 396). Additionally results between different established cell lines can also differ (385). Whilst PMBCs supposedly offer results most representative of *in vivo* environments, they cannot be maintained for long times, they can be highly variable between donors and they require stimulation prior to use (183, 386). These limitations can be addressed by using established cell lines.

1.13.3 Measuring viral replication capacity

Viral fitness can be measured by *in vivo* and *in vitro* techniques. *In vivo* techniques involve comparing the quantity of mutant and WT virus detected within *in vivo* populations with the most commonly used sample being blood (397). Whilst this technique mimics the hosts' natural environment and provides the best estimate of viral fitness, it is limited by variation in quantities of viral variants in different compartments within the host (398). For example the most dominant quasispecies in the blood may differ from that in the lung. This limits its use in studies involving host genetics, immune response and drug resistance (398).

In contrast to *in vivo* methods, *in vitro* techniques do not mimic the natural environment of the host. *In vitro* techniques employ either HIV-1 isolates or recombinant viruses in a specific controlled environment thereby making the method useful in the study of drug resistance. There are two types of *in vitro* assays: single cycle assays and multiple cycle assays. Both of which employs the use of recombinant viruses or pseudovirions.

1.13.4 Single cycle replication capacity assays

The single cycle assay involves the infection of a cell line with a recombinant virus, which encodes reporter genes, and the subsequent detection of the reporter gene in the cell line (either by luminescence or fluorescence) between 24–72 hours post infection (380, 383,

391, 399). Whilst this assay yields results quickly it is unable to measure the entire replication cycle and is thus less sensitive than multiple cycle assays (183).

1.13.5 Multiple cycle replication capacity assays

Multiple cycle replication capacity assays are divided into pairwise growth competition assays and parallel assays, both of which are described below.

1.13.5.1 Pairwise growth competition assays

For growth competition assays, two viral variants are mixed and added to the same experiment. Both viruses are exposed to identical experimental conditions and compete for the same resources (380). During several passages, the fitter virus out-competes the less fit variant to become the predominant population, the proportion of which can be measured by Sanger sequencing, heteroduplex tracking assays or real-time PCR (386). These detection techniques are however expensive, labour intensive and produce data which is not easily analysed (183, 386). In order to overcome these pitfalls, cell lines or backbones of viruses (i.e. only in the case of recombinant viruses) can be engineered to include reporter genes which can be detected by fluorescent antibodies using flow cytometry (388, 400).

1.13.5.2 Parallel assays

In contrast to pairwise growth competition assays, parallel assays involve the quantification of HIV-1 replication in parallel cultures (401, 402). Cell lines or PBMCs are infected with a particular virus; replication capacity is then quantified by measuring p24 levels or RT-activity in the supernatant at various time-points (403, 404). Reporter genes in the backbone may also be used in detection. This method is simpler and less labor intensive than growth competition assays but it does not allow for the identification of subtle differences in replication kinetics between viruses that are tested, since cell populations in the parallel cultures may grow at slightly different rates thereby influencing the calculation of replicative capacity (380). Parallel assays are however more sensitive than single cycle assays.

1.14 The current study: Rationale, aims and objectives

1.14.1 Study Rationale

South Africa has the highest prevalence of HIV-1 worldwide, with one of the largest global ARV programs comprising of two main treatment regimens and limited access to an extremely costly third-line regimen (12, 218, 231, 405, 406). The main challenge in having a large ARV programme in a resource limited setting however is the development of acquired drug resistance and subsequent transmission of resistant variants. Evidence showing that acquired and transmitted drug resistance is increasing in SA is emerging (229, 231).

A recent South African study showed that 40% of patients receiving a PI inclusive treatment regimen experienced virologic failure in the absence of PR mutations (229). Whist this could be attributed to poor treatment adherence, several studies have shown that mutations in Gag can confer primary resistance to PI's in the absence of PR mutations (182, 311, 316, 328). The majority of studies investigating the role of Gag in PI resistance have however been conducted primarily on HIV-1 subtype B (182, 311, 316, 328), despite HIV-1 subtype C being the most prevalent subtype globally (407). With the polymorphic nature of Gag (408), a study employing an HIV-1 subtype C cohort to investigate the role of Gag in PI resistance is required.

Despite the large number of patients experiencing virologic failure on a PI inclusive treatment regimen, TDR studies in SA have not reported the presence of TDR mutations in PR (231). This could be attributed to the high associated fitness cost of PR mutations which would either cause mutations to revert rapidly or exist at low frequencies in treatment naïve individuals (353, 409-417). The majority of studies in SA have utilised recently infected cohorts to study TDR (231, 242). Recent infection however is not well defined and in some instances can reflect samples collected from patients >3 months after onset of plasma viremia. This would mean that rapidly reverting TDR mutations are not identified. Since it has been reported that viral reservoirs form as early as 10 days after onset of clinical symptoms in primary HIV-1 infection, studying the earliest point of infection would be more representative of the TDR mutation population. Furthermore all TDR studies in SA to date have employed Sanger sequencing to identify TDR mutations. Sanger sequencing however can only detect mutations at frequencies >15-20% which

would mean that low frequency mutations (i.e. those present at a frequency of <15%) would remain undetected (357, 358). With several studies demonstrating that mutations present at frequencies as low as 1% can impact treatment outcomes (359-367), the use of deep sequencing technologies to detect these mutations is necessary.

The focus of this study was to identify mutations in Gag-Protease that were associated with PI resistance/exposure and to determine the impact of these mutations on replication capacity, drug susceptibility and polyprotein cleavage, in an HIV-1 subtype C cohort of patients failing a PI inclusive treatment regimen from KwaZulu-Natal, South Africa. The presence of both high frequency and low frequency TDR mutations was also investigated using an HIV-1 subtype C acute infection cohort from KwaZulu-Natal.

The data generated from this study, on acquired and transmitted drug resistance could be used to inform public health policy and to facilitate improved first and second-line treatment regimens in South Africa. It can also be used in the clinical management of patients. Such data could also contribute to the continued development of the national ARV roll-out programme.

1.14.2 Aims and objectives

This study comprised of two main aims.

1.14.2.1 Aim 1

The first aim was to examine the hypothesis that mutations in both the amino and carboxy terminal of Gag contribute to reduced PI drug susceptibility and altered viral fitness in HIV-1 subtype C infected participants who are failing a PI inclusive treatment regimen. The specific aims were to:

- Identify mutations in Gag-Protease associated with PI exposure or resistance in 80 HIV-1 subtype C infected participants who were failing a PI inclusive treatment regimen.
- Determine the effect of Gag-Protease mutations on replication capacity in 80 HIV-1 subtype C infected participants who were failing a PI inclusive treatment regimen.
- Determine the effect of selected Gag-Protease mutations on PI drug susceptibility.
- Determine the effect of particular mutations in Gag-Protease on replication capacity, polyprotein cleavage and drug susceptibility using 20 site-directed mutant viruses.

The following **specific objectives** were fulfilled to address each of the above mentioned aims:

- Resistance genotyping using Sanger sequencing was performed on 80 samples from patients failing a PI inclusive treatment regimen in order to identify PI associated RAMs and to elucidate mutations in Gag. This was conducted in collaboration with an MSC student (418).
- Recombinant viruses encoding patient derived Gag-Protease sequences from patients failing a PI inclusive treatment regimen (n=80) were constructed by homologous recombination in a GFP-reporter CEM-GXR cell line.
- Replication capacity of each recombinant virus was measured using a multiple cycle, parallel mono infection, flow cytometry based assay which employed a GFP reporter CEM-GXR cell line (n=80).

- Drug susceptibility of selected samples (n=18) was measured using a multiple cycle assay which employed a luciferase reporter TZMBL-cell line based phenotypic assay.
- Mutant viruses were produced by the process of SDM to generate a total of 20 mutants, which were subject to replication capacity, proteolytic cleavage and drug susceptibility assays.

1.14.2.2 Aim 2

The second main aim of the study was to examine the hypothesis that NRTI, NNRTI and PI associated RAMs exist at low frequencies in individuals acutely infected with HIV-1 subtype C. The specific aim was to:

- Determine the presence of low frequency mutations associated with TDR in a cohort of acutely infected participants.

The **specific objectives** fulfilled to address this aim included:

- Genotypic resistance testing (using Sanger sequencing) of RT and PR was conducted on 47 samples from an HIV-1 subtype C acute infection cohort.
- UDPS of a subset of these samples (n=14) was performed to detect low frequency mutations in RT, PR and INT.
- Exploratory analysis was conducted to determine the impact of low frequency mutations on treatment outcomes.

1.15 Structure of thesis

The current thesis is presented in seven chapters. Four of which are presented as manuscripts. The first chapter provides an introduction to the study and a literature view. It also highlights the study rationale and provides information on the aims and specific objectives addressed in this study. The second chapter identifies mutations in Gag and Protease which are associated with PI resistance/exposure and also identifies novel Gag mutations associated with PI resistance/exposure. In this chapter, the frequency of mutations was analysed and combinations of mutations were assessed. The third chapter investigates the impact of mutations identified in the second chapter on replication capacity and drug susceptibility. The fourth chapter validates the role of mutations

identified to be significantly associated with replication capacity or drug susceptibility in the third chapter. The fifth chapter comprises of a paper which has been submitted for review. This paper addresses the prevalence of TDR mutations in an acute cohort and includes an exploratory analysis on the impact of low frequency TDR mutations on treatment outcomes. The sixth chapter of this thesis comprises of a general discussion and the seventh chapter comprises of appendices.

1.16 References

1. Chiu IM, Yaniv A, Dahlberg JE, Gazit A, Skuntz SF, Tronick SR, Aaronson SA. 1985. Nucleotide sequence evidence for relationship of AIDS retrovirus to lentiviruses. Nature 317:366-368.
2. Gottlieb MS, Schroff R, Schanker HM, Weisman JD, Fan PT, Wolf RA, Saxon A. 1981. Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. N Engl J Med 305:1425-1431.
3. Siegal FP, Lopez C, Hammer GS, Brown AE, Kornfeld SJ, Gold J, Hassett J, Hirschman SZ, Cunningham-Rundles C, Adelsberg BR, et al. 1981. Severe acquired immunodeficiency in male homosexuals, manifested by chronic perianal ulcerative herpes simplex lesions. N Engl J Med 305:1439-1444.
4. Friedman-Kien AE. 1981. Disseminated Kaposi's sarcoma syndrome in young homosexual men. J Am Acad Dermatol 5:468-471.
5. Klimas N, Koneru AO, Fletcher MA. 2008. Overview of HIV. Psychosom Med 70:523-530.
6. Girard MP, Osmanov S, Assossou OM, Kieny MP. 2011. Human immunodeficiency virus (HIV) immunopathogenesis and vaccine development: a review. Vaccine 29:6191-6218.
7. World Health Organisation. 2014. Global health observatory (GHO) data. http://www.who.int/gho/hiv/en/. Accessed 15 August 2015.
8. UNAIDS. 2014. 2014 Global statistics - fact sheet.
9. Maartens G, Celum C, Lewin SR. 2014. HIV infection: epidemiology, pathogenesis, treatment, and prevention. Lancet 384:258-271.
10. Statistics South Africa. 2014. Mid year population estimates. South Africa.
11. Rossouw TM. 2014. Editorial commentary: monitoring early warning indicators for HIV drug resistance in South Africa: challenges and opportunities. Clin Infect Dis 58:1615-1617.
12. UNAIDS. 2012. Regional Fact Sheet - Sub Saharan Africa. http://www.unaids.org/en/media/unaids/contentassets/documents/epidemiology/2012/gr2012/2012_FS_regional_ssa_en.pdf. Accessed 15 March
13. World Health Organisation. 2014. Global Health Observatory Map Gallery. http://gamapserver.who.int/mapLibrary/Files/Maps/HIV_all_2013.png. Accessed 15 August 2015.

14. Greene WC. 2007. A history of AIDS: looking back to see ahead. Eur J Immunol 37 Suppl 1:S94-102.
15. Center for disease control, Morbidity & Mortality Weekly Report. 1981. Kaposi's sarcoma and pneumocystis pneumonia among homosexual men - New York City and California. MMWR Weekly 30:305–308.
16. Marx JL. 1982. New disease baffles medical community. Science 217:618-621.
17. Centre for disease control, Morbidity & Mortality Weekly Report. 1982. Epidemiologic notes and reports: possible transfusion-associated acquired immune deficiency syndrome, AIDS - California. MMWR Weekly 31:652-654.
18. Barre-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vezinet-Brun F, Rouzioux C, Rozenbaum W, Montagnier L. 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science 220:868-871.
19. Popovic M, Sarngadharan MG, Read E, Gallo RC. 1984. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. Science 224:497-500.
20. Levy JA, Hoffman AD, Kramer SM, Landis JA, Shimabukuro JM, Oshiro LS. 1984. Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. Science 225:840-842.
21. Ratner L, Gallo RC, Wong-Staal F. 1985. HTLV-III, LAV, ARV are variants of same AIDS virus. Nature 313:636-637.
22. Marx JL. 1985. A virus by any other name. Science 227:1449-1451.
23. Coffin J, Haase A, Levy JA, Montagnier L, Oroszlan S, Teich N, Temin H, Toyoshima K, Varmus H, Vogt P, et al. 1986. What to call the AIDS virus? Nature 321:10.
24. Clavel F, Guetard D, Brun-Vezinet F, Chamaret S, Rey MA, Santos-Ferreira MO, Laurent AG, Dauguet C, Katlama C, Rouzioux C, et al. 1986. Isolation of a new human retrovirus from West African patients with AIDS. Science 233:343-346.
25. Sharp PM, Hahn BH. 2011. Origins of HIV and the AIDS pandemic. Cold Spring Harb Perspect Med 1:a006841.
26. Tebit DM, Arts EJ. 2011. Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease. Lancet Infect Dis 11:45-56.
27. Ng OT, Eyzaguirre LM, Carr JK, Chew KK, Lin L, Chua A, Leo YS, Redd AD, Quinn TC, Laeyendecker O. 2012. Identification of new CRF51_01B in Singapore

using full genome analysis of three HIV type 1 isolates. AIDS Res Hum Retroviruses 28:527-530.

28. Plantier JC, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, Lemee V, Damond F, Robertson DL, Simon F. 2009. A new human immunodeficiency virus derived from gorillas. Nat Med 15:871-872.
29. Delatorre EO, Bello G. 2012. Phylodynamics of HIV-1 subtype C epidemic in east Africa. PLoS One 7:e41904.
30. Hemelaar J. 2013. Implications of HIV diversity for the HIV-1 pandemic. J Infect 66:391-400.
31. Korber B, Muldoon M, Theiler J, Gao F, Gupta R, Lapedes A, Hahn BH, Wolinsky S, Bhattacharya T. 2000. Timing the ancestor of the HIV-1 pandemic strains. Science 288:1789-1796.
32. Sharp PM, Hahn BH. 2008. AIDS: prehistory of HIV-1. Nature 455:605-606.
33. Keele BF, Van Heuverswyn F, Li Y, Bailes E, Takehisa J, Santiago ML, Bibollet-Ruche F, Chen Y, Wain LV, Liegeois F, Loul S, Ngole EM, Bienvenue Y, Delaporte E, Brookfield JF, Sharp PM, Shaw GM, Peeters M, Hahn BH. 2006. Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. Science 313:523-526.
34. Nerrienet E, Santiago ML, Foupouapouognigni Y, Bailes E, Mundy NI, Njinku B, Kfutwah A, Muller-Trutwin MC, Barre-Sinoussi F, Shaw GM, Sharp PM, Hahn BH, Ayouba A. 2005. Simian immunodeficiency virus infection in wild-caught chimpanzees from cameroon. J Virol 79:1312-1319.
35. Santiago ML, Range F, Keele BF, Li Y, Bailes E, Bibollet-Ruche F, Fruteau C, Noe R, Peeters M, Brookfield JF, Shaw GM, Sharp PM, Hahn BH. 2005. Simian immunodeficiency virus infection in free-ranging sooty mangabeys (Cercocebus atys atys) from the Tai Forest, Cote d'Ivoire: implications for the origin of epidemic human immunodeficiency virus type 2. J Virol 79:12515-12527.
36. Wertheim JO, Worobey M. 2009. Dating the age of the SIV lineages that gave rise to HIV-1 and HIV-2. PLoS Comput Biol 5:e1000377.
37. Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, Tatem AJ, Sousa JD, Arinaminpathy N, Pepin J, Posada D, Peeters M, Pybus OG, Lemey P. 2014. HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. Science 346:56-61.
38. Arien KK, Vanham G, Arts EJ. 2007. Is HIV-1 evolving to a less virulent form in humans? Nat Rev Microbiol 5:141-151.

39. Hemelaar J, Gouws E, Ghys PD, Osmanov S, Isolation W-UNfH, Characterisation. 2011. Global trends in molecular epidemiology of HIV-1 during 2000-2007. AIDS 25:679-689.
40. Hemelaar J. 2012. The origin and diversity of the HIV-1 pandemic. Trends Mol Med 18:182-192.
41. Fanales-Belasio E, Raimondo M, Suligoi B, Butto S. 2010. HIV virology and pathogenetic mechanisms of infection: a brief overview. Ann Ist Super Sanita 46:5-14.
42. Fiorentini S, Marini E, Caracciolo S, Caruso A. 2006. Functions of the HIV-1 matrix protein p17. New Microbiol 29:1-10.
43. Valsamakis A, Schek N, Alwine JC. 1992. Elements upstream of the AAUAAA within the human immunodeficiency virus polyadenylation signal are required for efficient polyadenylation in vitro. Mol Cell Biol 12:3699-3705.
44. Turner BG, Summers MF. 1999. Structural biology of HIV. J Mol Biol 285:1-32.
45. Mervis RJ, Ahmad N, Lillehoj EP, Raum MG, Salazar FH, Chan HW, Venkatesan S. 1988. The Gag gene products of human immunodeficiency virus type 1: alignment within the Gag open reading frame, identification of posttranslational modifications, and evidence for alternative Gag precursors. J Virol 62:3993-4002.
46. Henderson LE, Bowers MA, Sowder RC, 2nd, Serabyn SA, Johnson DG, Bess JW, Jr., Arthur LO, Bryant DK, Fenselau C. 1992. Gag proteins of the highly replicative MN strain of human immunodeficiency virus type 1: posttranslational modifications, proteolytic processings, and complete amino acid sequences. J Virol 66:1856-1865.
47. Watts JM, Dang KK, Gorelick RJ, Leonard CW, Bess JW, Jr., Swanstrom R, Burch CL, Weeks KM. 2009. Architecture and secondary structure of an entire HIV-1 RNA genome. Nature 460:711-716.
48. Ghanam RH, Samal AB, Fernandez TF, Saad JS. 2012. Role of the HIV-1 Matrix Protein in Gag Intracellular Trafficking and Targeting to the Plasma Membrane for Virus Assembly. Front Microbiol 3:55.
49. Ganser-Pornillos BK, Yeager M, Sundquist WI. 2008. The structural biology of HIV assembly. Curr Opin Struct Biol 18:203-217.
50. Darlix JL, de Rocquigny H, Mauffret O, Mely Y. 2014. Retrospective on the all-in-one retroviral nucleocapsid protein. Virus Res 193:2-15.
51. Meng B, Lever AM. 2013. Wrapping up the bad news: HIV assembly and release. Retrovirology 10:5.

52. Sadiq SK, Noe F, De Fabritiis G. 2012. Kinetic characterization of the critical step in HIV-1 Protease maturation. Proc Natl Acad Sci U S A 109:20449-20454.
53. Naicker P, Seele P, Dirr HW, Sayed Y. 2013. F99 is critical for dimerization and activation of South African HIV-1 subtype C Protease. Protein J 32:560-567.
54. Arnold E, Sarafianos SG. 2008. Molecular biology: an HIV secret uncovered. Nature 453:169-170.
55. Delelis O, Carayon K, Saib A, Deprez E, Mouscadet JF. 2008. Integrase and integration: biochemical activities of HIV-1 integrase. Retrovirology 5:114.
56. Hallenberger S, Bosch V, Angliker H, Shaw E, Klenk HD, Garten W. 1992. Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160. Nature 360:358-361.
57. Camerini D, Seed B. 1990. A CD4 domain important for HIV-mediated syncytium formation lies outside the virus binding site. Cell 60:747-754.
58. Das K, Bandwar RP, White KL, Feng JY, Sarafianos SG, Tuske S, Tu X, Clark AD, Jr., Boyer PL, Hou X, Gaffney BL, Jones RA, Miller MD, Hughes SH, Arnold E. 2009. Structural basis for the role of the K65R mutation in HIV-1 reverse transcriptase polymerization, excision antagonism, and tenofovir resistance. J Biol Chem 284:35092-35100.
59. Kim HY, Choi BS, Kim SS, Roh TY, Park J, Yoon CH. 2014. NUCKS1, a novel Tat coactivator, plays a crucial role in HIV-1 replication by increasing Tat-mediated viral transcription on the HIV-1 LTR promoter. Retrovirology 11:67.
60. Bagashev A, Sawaya BE. 2013. Roles and functions of HIV-1 Tat protein in the CNS: an overview. Virol J 10:358.
61. Kuzembayeva M, Dilley K, Sardo L, Hu WS. 2014. Life of psi: how full-length HIV-1 RNAs become packaged genomes in the viral particles. Virology 454-455:362-370.
62. Blissenbach M, Grewe B, Hoffmann B, Brandt S, Uberla K. 2010. Nuclear RNA export and packaging functions of HIV-1 Rev revisited. J Virol 84:6598-6604.
63. Wang Y, Kinlock BL, Shao Q, Turner TM, Liu B. 2014. HIV-1 Vif inhibits G to A hypermutations catalyzed by virus-encapsidated APOBEC3G to maintain HIV-1 infectivity. Retrovirology 11:89.
64. Feng Y, Baig TT, Love RP, Chelico L. 2014. Suppression of APOBEC3-mediated restriction of HIV-1 by Vif. Front Microbiol 5:450.
65. Sato K, Misawa N, Iwami S, Satou Y, Matsuoka M, Ishizaka Y, Ito M, Aihara K, An DS, Koyanagi Y. 2013. HIV-1 Vpr accelerates viral replication during acute

infection by exploitation of proliferating CD4+ T cells in vivo. PLoS Pathog 9:e1003812.

66. Kogan M, Rappaport J. 2011. HIV-1 accessory protein Vpr: relevance in the pathogenesis of HIV and potential for therapeutic intervention. Retrovirology 8:25.
67. Guenzel CA, Herate C, Benichou S. 2014. HIV-1 Vpr-a still "enigmatic multitasker". Front Microbiol 5:127.
68. Sato K, Misawa N, Fukuhara M, Iwami S, An DS, Ito M, Koyanagi Y. 2012. Vpu augments the initial burst phase of HIV-1 propagation and downregulates BST2 and CD4 in humanized mice. J Virol 86:5000-5013.
69. Bolduan S, Reif T, Schindler M, Schubert U. 2014. HIV-1 Vpu mediated downregulation of CD155 requires alanine residues 10, 14 and 18 of the transmembrane domain. Virology 464-465:375-384.
70. Fujii H, Ato M, Takahashi Y, Otake K, Hashimoto S, Kaji T, Tsunetsugu-Yokota Y, Fujita M, Adachi A, Nakayama T, Taniguchi M, Koyasu S, Takemori T. 2011. HIV-1 Nef impairs multiple T-cell functions in antigen-specific immune response in mice. Int Immunol 23:433-441.
71. Mwimanzi P, Markle TJ, Martin E, Ogata Y, Kuang XT, Tokunaga M, Mahiti M, Pereyra F, Miura T, Walker BD, Brumme ZL, Brockman MA, Ueno T. 2013. Attenuation of multiple Nef functions in HIV-1 elite controllers. Retrovirology 10:1.
72. Wilen CB, Tilton JC, Doms RW. 2012. Molecular mechanisms of HIV entry. Adv Exp Med Biol 726:223-242.
73. Zhu P, Liu J, Bess J, Jr., Chertova E, Lifson JD, Grisé H, Ofek GA, Taylor KA, Roux KH. 2006. Distribution and three-dimensional structure of AIDS virus envelope spikes. Nature 441:847-852.
74. Arhel N. 2010. Revisiting HIV-1 uncoating. Retrovirology 7:96.
75. Fassati A. 2012. Multiple roles of the capsid protein in the early steps of HIV-1 infection. Virus Res 170:15-24.
76. Ganser-Pornillos BK, Cheng A, Yeager M. 2007. Structure of full-length HIV-1 CA: a model for the mature capsid lattice. Cell 131:70-79.
77. Pornillos O, Ganser-Pornillos BK, Yeager M. 2011. Atomic-level modelling of the HIV capsid. Nature 469:424-427.
78. Vogt VM. 1997. Retroviral Virions and Genomes. *In* Coffin JM, Hughes SH, Varmus HE (ed), Retroviruses, Cold Spring Harbor (NY).

79. Bharat TA, Davey NE, Ulbrich P, Riches JD, de Marco A, Rumlova M, Sachse C, Ruml T, Briggs JA. 2012. Structure of the immature retroviral capsid at 8 A resolution by cryo-electron microscopy. Nature 487:385-389.
80. Briggs JA, Krausslich HG. 2011. The molecular architecture of HIV. J Mol Biol 410:491-500.
81. Mohammadi P, Desfarges S, Bartha I, Joos B, Zangger N, Munoz M, Gunthard HF, Beerenwinkel N, Telenti A, Ciuffi A. 2013. 24 hours in the life of HIV-1 in a T cell line. PLoS Pathog 9:e1003161.
82. Los Alamos National Laboratory. HIV sequence database. http://www.hiv.lanl.gov/content/sequence/HIV/MAP/landmark.html. Accessed 26 September.
83. Robinson HL. 2002. New hope for an AIDS vaccine. Nat Rev Immunol 2:239-250.
84. Engelman A, Cherepanov P. 2012. The structural biology of HIV-1: mechanistic and therapeutic insights. Nat Rev Microbiol 10:279-290.
85. Wilen CB, Tilton JC, Doms RW. 2012. HIV: cell binding and entry. Cold Spring Harb Perspect Med 2.
86. Saphire AC, Bobardt MD, Zhang Z, David G, Gallay PA. 2001. Syndecans serve as attachment receptors for human immunodeficiency virus type 1 on macrophages. J Virol 75:9187-9200.
87. Cicala C, Martinelli E, McNally JP, Goode DJ, Gopaul R, Hiatt J, Jelicic K, Kottilil S, Macleod K, O'Shea A, Patel N, Van Ryk D, Wei D, Pascuccio M, Yi L, McKinnon L, Izulla P, Kimani J, Kaul R, Fauci AS, Arthos J. 2009. The integrin alpha4beta7 forms a complex with cell-surface CD4 and defines a T-cell subset that is highly susceptible to infection by HIV-1. Proc Natl Acad Sci U S A 106:20877-20882.
88. Geijtenbeek TB, Kwon DS, Torensma R, van Vliet SJ, van Duijnhoven GC, Middel J, Cornelissen IL, Nottet HS, KewalRamani VN, Littman DR, Figdor CG, van Kooyk Y. 2000. DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. Cell 100:587-597.
89. Orloff GM, Orloff SL, Kennedy MS, Maddon PJ, McDougal JS. 1991. Penetration of CD4 T cells by HIV-1. The CD4 receptor does not internalize with HIV, and CD4-related signal transduction events are not required for entry. J Immunol 146:2578-2587.
90. Freed EO. 2001. HIV-1 replication. Somat Cell Mol Genet 26:13-33.

91. Berger EA, Doms RW, Fenyo EM, Korber BT, Littman DR, Moore JP, Sattentau QJ, Schuitemaker H, Sodroski J, Weiss RA. 1998. A new classification for HIV-1. Nature 391:240.
92. Rizzuto C, Sodroski J. 2000. Fine definition of a conserved CCR5-binding region on the human immunodeficiency virus type 1 glycoprotein 120. AIDS Res Hum Retroviruses 16:741-749.
93. Rizzuto CD, Wyatt R, Hernandez-Ramos N, Sun Y, Kwong PD, Hendrickson WA, Sodroski J. 1998. A conserved HIV gp120 glycoprotein structure involved in chemokine receptor binding. Science 280:1949-1953.
94. Tamm LK, Han X. 2000. Viral fusion peptides: a tool set to disrupt and connect biological membranes. Biosci Rep 20:501-518.
95. Chan DC, Fass D, Berger JM, Kim PS. 1997. Core structure of gp41 from the HIV envelope glycoprotein. Cell 89:263-273.
96. Weissenhorn W, Dessen A, Harrison SC, Skehel JJ, Wiley DC. 1997. Atomic structure of the ectodomain from HIV-1 gp41. Nature 387:426-430.
97. Campbell EM, Hope TJ. 2015. HIV-1 capsid: the multifaceted key player in HIV-1 infection. Nat Rev Microbiol 13:471-483.
98. Hu WS, Hughes SH. 2012. HIV-1 reverse transcription. Cold Spring Harb Perspect Med 2.
99. Rouzina I, Bruinsma R. 2015. 190 Reverse transcription driven uncoating of mature HIV capsids. J Biomol Struct Dyn 33 Suppl 1:125-126.
100. Sleiman D, Goldschmidt V, Barraud P, Marquet R, Paillart JC, Tisne C. 2012. Initiation of HIV-1 reverse transcription and functional role of nucleocapsid-mediated tRNA/viral genome interactions. Virus Res 169:324-339.
101. Gotte M, Li X, Wainberg MA. 1999. HIV-1 reverse transcription: a brief overview focused on structure-function relationships among molecules involved in initiation of the reaction. Arch Biochem Biophys 365:199-210.
102. Lanchy JM, Isel C, Keith G, Le Grice SF, Ehresmann C, Ehresmann B, Marquet R. 2000. Dynamics of the HIV-1 reverse transcription complex during initiation of DNA synthesis. J Biol Chem 275:12306-12312.
103. Arhel NJ, Souquere-Besse S, Munier S, Souque P, Guadagnini S, Rutherford S, Prevost MC, Allen TD, Charneau P. 2007. HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. EMBO J 26:3025-3037.

104. Strebel K, Luban J, Jeang KT. 2009. Human cellular restriction factors that target HIV-1 replication. BMC Med 7:48.
105. Zennou V, Petit C, Guetard D, Nerhbass U, Montagnier L, Charneau P. 2000. HIV-1 genome nuclear import is mediated by a central DNA flap. Cell 101:173-185.
106. Hulme AE, Perez O, Hope TJ. 2011. Complementary assays reveal a relationship between HIV-1 uncoating and reverse transcription. Proc Natl Acad Sci U S A 108:9975-9980.
107. McDonald D, Vodicka MA, Lucero G, Svitkina TM, Borisy GG, Emerman M, Hope TJ. 2002. Visualization of the intracellular behavior of HIV in living cells. J Cell Biol 159:441-452.
108. Hulme AE, Kelley Z, Foley D, Hope TJ. 2015. Complementary Assays Reveal a Low Level of CA Associated with Viral Complexes in the Nuclei of HIV-1-Infected Cells. J Virol 89:5350-5361.
109. Peng K, Muranyi W, Glass B, Laketa V, Yant SR, Tsai L, Cihlar T, Muller B, Krausslich HG. 2014. Quantitative microscopy of functional HIV post-entry complexes reveals association of replication with the viral capsid. Elife 3:e04114.
110. Matreyek KA, Yucel SS, Li X, Engelman A. 2013. Nucleoporin NUP153 phenylalanine-glycine motifs engage a common binding pocket within the HIV-1 capsid protein to mediate lentiviral infectivity. PLoS Pathog 9:e1003693.
111. Zhou L, Sokolskaja E, Jolly C, James W, Cowley SA, Fassati A. 2011. Transportin 3 promotes a nuclear maturation step required for efficient HIV-1 integration. PLoS Pathog 7:e1002194.
112. Le Sage V, Mouland AJ, Valiente-Echeverria F. 2014. Roles of HIV-1 capsid in viral replication and immune evasion. Virus Res 193:116-129.
113. Hatziioannou T, Perez-Caballero D, Cowan S, Bieniasz PD. 2005. Cyclophilin interactions with incoming human immunodeficiency virus type 1 capsids with opposing effects on infectivity in human cells. J Virol 79:176-183.
114. Bosco DA, Eisenmesser EZ, Pochapsky S, Sundquist WI, Kern D. 2002. Catalysis of cis/trans isomerization in native HIV-1 capsid by human cyclophilin A. Proc Natl Acad Sci U S A 99:5247-5252.
115. Bichel K, Price AJ, Schaller T, Towers GJ, Freund SM, James LC. 2013. HIV-1 capsid undergoes coupled binding and isomerization by the nuclear pore protein NUP358. Retrovirology 10:81.
116. Lukic Z, Dharan A, Fricke T, Diaz-Griffero F, Campbell EM. 2014. HIV-1 uncoating is facilitated by dynein and kinesin 1. J Virol 88:13613-13625.

117. Pawlica P, Berthoux L. 2014. Cytoplasmic dynein promotes HIV-1 uncoating. Viruses 6:4195-4211.
118. Labokha AA, Fassati A. 2013. Viruses challenge selectivity barrier of nuclear pores. Viruses 5:2410-2423.
119. Strambio-De-Castillia C, Niepel M, Rout MP. 2010. The nuclear pore complex: bridging nuclear transport and gene regulation. Nat Rev Mol Cell Biol 11:490-501.
120. Schaller T, Ocwieja KE, Rasaiyaah J, Price AJ, Brady TL, Roth SL, Hue S, Fletcher AJ, Lee K, KewalRamani VN, Noursadeghi M, Jenner RG, James LC, Bushman FD, Towers GJ. 2011. HIV-1 capsid-cyclophilin interactions determine nuclear import pathway, integration targeting and replication efficiency. PLoS Pathog 7:e1002439.
121. Yamashita M, Emerman M. 2006. Retroviral infection of non-dividing cells: old and new perspectives. Virology 344:88-93.
122. Yamashita M, Emerman M. 2004. Capsid is a dominant determinant of retrovirus infectivity in nondividing cells. J Virol 78:5670-5678.
123. Yamashita M, Perez O, Hope TJ, Emerman M. 2007. Evidence for direct involvement of the capsid protein in HIV infection of nondividing cells. PLoS Pathog 3:1502-1510.
124. Bushman FD, Malani N, Fernandes J, D'Orso I, Cagney G, Diamond TL, Zhou H, Hazuda DJ, Espeseth AS, Konig R, Bandyopadhyay S, Ideker T, Goff SP, Krogan NJ, Frankel AD, Young JA, Chanda SK. 2009. Host cell factors in HIV replication: meta-analysis of genome-wide studies. PLoS Pathog 5:e1000437.
125. Konig R, Zhou Y, Elleder D, Diamond TL, Bonamy GM, Irelan JT, Chiang CY, Tu BP, De Jesus PD, Lilley CE, Seidel S, Opaluch AM, Caldwell JS, Weitzman MD, Kuhen KL, Bandyopadhyay S, Ideker T, Orth AP, Miraglia LJ, Bushman FD, Young JA, Chanda SK. 2008. Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. Cell 135:49-60.
126. Brass AL, Dykxhoorn DM, Benita Y, Yan N, Engelman A, Xavier RJ, Lieberman J, Elledge SJ. 2008. Identification of host proteins required for HIV infection through a functional genomic screen. Science 319:921-926.
127. Zhao G, Perilla JR, Yufenyuy EL, Meng X, Chen B, Ning J, Ahn J, Gronenborn AM, Schulten K, Aiken C, Zhang P. 2013. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. Nature 497:643-646.

128. Price AJ, Fletcher AJ, Schaller T, Elliott T, Lee K, KewalRamani VN, Chin JW, Towers GJ, James LC. 2012. CPSF6 defines a conserved capsid interface that modulates HIV-1 replication. PLoS Pathog 8:e1002896.
129. Kataoka N, Bachorik JL, Dreyfuss G. 1999. Transportin-SR, a nuclear import receptor for SR proteins. J Cell Biol 145:1145-1152.
130. Di Nunzio F, Danckaert A, Fricke T, Perez P, Fernandez J, Perret E, Roux P, Shorte S, Charneau P, Diaz-Griffero F, Arhel NJ. 2012. Human nucleoporins promote HIV-1 docking at the nuclear pore, nuclear import and integration. PLoS One 7:e46037.
131. Di Nunzio F, Fricke T, Miccio A, Valle-Casuso JC, Perez P, Souque P, Rizzi E, Severgnini M, Mavilio F, Charneau P, Diaz-Griffero F. 2013. Nup153 and Nup98 bind the HIV-1 core and contribute to the early steps of HIV-1 replication. Virology 440:8-18.
132. Krishnan L, Engelman A. 2012. Retroviral integrase proteins and HIV-1 DNA integration. J Biol Chem 287:40858-40866.
133. Fujiwara T, Mizuuchi K. 1988. Retroviral DNA integration: structure of an integration intermediate. Cell 54:497-504.
134. Engelman A, Mizuuchi K, Craigie R. 1991. HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. Cell 67:1211-1221.
135. Nabel G, Baltimore D. 1987. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. Nature 326:711-713.
136. Starcich B, Ratner L, Josephs SF, Okamoto T, Gallo RC, Wong-Staal F. 1985. Characterization of long terminal repeat sequences of HTLV-III. Science 227:538-540.
137. Karn J, Stoltzfus CM. 2012. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. Cold Spring Harb Perspect Med 2:a006916.
138. Suzuki K, Ahlenstiel C, Marks K, Kelleher AD. 2015. Promoter Targeting RNAs: Unexpected Contributors to the Control of HIV-1 Transcription. Mol Ther Nucleic Acids 4:e222.
139. Kao SY, Calman AF, Luciw PA, Peterlin BM. 1987. Anti-termination of transcription within the long terminal repeat of HIV-1 by Tat gene product. Nature 330:489-493.
140. Bieniasz PD, Grdina TA, Bogerd HP, Cullen BR. 1998. Recruitment of a protein complex containing Tat and cyclin T1 to TAR governs the species specificity of HIV-1 Tat. EMBO J 17:7056-7065.

141. Fujinaga K, Cujec TP, Peng J, Garriga J, Price DH, Grana X, Peterlin BM. 1998. The ability of positive transcription elongation factor B to transactivate human immunodeficiency virus transcription depends on a functional kinase domain, cyclin T1, and Tat. J Virol 72:7154-7159.
142. Zhou M, Halanski MA, Radonovich MF, Kashanchi F, Peng J, Price DH, Brady JN. 2000. Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during human immunodeficiency virus type 1 transcription. Mol Cell Biol 20:5077-5086.
143. Freed EO. 2015. HIV-1 assembly, release and maturation. Nat Rev Microbiol 13:484-496.
144. Ono A. 2009. HIV-1 Assembly at the Plasma Membrane: Gag Trafficking and Localization. Future Virol 4:241-257.
145. Bouamr F, Scarlata S, Carter C. 2003. Role of myristylation in HIV-1 Gag assembly. Biochemistry 42:6408-6417.
146. Briggs JA, Riches JD, Glass B, Bartonova V, Zanetti G, Krausslich HG. 2009. Structure and assembly of immature HIV. Proc Natl Acad Sci U S A 106:11090-11095.
147. Wright ER, Schooler JB, Ding HJ, Kieffer C, Fillmore C, Sundquist WI, Jensen GJ. 2007. Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. EMBO J 26:2218-2226.
148. Fuller SD, Wilk T, Gowen BE, Krausslich HG, Vogt VM. 1997. Cryo-electron microscopy reveals ordered domains in the immature HIV-1 particle. Curr Biol 7:729-738.
149. Kutluay SB, Bieniasz PD. 2010. Analysis of the initiating events in HIV-1 particle assembly and genome packaging. PLoS Pathog 6:e1001200.
150. Kutluay SB, Zang T, Blanco-Melo D, Powell C, Jannain D, Errando M, Bieniasz PD. 2014. Global changes in the RNA binding specificity of HIV-1 Gag regulate virion genesis. Cell 159:1096-1109.
151. Lehmann M, Milev MP, Abrahamyan L, Yao XJ, Pante N, Mouland AJ. 2009. Intracellular transport of human immunodeficiency virus type 1 genomic RNA and viral production are dependent on dynein motor function and late endosome positioning. J Biol Chem 284:14572-14585.
152. Molle D, Segura-Morales C, Camus G, Berlioz-Torrent C, Kjems J, Basyuk E, Bertrand E. 2009. Endosomal trafficking of HIV-1 Gag and genomic RNAs regulates viral egress. J Biol Chem 284:19727-19743.

153. Scarlata S, Carter C. 2003. Role of HIV-1 Gag domains in viral assembly. Biochim Biophys Acta 1614:62-72.
154. Mateu MG. 2009. The capsid protein of human immunodeficiency virus: intersubunit interactions during virus assembly. FEBS J 276:6098-6109.
155. Bierman WF, van Agtmael MA, Nijhuis M, Danner SA, Boucher CA. 2009. HIV monotherapy with ritonavir-boosted Protease inhibitors: a systematic review. AIDS 23:279-291.
156. Leung K, Kim JO, Ganesh L, Kabat J, Schwartz O, Nabel GJ. 2008. HIV-1 assembly: viral glycoproteins segregate quantally to lipid rafts that associate individually with HIV-1 capsids and virions. Cell Host Microbe 3:285-292.
157. Dorfman T, Mammano F, Haseltine WA, Gottlinger HG. 1994. Role of the matrix protein in the virion association of the human immunodeficiency virus type 1 envelope glycoprotein. J Virol 68:1689-1696.
158. Freed EO, Martin MA. 1995. Virion incorporation of envelope glycoproteins with long but not short cytoplasmic tails is blocked by specific, single amino acid substitutions in the human immunodeficiency virus type 1 matrix. J Virol 69:1984-1989.
159. Freed EO, Martin MA. 1996. Domains of the human immunodeficiency virus type 1 matrix and gp41 cytoplasmic tail required for envelope incorporation into virions. J Virol 70:341-351.
160. Ono A, Huang M, Freed EO. 1997. Characterization of human immunodeficiency virus type 1 matrix revertants: effects on virus assembly, Gag processing, and Env incorporation into virions. J Virol 71:4409-4418.
161. Bell NM, Lever AM. 2013. HIV Gag polyprotein: processing and early viral particle assembly. Trends Microbiol 21:136-144.
162. Lingappa JR, Reed JC, Tanaka M, Chutiraka K, Robinson BA. 2014. How HIV-1 Gag assembles in cells: Putting together pieces of the puzzle. Virus Res 193:89-107.
163. Fujii K, Munshi UM, Ablan SD, Demirov DG, Soheilian F, Nagashima K, Stephen AG, Fisher RJ, Freed EO. 2009. Functional role of Alix in HIV-1 replication. Virology 391:284-292.
164. Shen QT, Schuh AL, Zheng Y, Quinney K, Wang L, Hanna M, Mitchell JC, Otegui MS, Ahlquist P, Cui Q, Audhya A. 2014. Structural analysis and modeling reveals new mechanisms governing ESCRT-III spiral filament assembly. J Cell Biol 206:763-777.

165. Hanson PI, Roth R, Lin Y, Heuser JE. 2008. Plasma membrane deformation by circular arrays of ESCRT-III protein filaments. J Cell Biol 180:389-402.
166. Wlodawer A, Erickson JW. 1993. Structure-based inhibitors of HIV-1 Protease. Annu Rev Biochem 62:543-585.
167. Pettit SC, Moody MD, Wehbie RS, Kaplan AH, Nantermet PV, Klein CA, Swanstrom R. 1994. The p2 domain of human immunodeficiency virus type 1 Gag regulates sequential proteolytic processing and is required to produce fully infectious virions. J Virol 68:8017-8027.
168. Sundquist WI, Krausslich HG. 2012. HIV-1 assembly, budding, and maturation. Cold Spring Harb Perspect Med 2:a006924.
169. Yang H, Nkeze J, Zhao RY. 2012. Effects of HIV-1 Protease on cellular functions and their potential applications in antiretroviral therapy. Cell Biosci 2:32.
170. Lv Z, Chu Y, Wang Y. 2015. HIV Protease inhibitors: a review of molecular selectivity and toxicity. HIV AIDS (Auckl) 7:95-104.
171. Toth G, Borics A. 2006. Closing of the flaps of HIV-1 Protease induced by substrate binding: a model of a flap closing mechanism in retroviral aspartic Proteases. Biochemistry 45:6606-6614.
172. Toth G, Borics A. 2006. Flap opening mechanism of HIV-1 Protease. J Mol Graph Model 24:465-474.
173. Konvalinka J, Krausslich HG, Muller B. 2015. Retroviral Proteases and their roles in virion maturation. Virology 479-480:403-417.
174. Ozer N, Ozen A, Schiffer CA, Haliloglu T. 2015. Drug-resistant HIV-1 Protease regains functional dynamics through cleavage site coevolution. Evol Appl 8:185-198.
175. Iyidogan P, Anderson KS. 2014. Current perspectives on HIV-1 antiretroviral drug resistance. Viruses 6:4095-4139.
176. Prabu-Jeyabalan M, Nalivaika E, Schiffer CA. 2002. Substrate shape determines specificity of recognition for HIV-1 Protease: analysis of crystal structures of six substrate complexes. Structure 10:369-381.
177. Prabu-Jeyabalan M, Nalivaika EA, Romano K, Schiffer CA. 2006. Mechanism of substrate recognition by drug-resistant human immunodeficiency virus type 1 Protease variants revealed by a novel structural intermediate. J Virol 80:3607-3616.

178. Potempa M, Nalivaika E, Ragland D, Lee SK, Schiffer CA, Swanstrom R. 2015. A Direct Interaction with RNA Dramatically Enhances the Catalytic Activity of the HIV-1 Protease In Vitro. J Mol Biol 427:2360-2378.
179. Pettit SC, Everitt LE, Choudhury S, Dunn BM, Kaplan AH. 2004. Initial cleavage of the human immunodeficiency virus type 1 GagPol precursor by its activated Protease occurs by an intramolecular mechanism. J Virol 78:8477-8485.
180. Tessmer U, Krausslich HG. 1998. Cleavage of human immunodeficiency virus type 1 proteinase from the N-terminally adjacent p6* protein is essential for efficient Gag polyprotein processing and viral infectivity. J Virol 72:3459-3463.
181. Tozser J, Yin FH, Cheng YS, Bagossi P, Weber IT, Harrison RW, Oroszlan S. 1997. Activity of tethered human immunodeficiency virus 1 Protease containing mutations in the flap region of one subunit. Eur J Biochem 244:235-241.
182. Fun A, Wensing AM, Verheyen J, Nijhuis M. 2012. Human Immunodeficiency Virus Gag and Protease: partners in resistance. Retrovirology 9:63.
183. Brockman MA, Tanzi GO, Walker BD, Allen TM. 2006. Use of a novel GFP reporter cell line to examine replication capacity of CXCR4- and CCR5-tropic HIV-1 by flow cytometry. J Virol Methods 131:134-142.
184. Blundell TL, Jhoti H, Abell C. 2002. High-throughput crystallography for lead discovery in drug design. Nat Rev Drug Discov 1:45-54.
185. de Oliveira T, Engelbrecht S, Janse van Rensburg E, Gordon M, Bishop K, zur Megede J, Barnett SW, Cassol S. 2003. Variability at Human Immunodeficiency Virus Type 1 Subtype C Protease Cleavage Sites: an Indication of Viral Fitness? Journal of Virology 77:9422-9430.
186. Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, Sun C, Grayson T, Wang S, Li H, Wei X, Jiang C, Kirchherr JL, Gao F, Anderson JA, Ping LH, Swanstrom R, Tomaras GD, Blattner WA, Goepfert PA, Kilby JM, Saag MS, Delwart EL, Busch MP, Cohen MS, Montefiori DC, Haynes BF, Gaschen B, Athreya GS, Lee HY, Wood N, Seoighe C, Perelson AS, Bhattacharya T, Korber BT, Hahn BH, Shaw GM. 2008. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. Proc Natl Acad Sci U S A 105:7552-7557.
187. Cohen MS, Shaw GM, McMichael AJ, Haynes BF. 2011. Acute HIV-1 Infection. N Engl J Med 364:1943-1954.

188. McMichael AJ, Borrow P, Tomaras GD, Goonetilleke N, Haynes BF. 2010. The immune response during acute HIV-1 infection: clues for vaccine development. Nat Rev Immunol 10:11-23.
189. Burgener A, McGowan I, Klatt NR. 2015. HIV and mucosal barrier interactions: consequences for transmission and pathogenesis. Curr Opin Immunol 36:22-30.
190. Suthar AB, Granich RM, Kato M, Nsanzimana S, Montaner JS, Williams BG. 2015. Programmatic implications of acute and early HIV infection. J Infect Dis doi:10.1093/infdis/jiv430.
191. Martinez-Bonet M, Puertas MC, Fortuny C, Ouchi D, Mellado MJ, Rojo P, Noguera-Julian A, Munoz-Fernandez MA, Martinez-Picado J. 2015. Establishment and Replenishment of the Viral Reservoir in Perinatally HIV-1-infected Children Initiating Very Early Antiretroviral Therapy. Clin Infect Dis 61:1169-1178.
192. Chun TW, Engel D, Berrey MM, Shea T, Corey L, Fauci AS. 1998. Early establishment of a pool of latently infected, resting CD4(+) T cells during primary HIV-1 infection. Proc Natl Acad Sci U S A 95:8869-8873.
193. Chun TW, Fauci AS. 2012. HIV reservoirs: pathogenesis and obstacles to viral eradication and cure. AIDS 26:1261-1268.
194. Brenchley JM, Schacker TW, Ruff LE, Price DA, Taylor JH, Beilman GJ, Nguyen PL, Khoruts A, Larson M, Haase AT, Douek DC. 2004. CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. J Exp Med 200:749-759.
195. Brenchley JM, Paiardini M, Knox KS, Asher AI, Cervasi B, Asher TE, Scheinberg P, Price DA, Hage CA, Kholi LM, Khoruts A, Frank I, Else J, Schacker T, Silvestri G, Douek DC. 2008. Differential Th17 CD4 T-cell depletion in pathogenic and nonpathogenic lentiviral infections. Blood 112:2826-2835.
196. Stacey AR, Norris PJ, Qin L, Haygreen EA, Taylor E, Heitman J, Lebedeva M, DeCamp A, Li D, Grove D, Self SG, Borrow P. 2009. Induction of a striking systemic cytokine cascade prior to peak viremia in acute human immunodeficiency virus type 1 infection, in contrast to more modest and delayed responses in acute hepatitis B and C virus infections. J Virol 83:3719-3733.
197. Goonetilleke N, Liu MK, Salazar-Gonzalez JF, Ferrari G, Giorgi E, Ganusov VV, Keele BF, Learn GH, Turnbull EL, Salazar MG, Weinhold KJ, Moore S, Letvin N, Haynes BF, Cohen MS, Hraber P, Bhattacharya T, Borrow P, Perelson AS, Hahn BH, Shaw GM, Korber BT, McMichael AJ. 2009. The first T cell response to

transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection. J Exp Med 206:1253-1272.

198. Trautmann L, Janbazian L, Chomont N, Said EA, Gimmig S, Bessette B, Boulassel MR, Delwart E, Sepulveda H, Balderas RS, Routy JP, Haddad EK, Sekaly RP. 2006. Upregulation of PD-1 expression on HIV-specific CD8+ T cells leads to reversible immune dysfunction. Nat Med 12:1198-1202.

199. Mattapallil JJ, Douek DC, Hill B, Nishimura Y, Martin M, Roederer M. 2005. Massive infection and loss of memory CD4+ T cells in multiple tissues during acute SIV infection. Nature 434:1093-1097.

200. Streeck H, Nixon DF. 2010. T cell immunity in acute HIV-1 infection. J Infect Dis 202 Suppl 2:S302-308.

201. Levy JA. 1993. Pathogenesis of human immunodeficiency virus infection. Microbiol Rev 57:183-289.

202. Fiebig EW, Wright DJ, Rawal BD, Garrett PE, Schumacher RT, Peddada L, Heldebrant C, Smith R, Conrad A, Kleinman SH, Busch MP. 2003. Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. AIDS 17:1871-1879.

203. Marchetti G, Tincati C, Silvestri G. 2013. Microbial translocation in the pathogenesis of HIV infection and AIDS. Clin Microbiol Rev 26:2-18.

204. Cao J, Park IW, Cooper A, Sodroski J. 1996. Molecular determinants of acute single-cell lysis by human immunodeficiency virus type 1. J Virol 70:1340-1354.

205. Lenardo MJ, Angleman SB, Bounkeua V, Dimas J, Duvall MG, Graubard MB, Hornung F, Selkirk MC, Speirs CK, Trageser C, Orenstein JO, Bolton DL. 2002. Cytopathic killing of peripheral blood CD4(+) T lymphocytes by human immunodeficiency virus type 1 appears necrotic rather than apoptotic and does not require Env. J Virol 76:5082-5093.

206. McMichael AJ, Rowland-Jones SL. 2001. Cellular immune responses to HIV. Nature 410:980-987.

207. Appay V, Lissina A, Sauce D. 2014. HIV-Associated Immune Exhaustion, p 1-8. *In* Hope TJ, Stevenson M, Richman D (ed), Encyclopedia of AIDS doi:10.1007/978-1-4614-9610-6_186-1. Springer New York.

208. Appay V, Sauce D. 2008. Immune activation and inflammation in HIV-1 infection: causes and consequences. J Pathol 214:231-241.

209. Cadogan M, Dalgleish AG. 2008. Pathogenesis of HIV: non-specific immune hyperactivity and its implications for vaccines. Clin Med 8:267-271.

210. Sodora DL, Allan JS, Apetrei C, Brenchley JM, Douek DC, Else JG, Estes JD, Hahn BH, Hirsch VM, Kaur A, Kirchhoff F, Muller-Trutwin M, Pandrea I, Schmitz JE, Silvestri G. 2009. Toward an AIDS vaccine: lessons from natural simian immunodeficiency virus infections of African nonhuman primate hosts. Nat Med 15:861-865.
211. Paiardini M, Cervasi B, Reyes-Aviles E, Micci L, Ortiz AM, Chahroudi A, Vinton C, Gordon SN, Bosinger SE, Francella N, Hallberg PL, Cramer E, Schlub T, Chan ML, Riddick NE, Collman RG, Apetrei C, Pandrea I, Else J, Munch J, Kirchhoff F, Davenport MP, Brenchley JM, Silvestri G. 2011. Low levels of SIV infection in sooty mangabey central memory CD(4)(+) T cells are associated with limited CCR5 expression. Nat Med 17:830-836.
212. Collman RG, Perno CF, Crowe SM, Stevenson M, Montaner LJ. 2003. HIV and cells of macrophage/dendritic lineage and other non-T cell reservoirs: new answers yield new questions. J Leukoc Biol 74:631-634.
213. Langford SE, Ananworanich J, Cooper DA. 2007. Predictors of disease progression in HIV infection: a review. AIDS Res Ther 4:11.
214. Kuritzkes DR, Walker BD. 2007. HIV-1: Pathogenesis, Clinical Manifestations, and Treatment, p 2187-2214. *In* Fields BN, Knipe DM, Howley PM (ed), Fields virology *5th edition*. Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia.
215. Yue L, Prentice HA, Farmer P, Song W, He D, Lakhi S, Goepfert P, Gilmour J, Allen S, Tang J, Kaslow RA, Hunter E. 2013. Cumulative impact of host and viral factors on HIV-1 viral-load control during early infection. J Virol 87:708-715.
216. Perlmutter BL, Glaser JB, Oyugi SO. 1999. How to recognize and treat acute HIV syndrome. Am Fam Physician 60:535-542, 545-536.
217. Thompson MA, Aberg JA, Hoy JF, Telenti A, Benson C, Cahn P, Eron JJ, Gunthard HF, Hammer SM, Reiss P, Richman DD, Rizzardini G, Thomas DL, Jacobsen DM, Volberding PA. 2012. Antiretroviral treatment of adult HIV infection: 2012 recommendations of the International Antiviral Society-USA panel. JAMA 308:387-402.
218. National Department of Health. December 2014 2014. National consolidated guidelines for the prevention of mother to child transmission (PMTCT) of HIV and the management of HIV in children, adolescents and adults, *on* National Department of Health http://www.sahivsoc.org/upload/documents/HIV%20guidelines%20_Jan%202015.pdf. Accessed 27 April

219. Lederman MM, Valdez H. 1998. Cytotoxic T-cell activity and progression of HIV disease: abstract and commentary. JAMA 279:242.
220. Komanduri KV, Viswanathan MN, Wieder ED, Schmidt DK, Bredt BM, Jacobson MA, McCune JM. 1998. Restoration of cytomegalovirus-specific CD4+ T-lymphocyte responses after ganciclovir and highly active antiretroviral therapy in individuals infected with HIV-1. Nat Med 4:953-956.
221. Broder S. 2010. The development of antiretroviral therapy and its impact on the HIV-1/AIDS pandemic. Antiviral Res 85:1-18.
222. Broder S. 2010. Twenty-five years of translational medicine in antiretroviral therapy: promises to keep. Sci Transl Med 2:39ps33.
223. Larder BA, Darby G, Richman DD. 1989. HIV with reduced sensitivity to zidovudine (AZT) isolated during prolonged therapy. Science 243:1731-1734.
224. Fischl MA, Richman DD, Grieco MH, Gottlieb MS, Volberding PA, Laskin OL, Leedom JM, Groopman JE, Mildvan D, Schooley RT, et al. 1987. The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. A double-blind, placebo-controlled trial. N Engl J Med 317:185-191.
225. Hammer SM, Katzenstein DA, Hughes MD, Gundacker H, Schooley RT, Haubrich RH, Henry WK, Lederman MM, Phair JP, Niu M, Hirsch MS, Merigan TC. 1996. A trial comparing nucleoside monotherapy with combination therapy in HIV-infected adults with CD4 cell counts from 200 to 500 per cubic millimeter. AIDS Clinical Trials Group Study 175 Study Team. N Engl J Med 335:1081-1090.
226. Laskey SB, Siliciano RF. 2014. A mechanistic theory to explain the efficacy of antiretroviral therapy. Nat Rev Microbiol 12:772-780.
227. Jilek BL, Zarr M, Sampah ME, Rabi SA, Bullen CK, Lai J, Shen L, Siliciano RF. 2012. A quantitative basis for antiretroviral therapy for HIV-1 infection. Nat Med 18:446-451.
228. Arts EJ, Hazuda DJ. 2012. HIV-1 antiretroviral drug therapy. Cold Spring Harb Perspect Med 2:a007161.
229. Levison JH, Orrell C, Gallien S, Kuritzkes DR, Fu N, Losina E, Freedberg KA, Wood R. 2012. Virologic failure of Protease inhibitor-based second-line antiretroviral therapy without resistance in a large HIV treatment program in South Africa. PLoS One 7:e32144.
230. World Health Organisation. 2010. Antiretoviral Drug Resistance. http://www.who.int/hiv/topics/drugresistance/en/index.html. Accessed 08 April.

231. World Health Organisation. 2012. Drug Resistance Report www.who.int/hiv/pub/drugresistance/report2012/en/index.html Accessed 13 March
232. Coffin JM. 1995 <HIV population dynamics *in vivo*: implications for genetic variation pathogenesis and therapy Science 267:483-489.
233. Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. 1996. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. Science 271:1582-1586.
234. Li JY, Li HP, Li L, Li H, Wang Z, Yang K, Bao ZY, Zhuang DM, Liu SY, Liu YJ, Xing H, Shao YM. 2005. Prevalence and evolution of drug resistance HIV-1 variants in Henan, China. Cell Res 15:843-849.
235. Freeman S, Herron JC. 2007. A case for evolutionary thinking: Understanding HIV, Evolutionary Analysis, 4th ed. Pearson Benjamin Cummings, San Francisco.
236. Kozal MJ. 2009. Drug Resistant Human Immunodeficiency Virus. Clinical Microbiology Infection 15:69-73.
237. Clavel F, Hance AJ. 2004. HIV drug resistance. N Engl J Med 350:1023-1035.
238. Daar ES, Richman DD. 2005. Confronting the emergence of drug-resistant HIV type 1: impact of antiretroviral therapy on individual and population resistance. AIDS Res Hum Retroviruses 21:343-357.
239. Johnson VA, Brun-Vezinet F, Clotet B, Gunthard HF, Kuritzkes DR, Pillay D, Schapiro JM, Richman DD. 2007. Update of the drug resistance mutations in HIV-1: 2007. Top HIV Med 15:119-125.
240. Davey RT, Jr., Bhat N, Yoder C, Chun TW, Metcalf JA, Dewar R, Natarajan V, Lempicki RA, Adelsberger JW, Miller KD, Kovacs JA, Polis MA, Walker RE, Falloon J, Masur H, Gee D, Baseler M, Dimitrov DS, Fauci AS, Lane HC. 1999. HIV-1 and T cell dynamics after interruption of highly active antiretroviral therapy (HAART) in patients with a history of sustained viral suppression. Proc Natl Acad Sci U S A 96:15109-15114.
241. Nwobegahay J, Selabe G, Ndjeka ON, Manhaeve C, Bessong PO. 2012. Low prevalence of transmitted genetic drug resistance in a cohort of HIV infected Naive patients entering Antiretroviral treatment programs at two sites in Norther South Africa. Journal of Medical Virology 84:1839-1843.
242. Parboosing R, Naidoo A, Gordon M, Taylor M, Vella V. 2011. Resistance to antiretroviral drugs in newly diagnosed, young treatment-naive HIV-positive pregnant women in the province of KwaZulu-Natal, South Africa. J Med Virol 83:1508-1513.

243. Bennett DE, Myatt M, Bertagnolio S, Sutherland D, Gilks CF. 2008. Recommendations for surveillance of transmitted HIV drug resistance in countries scaling up antiretroviral treatment. Antivir Ther 13 Suppl 2:25-36.
244. De Feo CJ, Weiss CD. 2012. Escape from human immunodeficiency virus type 1 (HIV-1) entry inhibitors. Viruses 4:3859-3911.
245. Dragic T, Trkola A, Thompson DA, Cormier EG, Kajumo FA, Maxwell E, Lin SW, Ying W, Smith SO, Sakmar TP, Moore JP. 2000. A binding pocket for a small molecule inhibitor of HIV-1 entry within the transmembrane helices of CCR5. Proc Natl Acad Sci U S A 97:5639-5644.
246. Roche M, Jakobsen MR, Sterjovski J, Ellett A, Posta F, Lee B, Jubb B, Westby M, Lewin SR, Ramsland PA, Churchill MJ, Gorry PR. 2011. HIV-1 escape from the CCR5 antagonist maraviroc associated with an altered and less-efficient mechanism of gp120-CCR5 engagement that attenuates macrophage tropism. J Virol 85:4330-4342.
247. Moore JP, Kuritzkes DR. 2009. A piece de resistance: how HIV-1 escapes small molecule CCR5 inhibitors. Curr Opin HIV AIDS 4:118-124.
248. Westby M, Smith-Burchnell C, Mori J, Lewis M, Mosley M, Stockdale M, Dorr P, Ciaramella G, Perros M. 2007. Reduced maximal inhibition in phenotypic susceptibility assays indicates that viral strains resistant to the CCR5 antagonist maraviroc utilize inhibitor-bound receptor for entry. J Virol 81:2359-2371.
249. Yuan Y, Maeda Y, Terasawa H, Monde K, Harada S, Yusa K. 2011. A combination of polymorphic mutations in V3 loop of HIV-1 gp120 can confer noncompetitive resistance to maraviroc. Virology 413:293-299.
250. Berro R, Klasse PJ, Jakobsen MR, Gorry PR, Moore JP, Sanders RW. 2012. V3 determinants of HIV-1 escape from the CCR5 inhibitors Maraviroc and Vicriviroc. Virology 427:158-165.
251. Anastassopoulou CG, Ketas TJ, Klasse PJ, Moore JP. 2009. Resistance to CCR5 inhibitors caused by sequence changes in the fusion peptide of HIV-1 gp41. Proc Natl Acad Sci U S A 106:5318-5323.
252. Kilby JM, Hopkins S, Venetta TM, DiMassimo B, Cloud GA, Lee JY, Alldredge L, Hunter E, Lambert D, Bolognesi D, Matthews T, Johnson MR, Nowak MA, Shaw GM, Saag MS. 1998. Potent suppression of HIV-1 replication in humans by T-20, a peptide inhibitor of gp41-mediated virus entry. Nat Med 4:1302-1307.

253. Matthews T, Salgo M, Greenberg M, Chung J, DeMasi R, Bolognesi D. 2004. Enfuvirtide: the first therapy to inhibit the entry of HIV-1 into host CD4 lymphocytes. Nat Rev Drug Discov 3:215-225.
254. Rimsky LT, Shugars DC, Matthews TJ. 1998. Determinants of human immunodeficiency virus type 1 resistance to gp41-derived inhibitory peptides. J Virol 72:986-993.
255. Mink M, Mosier SM, Janumpalli S, Davison D, Jin L, Melby T, Sista P, Erickson J, Lambert D, Stanfield-Oakley SA, Salgo M, Cammack N, Matthews T, Greenberg ML. 2005. Impact of human immunodeficiency virus type 1 gp41 amino acid substitutions selected during enfuvirtide treatment on gp41 binding and antiviral potency of enfuvirtide in vitro. J Virol 79:12447-12454.
256. Menzo S, Castagna A, Monachetti A, Hasson H, Danise A, Carini E, Bagnarelli P, Lazzarin A, Clementi M. 2004. Genotype and phenotype patterns of human immunodeficiency virus type 1 resistance to enfuvirtide during long-term treatment. Antimicrob Agents Chemother 48:3253-3259.
257. Nameki D, Kodama E, Ikeuchi M, Mabuchi N, Otaka A, Tamamura H, Ohno M, Fujii N, Matsuoka M. 2005. Mutations conferring resistance to human immunodeficiency virus type 1 fusion inhibitors are restricted by gp41 and Rev-responsive element functions. J Virol 79:764-770.
258. Greenberg ML, Cammack N. 2004. Resistance to enfuvirtide, the first HIV fusion inhibitor. J Antimicrob Chemother 54:333-340.
259. Metifiot M, Marchand C, Pommier Y. 2013. HIV integrase inhibitors: 20-year landmark and challenges. Adv Pharmacol 67:75-105.
260. Hare S, Vos AM, Clayton RF, Thuring JW, Cummings MD, Cherepanov P. 2010. Molecular mechanisms of retroviral integrase inhibition and the evolution of viral resistance. Proc Natl Acad Sci U S A 107:20057-20062.
261. Hazuda DJ, Anthony NJ, Gomez RP, Jolly SM, Wai JS, Zhuang L, Fisher TE, Embrey M, Guare JP, Jr., Egbertson MS, Vacca JP, Huff JR, Felock PJ, Witmer MV, Stillmock KA, Danovich R, Grobler J, Miller MD, Espeseth AS, Jin L, Chen IW, Lin JH, Kassahun K, Ellis JD, Wong BK, Xu W, Pearson PG, Schleif WA, Cortese R, Emini E, Summa V, Holloway MK, Young SD. 2004. A naphthyridine carboxamide provides evidence for discordant resistance between mechanistically identical inhibitors of HIV-1 integrase. Proc Natl Acad Sci U S A 101:11233-11238.
262. Fransen S, Gupta S, Danovich R, Hazuda D, Miller M, Witmer M, Petropoulos CJ, Huang W. 2009. Loss of raltegravir susceptibility by human immunodeficiency virus

type 1 is conferred via multiple nonoverlapping genetic pathways. J Virol 83:11440-11446.

263. Hicks C, Gulick RM. 2009. Raltegravir: the first HIV type 1 integrase inhibitor. Clin Infect Dis 48:931-939.

264. Hurt CB, Sebastian J, Hicks CB, Eron JJ. 2014. Resistance to HIV integrase strand transfer inhibitors among clinical specimens in the United States, 2009-2012. Clin Infect Dis 58:423-431.

265. Canducci F, Sampaolo M, Marinozzi MC, Boeri E, Spagnuolo V, Galli A, Castagna A, Lazzarin A, Clementi M, Gianotti N. 2009. Dynamic patterns of human immunodeficiency virus type 1 integrase gene evolution in patients failing raltegravir-based salvage therapies. AIDS 23:455-460.

266. Cooper DA, Steigbigel RT, Gatell JM, Rockstroh JK, Katlama C, Yeni P, Lazzarin A, Clotet B, Kumar PN, Eron JE, Schechter M, Markowitz M, Loutfy MR, Lennox JL, Zhao J, Chen J, Ryan DM, Rhodes RR, Killar JA, Gilde LR, Strohmaier KM, Meibohm AR, Miller MD, Hazuda DJ, Nessly ML, DiNubile MJ, Isaacs RD, Teppler H, Nguyen BY. 2008. Subgroup and resistance analyses of raltegravir for resistant HIV-1 infection. N Engl J Med 359:355-365.

267. Blanco JL, Varghese V, Rhee SY, Gatell JM, Shafer RW. 2011. HIV-1 integrase inhibitor resistance and its clinical implications. J Infect Dis 203:1204-1214.

268. Goethals O, Clayton R, Van Ginderen M, Vereycken I, Wagemans E, Geluykens P, Dockx K, Strijbos R, Smits V, Vos A, Meersseman G, Jochmans D, Vermeire K, Schols D, Hallenberger S, Hertogs K. 2008. Resistance mutations in human immunodeficiency virus type 1 integrase selected with elvitegravir confer reduced susceptibility to a wide range of integrase inhibitors. J Virol 82:10366-10374.

269. Le Grice SF. 2012. Human immunodeficiency virus reverse transcriptase: 25 years of research, drug discovery, and promise. J Biol Chem 287:40850-40857.

270. Menendez-Arias L. 2013. Molecular basis of human immunodeficiency virus type 1 drug resistance: Overview and recent developments. Antiviral Res 98:93-120.

271. Huber-Ruano I, Pastor-Anglada M. 2009. Transport of nucleoside analogs across the plasma membrane: a clue to understanding drug-induced cytotoxicity. Curr Drug Metab 10:347-358.

272. Sarafianos SG, Marchand B, Das K, Himmel DM, Parniak MA, Hughes SH, Arnold E. 2009. Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. J Mol Biol 385:693-713.

273. Richman DD. 2001. HIV chemotherapy. Nature 410:995-1001.

274. Stanford University. 2015. Human immunodeficiency virus: A global pandemic. https://web.stanford.edu/group/virus/retro/2005gongishmail/HIV.html. Accessed 16 August 2015, .
275. Menendez-Arias L. 2010. Molecular basis of human immunodeficiency virus drug resistance: an update. Antiviral Res 85:210-231.
276. Arion D, Kaushik N, McCormick S, Borkow G, Parniak MA. 1998. Phenotypic mechanism of HIV-1 resistance to 3'-azido-3'-deoxythymidine (AZT): increased polymerization processivity and enhanced sensitivity to pyrophosphate of the mutant viral reverse transcriptase. Biochemistry 37:15908-15917.
277. Meyer PR, Matsuura SE, So AG, Scott WA. 1998. Unblocking of chain-terminated primer by HIV-1 reverse transcriptase through a nucleotide-dependent mechanism. Proc Natl Acad Sci U S A 95:13471-13476.
278. Yahi N, Fantini J, Henry M, Tourres C, Tamalet C. 2005. Structural analysis of reverse transcriptase mutations at codon 215 explains the predominance of T215Y over T215F in HIV-1 variants selected under antiretroviral therapy. J Biomed Sci 12:701-710.
279. Marcelin AG, Delaugerre C, Wirden M, Viegas P, Simon A, Katlama C, Calvez V. 2004. Thymidine analogue reverse transcriptase inhibitors resistance mutations profiles and association to other nucleoside reverse transcriptase inhibitors resistance mutations observed in the context of virological failure. J Med Virol 72:162-165.
280. Harrigan PR, Kinghorn I, Bloor S, Kemp SD, Najera I, Kohli A, Larder BA. 1996. Significance of amino acid variation at human immunodeficiency virus type 1 reverse transcriptase residue 210 for zidovudine susceptibility. J Virol 70:5930-5934.
281. Larder BA, Coates KE, Kemp SD. 1991. Zidovudine-resistant human immunodeficiency virus selected by passage in cell culture. J Virol 65:5232-5236.
282. Hanna GJ, Johnson VA, Kuritzkes DR, Richman DD, Brown AJ, Savara AV, Hazelwood JD, D'Aquila RT. 2000. Patterns of resistance mutations selected by treatment of human immunodeficiency virus type 1 infection with zidovudine, didanosine, and nevirapine. J Infect Dis 181:904-911.
283. Das K, Arnold E. 2013. HIV-1 reverse transcriptase and antiviral drug resistance. Part 1. Curr Opin Virol 3:111-118.

284. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA. 1992. Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science 256:1783-1790.

285. Xia Q, Radzio J, Anderson KS, Sluis-Cremer N. 2007. Probing nonnucleoside inhibitor-induced active-site distortion in HIV-1 reverse transcriptase by transient kinetic analyses. Protein Sci 16:1728-1737.

286. Rittinger K, Divita G, Goody RS. 1995. Human immunodeficiency virus reverse transcriptase substrate-induced conformational changes and the mechanism of inhibition by nonnucleoside inhibitors. Proc Natl Acad Sci U S A 92:8046-8049.

287. Rodgers DW, Gamblin SJ, Harris BA, Ray S, Culp JS, Hellmig B, Woolf DJ, Debouck C, Harrison SC. 1995. The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1. Proc Natl Acad Sci U S A 92:1222-1226.

288. Sluis-Cremer N, Temiz NA, Bahar I. 2004. Conformational changes in HIV-1 reverse transcriptase induced by nonnucleoside reverse transcriptase inhibitor binding. Curr HIV Res 2:323-332.

289. Tantillo C, Ding J, Jacobo-Molina A, Nanni RG, Boyer PL, Hughes SH, Pauwels R, Andries K, Janssen PA, Arnold E. 1994. Locations of anti-AIDS drug binding sites and resistance mutations in the three-dimensional structure of HIV-1 reverse transcriptase. Implications for mechanisms of drug inhibition and resistance. J Mol Biol 243:369-387.

290. Tambuyzer L, Azijn H, Rimsky LT, Vingerhoets J, Lecocq P, Kraus G, Picchio G, de Bethune MP. 2009. Compilation and prevalence of mutations associated with resistance to non-nucleoside reverse transcriptase inhibitors. Antivir Ther 14:103-109.

291. Pomerantz RJ, Horn DL. 2003. Twenty years of therapy for HIV-1 infection. Nat Med 9:867-873.

292. Hull MW, Montaner JS. 2011. Ritonavir-boosted Protease inhibitors in HIV therapy. Ann Med 43:375-388.

293. Kempf DJ, Marsh KC, Kumar G, Rodrigues AD, Denissen JF, McDonald E, Kukulka MJ, Hsu A, Granneman GR, Baroldi PA, Sun E, Pizzuti D, Plattner JJ, Norbeck DW, Leonard JM. 1997. Pharmacokinetic enhancement of inhibitors of the human immunodeficiency virus Protease by coadministration with ritonavir. Antimicrob Agents Chemother 41:654-660.

294. Turner SR, Strohbach JW, Tommasi RA, Aristoff PA, Johnson PD, Skulnick HI, Dolak LA, Seest EP, Tomich PK, Bohanon MJ, Horng MM, Lynn JC, Chong KT, Hinshaw RR, Watenpaugh KD, Janakiraman MN, Thaisrivongs S. 1998. Tipranavir (PNU-140690): a potent, orally bioavailable nonpeptidic HIV Protease inhibitor of the 5,6-dihydro-4-hydroxy-2-pyrone sulfonamide class. J Med Chem 41:3467-3476.

295. Tang C, Louis JM, Aniana A, Suh JY, Clore GM. 2008. Visualizing transient events in amino-terminal autoprocessing of HIV-1 Protease. Nature 455:693-696.

296. Koh Y, Matsumi S, Das D, Amano M, Davis DA, Li J, Leschenko S, Baldridge A, Shioda T, Yarchoan R, Ghosh AK, Mitsuya H. 2007. Potent inhibition of HIV-1 replication by novel non-peptidyl small molecule inhibitors of Protease dimerization. J Biol Chem 282:28709-28720.

297. Shafer RW, Dupnik K, Winters MA, Eshleman SH. 2001. A Guide to HIV-1 Reverse Transcriptase and Protease Sequencing for Drug Resistance Studies. HIV Seq Compend 2001:1-51.

298. Molla A, Korneyeva M, Gao Q, Vasavanonda S, Schipper PJ, Mo HM, Markowitz M, Chernyavskiy T, Niu P, Lyons N, Hsu A, Granneman GR, Ho DD, Boucher CA, Leonard JM, Norbeck DW, Kempf DJ. 1996. Ordered accumulation of mutations in HIV Protease confers resistance to ritonavir. Nat Med 2:760-766.

299. Mahalingam B, Louis JM, Reed CC, Adomat JM, Krouse J, Wang YF, Harrison RW, Weber IT. 1999. Structural and kinetic analysis of drug resistant mutants of HIV-1 Protease. Eur J Biochem 263:238-245.

300. Croteau G, Doyon L, Thibeault D, McKercher G, Pilote L, Lamarre D. 1997. Impaired fitness of human immunodeficiency virus type 1 variants with high-level resistance to Protease inhibitors. J Virol 71:1089-1096.

301. Gulnik SV, Suvorov LI, Liu B, Yu B, Anderson B, Mitsuya H, Erickson JW. 1995. Kinetic characterization and cross-resistance patterns of HIV-1 Protease mutants selected under drug pressure. Biochemistry 34:9282-9287.

302. Sugiura W, Matsuda Z, Yokomaku Y, Hertogs K, Larder B, Oishi T, Okano A, Shiino T, Tatsumi M, Matsuda M, Abumi H, Takata N, Shirahata S, Yamada K, Yoshikura H, Nagai Y. 2002. Interference between D30N and L90M in selection and development of Protease inhibitor-resistant human immunodeficiency virus type 1. Antimicrob Agents Chemother 46:708-715.

303. Menendez-Arias L, Martinez MA, Quinones-Mateu ME, Martinez-Picado J. 2003. Fitness variations and their impact on the evolution of antiretroviral drug resistance. Curr Drug Targets Infect Disord 3:355-371.
304. Chang MW, Torbett BE. 2011. Accessory mutations maintain stability in drug-resistant HIV-1 Protease. J Mol Biol 410:756-760.
305. Zhang YM, Imamichi H, Imamichi T, Lane HC, Falloon J, Vasudevachari MB, Salzman NP. 1997. Drug resistance during indinavir therapy is caused by mutations in the Protease gene and in its Gag substrate cleavage sites. J Virol 71:6662-6670.
306. Doyon L, Payant C, Brakier-Gingras L, Lamarre D. 1998. Novel Gag-Pol frameshift site in human immunodeficiency virus type 1 variants resistant to Protease inhibitors. J Virol 72:6146-6150.
307. Robinson LH, Myers RE, Snowden BW, Tisdale M, Blair ED. 2000. HIV type 1 Protease cleavage site mutations and viral fitness: implications for drug susceptibility phenotyping assays. AIDS Res Hum Retroviruses 16:1149-1156.
308. Pettit SC, Henderson GJ, Schiffer CA, Swanstrom R. 2002. Replacement of the P1 amino acid of human immunodeficiency virus type 1 Gag processing sites can inhibit or enhance the rate of cleavage by the viral Protease. J Virol 76:10226-10233.
309. Nijhuis M, Deeks S, Boucher C. 2001. Implications of antiretroviral resistance on viral fitness. Curr Opin Infect Dis 14:23-28.
310. Gatanaga H, Suzuki Y, Tsang H, Yoshimura K, Kavlick MF, Nagashima K, Gorelick RJ, Mardy S, Tang C, Summers MF, Mitsuya H. 2002. Amino acid substitutions in Gag protein at non-cleavage sites are indispensable for the development of a high multitude of HIV-1 resistance against Protease inhibitors. J Biol Chem 277:5952-5961.
311. Dam E, Quercia R, Glass B, Descamps D, Launay O, Duval X, Krausslich HG, Hance AJ, Clavel F, Group AS. 2009. Gag mutations strongly contribute to HIV-1 resistance to Protease inhibitors in highly drug-experienced patients besides compensating for fitness loss. PLoS Pathog 5:e1000345.
312. Myint L, Matsuda M, Matsuda Z, Yokomaku Y, Chiba T, Okano A, Yamada K, Sugiura W. 2004. Gag non-cleavage site mutations contribute to full recovery of viral fitness in Protease inhibitor-resistant human immunodeficiency virus type 1. Antimicrob Agents Chemother 48:444-452.

313. Prado JG, Wrin T, Beauchaine J, Ruiz L, Petropoulos CJ, Frost SD, Clotet B, D'Aquila RT, Martinez-Picado J. 2002. Amprenavir-resistant HIV-1 exhibits lopinavir cross-resistance and reduced replication capacity. AIDS 16:1009-1017.
314. Yates PJ, Hazen R, St Clair M, Boone L, Tisdale M, Elston RC. 2006. In vitro development of resistance to human immunodeficiency virus Protease inhibitor GW640385. Antimicrob Agents Chemother 50:1092-1095.
315. Kolli M, Stawiski E, Chappey C, Schiffer CA. 2009. Human immunodeficiency virus type 1 Protease-correlated cleavage site mutations enhance inhibitor resistance. J Virol 83:11027-11042.
316. Nijhuis M, van Maarseveen NM, Lastere S, Schipper P, Coakley E, Glass B, Rovenska M, de Jong D, Chappey C, Goedegebuure IW, Heilek-Snyder G, Dulude D, Cammack N, Brakier-Gingras L, Konvalinka J, Parkin N, Krausslich HG, Brun-Vezinet F, Boucher CA. 2007. A novel substrate-based HIV-1 Protease inhibitor drug resistance mechanism. PLOS Medicine 4:e36.
317. Gupta RK, Kohli A, McCormick AL, Towers GJ, Pillay D, Parry CM. 2010. Full-length HIV-1 Gag determines Protease inhibitor susceptibility within in vitro assays. AIDS 24:1651-1655.
318. Kolli M, Lastere S, Schiffer CA. 2006. Co-evolution of nelfinavir-resistant HIV-1 Protease and the p1-p6 substrate. Virology 347:405-409.
319. van Maarseveen NM, Andersson D, Lepsik M, Fun A, Schipper PJ, de Jong D, Boucher CA, Nijhuis M. 2012. Modulation of HIV-1 Gag NC/p1 cleavage efficiency affects Protease inhibitor resistance and viral replicative capacity. Retrovirology 9:29.
320. Cote HC, Brumme ZL, Harrigan PR. 2001. Human immunodeficiency virus type 1 Protease cleavage site mutations associated with Protease inhibitor cross-resistance selected by indinavir, ritonavir, and/or saquinavir. J Virol 75:589-594.
321. Stanford University HIV Drug Resistance Database. 9 March 2015 2015. Drug resistance summaries. http://hivdb.stanford.edu/pages/download/resistanceMutations_handout.pdf. Accessed 25 September 2015.
322. Wensing AM, Calvez V, Gunthard HF, Johnson VA, Paredes R, Pillay D, Shafer RW, Richman DD. 2014. 2014 Update of the drug resistance mutations in HIV-1. Top Antivir Med 22:642-650.
323. Li G, Verheyen J, Theys K, Piampongsant S, Van Laethem K, Vandamme AM. 2014. HIV-1 Gag C-terminal amino acid substitutions emerging under selective

pressure of Protease inhibitors in patient populations infected with different HIV-1 subtypes. Retrovirology 11:79.

324. Aoki M, Venzon DJ, Koh Y, Aoki-Ogata H, Miyakawa T, Yoshimura K, Maeda K, Mitsuya H. 2009. Non-cleavage site Gag mutations in amprenavir-resistant human immunodeficiency virus type 1 (HIV-1) predispose HIV-1 to rapid acquisition of amprenavir resistance but delay development of resistance to other Protease inhibitors. J Virol 83:3059-3068.
325. Koh Y, Das D, Leschenko S, Nakata H, Ogata-Aoki H, Amano M, Nakayama M, Ghosh AK, Mitsuya H. 2009. GRL-02031, a novel nonpeptidic Protease inhibitor (PI) containing a stereochemically defined fused cyclopentanyltetrahydrofuran potent against multi-PI-resistant human immunodeficiency virus type 1 in vitro. Antimicrob Agents Chemother 53:997-1006.
326. Chang MW, Oliveira G, Yuan J, Okulicz JF, Levy S, Torbett BE. 2013. Rapid deep sequencing of patient-derived HIV with ion semiconductor technology. J Virol Methods 189:232-234.
327. Parry CM, Kohli A, Boinett CJ, Towers GJ, McCormick AL, Pillay D. 2009. Gag determinants of fitness and drug susceptibility in Protease inhibitor-resistant human immunodeficiency virus type 1. J Virol 83:9094-9101.
328. Parry CM, Kolli M, Myers RE, Cane PA, Schiffer C, Pillay D. 2011. Three residues in HIV-1 matrix contribute to Protease inhibitor susceptibility and replication capacity. Antimicrob Agents Chemother 55:1106-1113.
329. Larrouy L, Chazallon C, Landman R, Capitant C, Peytavin G, Collin G, Charpentier C, Storto A, Pialoux G, Katlama C, Girard PM, Yeni P, Aboulker JP, Brun-Vezinet F, Descamps D, Group AS. 2010. Gag mutations can impact virological response to dual-boosted Protease inhibitor combinations in antiretroviral-naive HIV-infected patients. Antimicrob Agents Chemother 54:2910-2919.
330. Knops E, Daumer M, Awerkiew S, Kartashev V, Schulter E, Kutsev S, Brakier-Gingras L, Kaiser R, Pfister H, Verheyen J. 2010. Evolution of Protease inhibitor resistance in the Gag and pol genes of HIV subtype G isolates. J Antimicrob Chemother 65:1472-1476.
331. Mammano F, Petit C, Clavel F. 1998. Resistance-associated loss of viral fitness in human immunodeficiency virus type 1: phenotypic analysis of Protease and Gag coevolution in Protease inhibitor-treated patients. J Virol 72:7632-7637.
332. Nijhuis M, Schuurman R, de Jong D, Erickson J, Gustchina E, Albert J, Schipper P, Gulnik S, Boucher CA. 1999. Increased fitness of drug resistant HIV-1 Protease as

a result of acquisition of compensatory mutations during suboptimal therapy. AIDS 13:2349-2359.

333. Doyon L, Tremblay S, Bourgon L, Wardrop E, Cordingley MG. 2005. Selection and characterization of HIV-1 showing reduced susceptibility to the non-peptidic Protease inhibitor tipranavir. Antiviral Res 68:27-35.

334. Verheyen J, Verhofstede C, Knops E, Vandekerckhove L, Fun A, Brunen D, Dauwe K, Wensing AM, Pfister H, Kaiser R, Nijhuis M. 2010. High prevalence of bevirimat resistance mutations in Protease inhibitor-resistant HIV isolates. AIDS 24:669-673.

335. Seclen E, Gonzalez Mdel M, Corral A, de Mendoza C, Soriano V, Poveda E. 2010. High prevalence of natural polymorphisms in Gag (CA-SP1) associated with reduced response to Bevirimat, an HIV-1 maturation inhibitor. AIDS 24:467-469.

336. Malet I, Roquebert B, Dalban C, Wirden M, Amellal B, Agher R, Simon A, Katlama C, Costagliola D, Calvez V, Marcelin AG. 2007. Association of Gag cleavage sites to Protease mutations and to virological response in HIV-1 treated patients. J Infect 54:367-374.

337. Ghosn J, Pellegrin I, Goujard C, Deveau C, Viard JP, Galimand J, Harzic M, Tamalet C, Meyer L, Rouzioux C, Chaix ML. 2006. HIV-1 resistant strains acquired at the time of primary infection massively fuel the cellular reservoir and persist for lengthy periods of time. AIDS 20:159-170.

338. Ghosn J, Delaugerre C, Flandre P, Galimand J, Cohen-Codar I, Raffi F, Delfraissy JF, Rouzioux C, Chaix ML. 2011. Polymorphism in Gag gene cleavage sites of HIV-1 non-B subtype and virological outcome of a first-line lopinavir/ritonavir single drug regimen. PLoS One 6:e24798.

339. Larrouy L, Charpentier C, Landman R, Capitant C, Chazallon C, Yeni P, Peytavin G, Damond F, Brun-Vezinet F, Descamps D, group As. 2011. Dynamics of Gag-pol minority viral populations in naive HIV-1-infected patients failing Protease inhibitor regimen. AIDS 25:2143-2148.

340. Doyon L, Croteau G, Thibeault D, Poulin F, Pilote L, Lamarre D. 1996. Second locus involved in human immunodeficiency virus type 1 resistance to Protease inhibitors. J Virol 70:3763-3769.

341. Bally F, Martinez R, Peters S, Sudre P, Telenti A. 2000. Polymorphism of HIV type 1 Gag p7/p1 and p1/p6 cleavage sites: clinical significance and implications for resistance to Protease inhibitors. AIDS Res Hum Retroviruses 16:1209-1213.

342. Knops E, Kemper I, Schulter E, Pfister H, Kaiser R, Verheyen J. 2010. The evolution of Protease mutation 76V is associated with Protease mutation 46I and Gag mutation 431V. AIDS 24:779-781.
343. Verheyen J, Litau E, Sing T, Daumer M, Balduin M, Oette M, Fatkenheuer G, Rockstroh JK, Schuldenzucker U, Hoffmann D, Pfister H, Kaiser R. 2006. Compensatory mutations at the HIV cleavage sites p7/p1 and p1/p6-Gag in therapy-naive and therapy-experienced patients. Antivir Ther 11:879-887.
344. Nijhuis M, Wensing AM, Bierman WF, de Jong D, Kagan R, Fun A, Jaspers CA, Schurink KA, van Agtmael MA, Boucher CA. 2009. Failure of treatment with first-line lopinavir boosted with ritonavir can be explained by novel resistance pathways with Protease mutation 76V. J Infect Dis 200:698-709.
345. Banke S, Lillemark MR, Gerstoft J, Obel N, Jorgensen LB. 2009. Positive selection pressure introduces secondary mutations at Gag cleavage sites in human immunodeficiency virus type 1 harboring major Protease resistance mutations. J Virol 83:8916-8924.
346. Knops E, Brakier-Gingras L, Schulter E, Pfister H, Kaiser R, Verheyen J. 2012. Mutational patterns in the frameshift-regulating site of HIV-1 selected by Protease inhibitors. Med Microbiol Immunol 201:213-218.
347. Larrouy L, Vivot A, Charpentier C, Benard A, Visseaux B, Damond F, Matheron S, Chene G, Brun-Vezinet F, Descamps D, Cohort ACH-. 2013. Impact of Gag genetic determinants on virological outcome to boosted lopinavir-containing regimen in HIV-2-infected patients. AIDS 27:69-80.
348. Mo H, Parkin N, Stewart KD, Lu L, Dekhtyar T, Kempf DJ, Molla A. 2007. Identification and structural characterization of I84C and I84A mutations that are associated with high-level resistance to human immunodeficiency virus Protease inhibitors and impair viral replication. Antimicrob Agents Chemother 51:732-735.
349. Brann TW, Dewar RL, Jiang MK, Shah A, Nagashima K, Metcalf JA, Falloon J, Lane HC, Imamichi T. 2006. Functional correlation between a novel amino acid insertion at codon 19 in the Protease of human immunodeficiency virus type 1 and polymorphism in the p1/p6 Gag cleavage site in drug resistance and replication fitness. J Virol 80:6136-6145.
350. Kaufmann GR, Suzuki K, Cunningham P, Mukaide M, Kondo M, Imai M, Zaunders J, Cooper DA. 2001. Impact of HIV type 1 Protease, reverse transcriptase, cleavage site, and p6 mutations on the virological response to quadruple therapy

with saquinavir, ritonavir, and two nucleoside analogs. AIDS Res Hum Retroviruses 17:487-497.

351. Maguire MF, Guinea R, Griffin P, Macmanus S, Elston RC, Wolfram J, Richards N, Hanlon MH, Porter DJ, Wrin T, Parkin N, Tisdale M, Furfine E, Petropoulos C, Snowden BW, Kleim JP. 2002. Changes in human immunodeficiency virus type 1 Gag at positions L449 and P453 are linked to I50V Protease mutants in vivo and cause reduction of sensitivity to amprenavir and improved viral fitness in vitro. J Virol 76:7398-7406.

352. Watkins T, Resch W, Irlbeck D, Swanstrom R. 2003. Selection of high-level resistance to human immunodeficiency virus type 1 Protease inhibitors. Antimicrob Agents Chemother 47:759-769.

353. Hirsch MS, Gunthard HF, Schapiro JM, Brun-Vezinet F, Clotet B, Hammer SM, Johnson VA, Kuritzkes DR, Mellors JW, Pillay D, Yeni PG, Jacobsen DM, Richman DD. 2008. Antiretroviral drug resistance testing in adult HIV-1 infection: 2008 recommendations of an International AIDS Society-USA panel. Clin Infect Dis 47:266-285.

354. Chang S, Zhuang D, Li J, Liu S, Li H, Han J, Li L, Liu Y, Bao Z, Li T, Song H, Zhang W. 2015. Comparison of susceptibility of HIV-1 variants to antiretroviral drugs by genotypic and recombinant virus phenotypic analyses. Int J Infect Dis 37:86-92.

355. Slatko BE, Albright LM, Tabor S, Ju J. 2001. DNA sequencing by the dideoxy method. Curr Protoc Mol Biol Chapter 7:Unit7 4A.

356. Frentz D, Boucher CA, Assel M, De Luca A, Fabbiani M, Incardona F, Libin P, Manca N, Muller V, B ON, Paredes R, Prosperi M, Quiros-Roldan E, Ruiz L, Sloot PM, Torti C, Vandamme AM, Van Laethem K, Zazzi M, van de Vijver DA. 2010. Comparison of HIV-1 genotypic resistance test interpretation systems in predicting virological outcomes over time. PLoS One 5:e11505.

357. Xiaobai Z, Xi C, Tian H, Williams AB, Wang H, He J, Zhen J, Chiarella J, Blake LA, Turenchalk G, Kozal MJ. 2014. Prevalence of WHO transmitted drug resistance mutations by deep sequencing in antiretroviral-naive subjects in Hunan Province, China. PLoS One 9:e98740.

358. Palmer S, Kearney M, Maldarelli F, Halvas EK, Bixby CJ, Bazmi H, Rock D, Falloon J, Davey RT, Jr., Dewar RL, Metcalf JA, Hammer S, Mellors JW, Coffin JM. 2005. Multiple, linked human immunodeficiency virus type 1 drug resistance

mutations in treatment-experienced patients are missed by standard genotype analysis. J Clin Microbiol 43:406-413.

359. Johnson JA, Li JF, Wei X, Lipscomb J, Irlbeck D, Craig C, Smith A, Bennett DE, Monsour M, Sandstrom P, Lanier ER, Heneine W. 2008. Minority HIV-1 drug resistance mutations are present in antiretroviral treatment-naive populations and associate with reduced treatment efficacy. PLoS Med 5:e158.

360. Svarovskaia ES, Margot NA, Bae AS, Waters JM, Goodman D, Zhong L, Borroto-Esoda K, Miller MD. 2007. Low-level K65R mutation in HIV-1 reverse transcriptase of treatment-experienced patients exposed to abacavir or didanosine. J Acquir Immune Defic Syndr 46:174-180.

361. Palmer S, Boltz V, Maldarelli F, Kearney M, Halvas EK, Rock D, Falloon J, Davey RT, Jr., Dewar RL, Metcalf JA, Mellors JW, Coffin JM. 2006. Selection and persistence of non-nucleoside reverse transcriptase inhibitor-resistant HIV-1 in patients starting and stopping non-nucleoside therapy. AIDS 20:701-710.

362. Paredes R, Lalama CM, Ribaudo HJ, Schackman BR, Shikuma C, Giguel F, Meyer WA, 3rd, Johnson VA, Fiscus SA, D'Aquila RT, Gulick RM, Kuritzkes DR. 2010. Pre-existing minority drug-resistant HIV-1 variants, adherence, and risk of antiretroviral treatment failure. J Infect Dis 201:662-671.

363. Li JZ, Paredes R, Ribaudo HJ, Svarovskaia ES, Metzner KJ, Kozal MJ, Hullsiek KH, Balduin M, Jakobsen MR, Geretti AM, Thiebaut R, Ostergaard L, Masquelier B, Johnson JA, Miller MD, Kuritzkes DR. 2011. Low-frequency HIV-1 drug resistance mutations and risk of NNRTI-based antiretroviral treatment failure: a systematic review and pooled analysis. JAMA 305:1327-1335.

364. Palmer S, Boltz V, Martinson N, Maldarelli F, Gray G, McIntyre J, Mellors J, Morris L, Coffin J. 2006. Persistence of nevirapine-resistant HIV-1 in women after single-dose nevirapine therapy for prevention of maternal-to-fetal HIV-1 transmission. Proc Natl Acad Sci U S A 103:7094-7099.

365. Jourdain G, Ngo-Giang-Huong N, Le Coeur S, Bowonwatanuwong C, Kantipong P, Leechanachai P, Ariyadej S, Leenasirimakul P, Hammer S, Lallemant M. 2004. Intrapartum exposure to nevirapine and subsequent maternal responses to nevirapine-based antiretroviral therapy. N Engl J Med 351:229-240.

366. Kapoor A, Jones M, Shafer RW, Rhee SY, Kazanjian P, Delwart EL. 2004. Sequencing-based detection of low-frequency human immunodeficiency virus type 1 drug-resistant mutants by an RNA/DNA heteroduplex generator-tracking assay. J Virol 78:7112-7123.

367. Lecossier D, Shulman NS, Morand-Joubert L, Shafer RW, Joly V, Zolopa AR, Clavel F, Hance AJ. 2005. Detection of minority populations of HIV-1 expressing the K103N resistance mutation in patients failing nevirapine. J Acquir Immune Defic Syndr 38:37-42.

368. Halvas EK, Aldrovandi GM, Balfe P, Beck IA, Boltz VF, Coffin JM, Frenkel LM, Hazelwood JD, Johnson VA, Kearney M, Kovacs A, Kuritzkes DR, Metzner KJ, Nissley DV, Nowicki M, Palmer S, Ziermann R, Zhao RY, Jennings CL, Bremer J, Brambilla D, Mellors JW. 2006. Blinded, multicenter comparison of methods to detect a drug-resistant mutant of human immunodeficiency virus type 1 at low frequency. J Clin Microbiol 44:2612-2614.

369. Lange CM, Hue S, Violari A, Cotton M, Gibb D, Babiker A, Otwombe K, Panchia R, Dobbels E, Jean-Philippe P, McIntyre JA, Pillay D, Gupta RK. 2015. Single Genome Analysis for the Detection of Linked Multiclass Drug Resistance Mutations in HIV-1-Infected Children After Failure of Protease Inhibitor-Based First-Line Therapy. J Acquir Immune Defic Syndr 69:138-144.

370. Avidor B, Girshengorn S, Matus N, Talio H, Achsanov S, Zeldis I, Fratty IS, Katchman E, Brosh-Nissimov T, Hassin D, Alon D, Bentwich Z, Yust I, Amit S, Forer R, Vulih Shultsman I, Turner D. 2013. Evaluation of a benchtop HIV ultradeep pyrosequencing drug resistance assay in the clinical laboratory. J Clin Microbiol 51:880-886.

371. Etemad B, Ghulam-Smith M, Gonzalez O, White LF, Sagar M. 2015. Single genome amplification and standard bulk PCR yield HIV-1 envelope products with similar genotypic and phenotypic characteristics. J Virol Methods 214:46-53.

372. Grada A, Weinbrecht K. 2013. Next-generation sequencing: methodology and application. J Invest Dermatol 133:e11.

373. Hanna GJ, Caliendo AM. 2001. Testing for HIV-1 drug resistance. Mol Diagn 6:253-263.

374. Deeks SG, Barbour JD, Martin JN, Swanson MS, Grant RM. 2000. Sustained CD4+ T cell response after virologic failure of Protease inhibitor-based regimens in patients with human immunodeficiency virus infection. J Infect Dis 181:946-953.

375. Dauber DS, Ziermann R, Parkin N, Maly DJ, Mahrus S, Harris JL, Ellman JA, Petropoulos C, Craik CS. 2002. Altered substrate specificity of drug-resistant human immunodeficiency virus type 1 Protease. J Virol 76:1359-1368.

376. Malim MH, Emerman M. 2001. HIV-1 sequence variation: drift, shift, and attenuation. Cell 104:469-472.

377. Prado JG, Parkin NT, Clotet B, Ruiz L, Martinez-Picado J. 2005. HIV type 1 fitness evolution in antiretroviral-experienced patients with sustained CD4+ T cell counts but persistent virologic failure. Clin Infect Dis 41:729-737.

378. Deeks SG, Wrin T, Liegler T, Hoh R, Hayden M, Barbour JD, Hellmann NS, Petropoulos CJ, McCune JM, Hellerstein MK, Grant RM. 2001. Virologic and immunologic consequences of discontinuing combination antiretroviral-drug therapy in HIV-infected patients with detectable viremia. N Engl J Med 344:472-480.

379. De Luca A. 2006. The impact of resistance on viral fitness and its clinical implications. *In* Geretti AM (ed), Antiretroviral Resistance in Clinical Practice, London.

380. van Opijnen T, Berkhout B. 2005. The host environment drives HIV-1 fitness. Rev Med Virol 15:219-233.

381. Barbour J, Grant R. 2005. The role of viral fitness in HIV pathogenesis. Current HIV/AIDS Reports 2:29-34.

382. Barbour JD, Deeks SG. 2008. Clinical implications of HIV fitness and virulence, p 161. *In* Bell M, Bowler K (ed), Global HIV/AIDS medicine, vol 1. Elsevier, Philadelphia.

383. Martinez-Picado J, Prado JG, Fry EE, Pfafferott K, Leslie A, Chetty S, Thobakgale C, Honeyborne I, Crawford H, Matthews P, Pillay T, Rousseau C, Mullins JI, Brander C, Walker BD, Stuart DI, Kiepiela P, Goulder P. 2006. Fitness cost of escape mutations in p24 Gag in association with control of human immunodeficiency virus type 1. J Virol 80:3617-3623.

384. Crawford H, Lumm W, Leslie A, Schaefer M, Boeras D, Prado JG, Tang J, Farmer P, Ndung'u T, Lakhi S, Gilmour J, Goepfert P, Walker BD, Kaslow R, Mulenga J, Allen S, Goulder PJ, Hunter E. 2009. Evolution of HLA-B*5703 HIV-1 escape mutations in HLA-B*5703-positive individuals and their transmission recipients. J Exp Med 206:909-921.

385. Brockman MA, Schneidewind A, Lahaie M, Schmidt A, Miura T, Desouza I, Ryvkin F, Derdeyn CA, Allen S, Hunter E, Mulenga J, Goepfert PA, Walker BD, Allen TM. 2007. Escape and compensation from early HLA-B57-mediated cytotoxic T-lymphocyte pressure on human immunodeficiency virus type 1 Gag alter capsid interactions with cyclophilin A. J Virol 81:12608-12618.

386. Dykes C, Demeter LM. 2007. Clinical significance of human immunodeficiency virus type 1 replication fitness. Clin Microbiol Rev 20:550-578.

387. Singh A, Page T, Moore PL, Allgaier RL, Hiramen K, Coovadia HM, Walker BD, Morris L, Ndung'u T. 2009. Functional and genetic analysis of coreceptor usage by dualtropic HIV-1 subtype C isolates. Virology 393:56-67.
388. Weber J, Weberova J, Carobene M, Mirza M, Martinez-Picado J, Kazanjian P, Quinones-Mateu ME. 2006. Use of a novel assay based on intact recombinant viruses expressing green (EGFP) or red (DsRed2) fluorescent proteins to examine the contribution of pol and Env genes to overall HIV-1 replicative fitness. J Virol Methods 136:102-117.
389. Fikkert V, Cherepanov P, Van Laethem K, Hantson A, Van Remoortel B, Pannecouque C, De Clercq E, Debyser Z, Vandamme AM, Witvrouw M. 2002. Env chimeric virus technology for evaluating human immunodeficiency virus susceptibility to entry inhibitors. Antimicrob Agents Chemother 46:3954-3962.
390. Dudley DM, Gao Y, Nelson KN, Henry KR, Nankya I, Gibson RM, Arts EJ. 2009. A novel yeast-based recombination method to clone and propagate diverse HIV-1 isolates. Biotechniques 46:458-467.
391. Lobritz MA, Marozsan AJ, Troyer RM, Arts EJ. 2007. Natural variation in the V3 crown of human immunodeficiency virus type 1 affects replicative fitness and entry inhibitor sensitivity. J Virol 81:8258-8269.
392. Marozsan AJ, Moore DM, Lobritz MA, Fraundorf E, Abraha A, Reeves JD, Arts EJ. 2005. Differences in the fitness of two diverse wild-type human immunodeficiency virus type 1 isolates are related to the efficiency of cell binding and entry. J Virol 79:7121-7134.
393. Marozsan AJ, Arts EJ. 2003. Development of a yeast-based recombination cloning/system for the analysis of gene products from diverse human immunodeficiency virus type 1 isolates. J Virol Methods 111:111-120.
394. Costa CI, Morgado MG, Santos VG, Bongertz V. 1996. HIV-1 isolation from plasma specimens. HEC/FIOCRUZ AIDS Clinical Research Group. Mem Inst Oswaldo Cruz 91:745-746.
395. Gervaix A, West D, Leoni LM, Richman DD, Wong-Staal F, Corbeil J. 1997. A new reporter cell line to monitor HIV infection and drug susceptibility in vitro. Proc Natl Acad Sci U S A 94:4653-4658.
396. Back NK, Nijhuis M, Keulen W, Boucher CA, Oude Essink BO, van Kuilenburg AB, van Gennip AH, Berkhout B. 1996. Reduced replication of 3TC-resistant HIV-1 variants in primary cells due to a processivity defect of the reverse transcriptase enzyme. EMBO J 15:4040-4049.

397. Eastman PS, Mittler J, Kelso R, Gee C, Boyer E, Kolberg J, Urdea M, Leonard JM, Norbeck DW, Mo H, Markowitz M. 1998. Genotypic changes in human immunodeficiency virus type 1 associated with loss of suppression of plasma viral RNA levels in subjects treated with ritonavir (Norvir) monotherapy. J Virol 72:5154-5164.

398. Devereux HL, Emery VC, Johnson MA, Loveday C. 2001. Replicative fitness in vivo of HIV-1 variants with multiple drug resistance-associated mutations. J Med Virol 65:218-224.

399. Prado JG, Prendergast A, Thobakgale C, Molina C, Tudor-Williams G, Ndung'u T, Walker BD, Goulder P. 2010. Replicative capacity of human immunodeficiency virus type 1 transmitted from mother to child is associated with pediatric disease progression rate. J Virol 84:492-502.

400. Dykes C, Wang J, Jin X, Planelles V, An DS, Tallo A, Huang Y, Wu H, Demeter LM. 2006. Evaluation of a multiple-cycle, recombinant virus, growth competition assay that uses flow cytometry to measure replication efficiency of human immunodeficiency virus type 1 in cell culture. J Clin Microbiol 44:1930-1943.

401. Resch W, Ziermann R, Parkin N, Gamarnik A, Swanstrom R. 2002. Nelfinavir-resistant, amprenavir-hypersusceptible strains of human immunodeficiency virus type 1 carrying an N88S mutation in Protease have reduced infectivity, reduced replication capacity, and reduced fitness and process the Gag polyprotein precursor aberrantly. J Virol 76:8659-8666.

402. Clavel F, Race E, Mammano F. 2000. HIV drug resistance and viral fitness. Adv Pharmacol 49:41-66.

403. Martinez-Picado J, Savara AV, Sutton L, D'Aquila RT. 1999. Replicative fitness of Protease inhibitor-resistant mutants of human immunodeficiency virus type 1. J Virol 73:3744-3752.

404. Nicastri E, Sarmati L, d'Ettorre G, Palmisano L, Parisi SG, Uccella I, Rianda A, Concia E, Vullo V, Vella S, Andreoni M. 2003. Replication capacity, biological phenotype, and drug resistance of HIV strains isolated from patients failing antiretroviral therapy. J Med Virol 69:1-6.

405. SA AIDS. 2013. Southern African HIV and AIDS information dissemination service. South Africa" National Strategic Plan 2012-2016 – Living with HIV. http://www.safaids.net/content/south-africa-national-strategic-plan-2012-2016-living-hiv Accessed 17 March.

406. National department of Health. 2013. The South African National Department of Health Treatment Guidelines.
407. Alcalde R, Guimaraes ML, Duarte AJ, Casseb J. 2012. Clinical, epidemiological and molecular features of the HIV-1 subtype C and recombinant forms that are circulating in the city of Sao Paulo, Brazil. Virol J 9:156.
408. Sutherland KA, Mbisa JL, Cane PA, Pillay D, Parry CM. 2014. Contribution of Gag and Protease to variation in susceptibility to Protease inhibitors between different strains of subtype B human immunodeficiency virus type 1. J Gen Virol 95:190-200.
409. Gadhamsetty S, Dixit NM. 2010. Estimating frequencies of minority nevirapine-resistant strains in chronically HIV-1-infected individuals naive to nevirapine by using stochastic simulations and a mathematical model. J Virol 84:10230-10240.
410. Little SJ, Frost SD, Wong JK, Smith DM, Pond SL, Ignacio CC, Parkin NT, Petropoulos CJ, Richman DD. 2008. Persistence of transmitted drug resistance among subjects with primary human immunodeficiency virus infection. J Virol 82:5510-5518.
411. Garcia-Lerma JG, Nidtha S, Blumoff K, Weinstock H, Heneine W. 2001. Increased ability for selection of zidovudine resistance in a distinct class of wild-type HIV-1 from drug-naive persons. Proc Natl Acad Sci U S A 98:13907-13912.
412. Brenner B, Routy JP, Quan Y, Moisi D, Oliveira M, Turner D, Wainberg MA. 2004. Persistence of multidrug-resistant HIV-1 in primary infection leading to superinfection. AIDS 18:1653-1660.
413. Yang WL, Kouyos RD, Boni J, Yerly S, Klimkait T, Aubert V, Scherrer AU, Shilaih M, Hinkley T, Petropoulos C, Bonhoeffer S, Gunthard HF. 2015. Persistence of transmitted HIV-1 drug resistance mutations associated with fitness costs and viral genetic backgrounds. PLoS Pathog 11:e1004722.
414. Martinez-Picado J, Martinez MA. 2008. HIV-1 reverse transcriptase inhibitor resistance mutations and fitness: a view from the clinic and ex vivo. Virus Res 134:104-123.
415. Cong ME, Heneine W, Garcia-Lerma JG. 2007. The fitness cost of mutations associated with human immunodeficiency virus type 1 drug resistance is modulated by mutational interactions. J Virol 81:3037-3041.
416. Li JZ, Kuritzkes DR. 2013. Clinical implications of HIV-1 minority variants. Clin Infect Dis 56:1667-1674.

417. Gandhi RT, Wurcel A, Rosenberg ES, Johnston MN, Hellmann N, Bates M, Hirsch MS, Walker BD. 2003. Progressive reversion of human immunodeficiency virus type 1 resistance mutations in vivo after transmission of a multiply drug-resistant virus. Clin Infect Dis 37:1693-1698.
418. Pillay K. 2015. The impact of the p7/p1 cleavage site mutations on replication capacity and drug resistance in HIV-1 subtype C. Masters of Science. University o Kwa-Zulu Natal, South Africa.

CHAPTER 2

Identification of Gag mutations associated with PI drug resistance/exposure in HIV-1 subtype C

2 CHAPTER 2: IDENTIFICATION OF GAG MUTATIONS ASSOCIATED WITH PI DRUG RESISTANCE/EXPOSURE IN HIV-1 SUBTYPE C

2.1 Introduction

Protease inhibitors have been described as the most effective class of drug employed in the treatment of HIV-1 (1-3), nonetheless their clinical benefits can be compromised by the development of drug resistance.

A key feature in PI resistance is that mutations which confer resistance are not only confined to PR itself but also occur in its' natural substrate (i.e. Gag) (4-7). Mutations in Gag can either confer primary resistance to PI's (8-10) or work synergistically with PR to: restore, improve or maintain the replicative capacity of HIV-1 (4, 11). As such PR and Gag have been labelled as "partners in resistance" (4).

Besides being "partners in resistance", Gag and PR are also partners in viral maturation, whereby PR is required to cleave the Gag and Gag-pol polyproteins into their respective structural (MA, CA, NC and P6) and enzymatic (RT, PR and INT) proteins, in order to generate mature infectious virions (described in section 1.6.9). This essential role of PR has driven the design of PI's, most of which bind to the active site of PR where they competitively inhibit substrate binding and cleavage thereby preventing the production of mature infectious viral particles (described in section 1.10.1.6).

During exposure to PI's, mutations can occur in Gag or PR to confer resistance (described in section 1.10.1.6) (8, 12, 13). Mutations at the PR active site alter interactions between PI's and binding site amino acids, resulting in a decreased affinity of the binding site for PI's. Consequently, the binding affinity for the natural substrate and subsequent viral replication is reduced (13-16). As a compensatory measure, secondary mutations are selected in PR or Gag in order to restore/improve cleavage and replicative capability, and/or in some cases enhance drug resistance (17-19). Compensatory mutations in Gag can occur at either CS's or non-CS's. In the case of CSMs, the Gag-pol frameshift is altered resulting in increased expression of pol products (i.e. more production of PR) which directly improves production of mature infectious virus (20). In contrast, non-CSMs

improve the access of PR to CS thereby improving cleavage and viral replication (17, 21, 22).

The majority of documented CSM's have been identified within the NC/p1 and p1/p6 Gag cleavage sites, since commercial kits were designed to include coverage of these regions when amplifying and sequencing the PR region (4, 17, 19). The L449F and R452K mutations are examples of Gag mutations at the p1/p6 cleavage site which work synergistically with PR RAMs to enhance resistance to PI's (23, 24). Similarly, compensatory mutations in Gag non-cleavage sites such as R76K, Y79F and T81A have also been reported to occur in conjunction with PR RAMs to confer reduced susceptibility to PI's and improve viral replication (11, 25).

Mutations such as A431V, I437V and the double mutations K436E/R and I437T have been shown to cause primary resistance to PI's in the absence of PR RAMs (8-10, 12, 22, 24, 26, 27). Studies have also suggested that mutations in Gag could possibly lead to the development of mutations in PR (4).

Even though HIV-1 subtype C is the most prevalent subtype globally (28), the majority of studies investigating the role of Gag in PI resistance have focused on HIV-1 subtype B (11, 23-26, 29, 30) and to a lesser extent, HIV-1 subtype A and D (6). Currently there are only two studies which investigated Gag mutations associated with PI resistance in HIV-1 subtype C (31, 32). The first study used sequences downloaded from the Los Alamos database, to identify Gag amino acids associated with PI resistance in HIV-1 subtypes A1, B, C, D, F1, G, CRF01_AE and CRF01_AG. The authors showed that most Gag mutations associated with PI resistance occurred in the carboxy terminal of Gag (31). This study however did not have matched treatment data for each sequence therefore associations between Gag mutations and specific PI's could not be established. Such, information is integral for the development of gag resistance testing algorithms. The second study investigated Gag mutations associated with PI resistance in 20 paediatric patients (32). This study showed that the majority of paediatric patients failing treatment did not harbour PR mutations, but did harbour mutations in the Gag CS and non-CS, which could directly contribute to PI resistance. The limitation of this study however was the small sample size.

In light of the limited number of studies on HIV-1 subtype C and the fact that subtype variation is associated with variance in the development of drug resistance (33), the need

for more HIV-1 subtype C studies investigating PI associated resistance mutations in gag, using larger cohorts with known treatment history's', is recognised. Such data would be useful for the inclusion of Gag in resistance testing and interpretation algorithms.

This current study aimed to identify Gag-Protease mutations associated with PI resistance/exposure in an HIV-1 subtype C cohort of 80 patients failing a PI inclusive treatment regimen between 2009-2013, from KwaZulu-Natal, South Africa. Briefly, the Gag-Protease region from a total of 80 HIV-1 subtype C infected participants failing a PI inclusive treatment regimen was genotyped. Mutations in PR and Gag known to be associated with PI resistance were identified. Novel Gag mutations possibly associated with PI resistance were identified by comparing Gag sequences from 80 HIV-1 subtype C infected individuals failing a PI inclusive treatment regimen (PCS cohort) to Gag sequences from: 54 HIV-1 subtype C acutely infected individuals, 954 HIV-1 subtype C treatment naïve individuals and 2,481 HIV-1 subtype B treatment naïve individuals. Once novel mutations were identified, we determined if these occurred in the absence or presence of PR DRMs.

2.2 Methods

2.2.1 Study participants/ sequences

Study samples comprised of 80 samples from patients failing a PI inclusive treatment regimen, 2,481 HIV-1 subtype B treatment naïve sequences, 954 HIV-1 subtype C treatment naïve sequences and 54 HIV-1 subtype C sequences from acutely infected individuals, as discussed below.

2.2.1.1 HIV-1 subtype C infected participants failing a PI inclusive treatment regimen (PCS cohort)

Stored samples were obtained from 80 patients failing a PI inclusive treatment regimen. These patients were enrolled in the Protease Cleavage Site (PCS) study between 2009 and 2013 and were recruited from King Edward VIII and McCords hospitals in Durban, South Africa. All participants received a PI- inclusive treatment regimen for at least six months and had a viral load >1000 copies/ml at the time of enrolment. Participant characteristics are presented in Table 2.1.

Viral load data was available for patients at virologic failure as part of routine clinical assessment. The Roche Amplicor version 1.5 assay (Roche Molecular Systems, Branchburg, New Jersey) was used to measure viral load. Adherence to cART was assessed by measuring levels of LPV of all participant samples using a mass spectrometry method validated and developed by the Division of Clinical Pharmacology, University of Cape Town, SA (34).

The protocol for this study was approved by the biomedical research ethics committee of the University of Kwa-Zulu Natal (BREC: BE347/13) (Appendix 7.3). Written informed consent was obtained from all study participants.

2.2.1.2 Control groups

Control groups comprised of sequences from: HIV-1 subtype C acutely infected individuals, HIV-1 subtype C treatment naïve individuals and HIV-1 subtype B treatment naïve individuals as discussed below.

2.2.1.2.1 HIV-1 subtype C acute sequences

Fifty-four Gag-Protease sequences from HIV-1 subtype C acutely infected individuals were obtained from Dr Jaclyn Mann (HIV Pathogenesis Programme, University of Kwa-Zulu Natal). Accession numbers for 32 sequences are available (HQ696791-HQ696822), whilst the remaining sequences constitute a dataset which has not been included in GenBank as yet due to ongoing work and analysis. The median number of days post infection for all acute samples was 14 days (range: 14–31.75 days) (35).

2.2.1.2.2 HIV-1 subtype C and B treatment naïve sequences

A total of 954 plasma sequences from treatment naïve individuals infected with HIV-1 subtype C from South Africa were downloaded from the Los Alamos sequence database (36). Additionally, 2,481 plasma sequences from treatment naïve HIV-1 subtype B infected individuals were downloaded from the Los Alamos sequence database (36). Accession numbers for all sequences from Los Alamos are provided in Appendix 7.1 and 7.2. The number of sequences which were used in this study as control groups represent all the

sequences found using the specific search criteria (i.e. all sequences had to be derived from plasma and had to be treatment naïve) that were available in Los Alamos.

All sequences were aligned using ClustalX version 2.1 (37) and manually edited in Bioedit version 7.2.5 (38). Duplicate sequences were removed using the ElimDupes tool in Los Alamos (39). The Rega subtyping tool was used to confirm the subtype of each sequence (40, 41).

An overview of all samples/sequences used in this study is depicted in Figure 2.1.

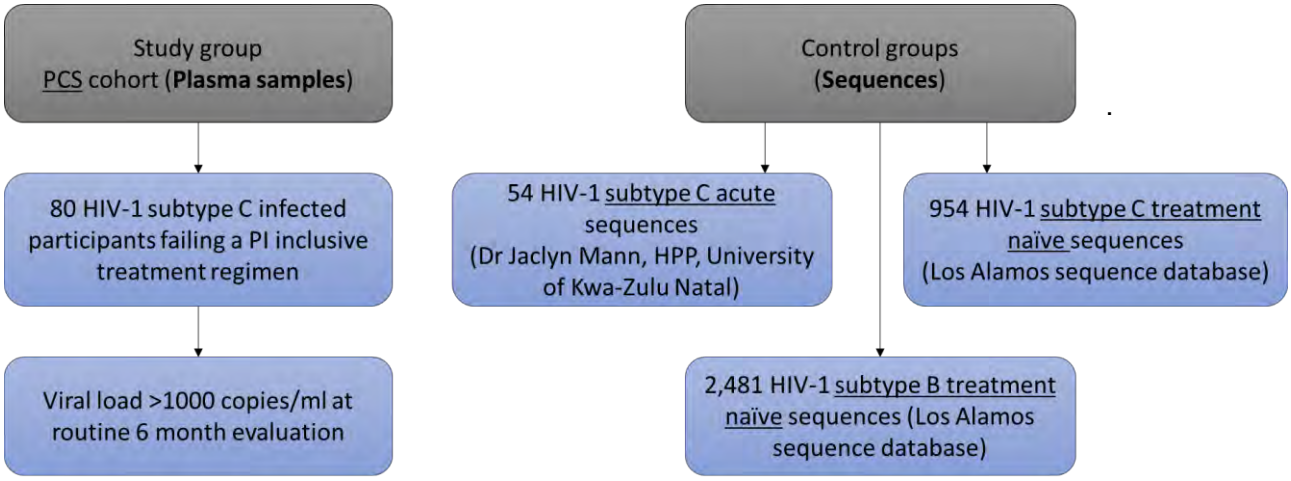


Figure 2-1 Overview of study participants and control groups utilized in this study. Study samples comprised of 80 participants failing a PI inclusive treatment regimen (PCS cohort), whilst the control group comprised of: 54 sequences from HIV-1 subtype C acutely infected individuals, 954 sequences from HIV-1 subtype C treatment naïve individuals and 2,481 sequences from HIV-1 subtype B treatment naïve sequences. Accession numbers for all sequences downloaded and used in this study is provided in Appendix 7.1 and 7.2.

All sequencing associated laboratory work, described below, for patients failing a PI inclusive treatment regimen (i.e. PCS cohort) was conducted in collaboration with MSC student K. Pillay (HIV Pathogenesis Programme, University of Kwa-Zulu Natal) (42).

2.2.2 RNA Extraction

Viral RNA from 80 PCS samples was extracted from 140 μl of stored plasma using the Qiamp Viral RNA Mini Kit, as per manufacturer's instructions (Qiagen, Valencia, USA). Prior to commencement of manufacturer's instructions, 500 μl of stored plasma was

centrifuged at 25,000 RCF (Jouan MR23i, Thermo Scientific, USA) for 90 minutes at 4°C in order to concentrate the plasma. Thereafter, as per manufacturer's instructions, 360 μl of supernatant was removed and 560 μl of lysis solution comprising of carrier-RNA[16] and buffer AVL[17] was added to the tube and incubated at room temperature for 10 minutes. This served to concentrate RNA and denature RNase molecules thereby facilitating the recovery of intact viral RNA.

A total of 560 μl of 96% ethanol was then added to each tube, to remove all degraded RNA. Tubes were vortexed for 5 seconds and centrifuged at 13,000 rpm's for 10 seconds. Supernatant was added to mini-columns and centrifuged at 13,000 rpm's for 1 minute. Thereafter 500 μl of buffer AW1 was added to each column and centrifuged at 13,000 rpm for 1 minute. This was followed by the addition of 500 μl of buffer AW2 to the mini-column, which was subsequently centrifuged at 13,000 rpm's for 1 minute. Viral RNA was then eluted in 50 μl of buffer AVE and stored at -80°C.

2.2.3 Amplification of the Gag-Protease region of HIV-1

Amplification of patient-derived Gag-Protease was performed using a previously described method (43) with minor amendments. Briefly, extracted viral RNA was reverse transcribed using the Superscript III one-step PCR kit (Invitrogen, Carlsbad, USA) and the following Gag-Protease specific primers: 5' GAG ATC TCT CGA CGC AGG AC 3' (HXB2 nucleotide: 675 to 697, forward primer) and 5' GGA GTG TTA Tat GGA TTT TCA GGC CCA ATT 3' (HXB2 nucleotides: 2,696 to 2,725, reverse primer). The reverse transcription PCR (RT-PCR) reaction comprised of: 3 μl diethyl pyrocarbonate (DEPC) treated water, 12.5 μl 2X reaction buffer (containing 0.4 mM of each dNTP and 3.2 mM $MgSO_4$), 0.5 μl of each primer (10 pmol/μl), 1 μl of Superscript III RT/Platinum Taq enzyme (2 U/μl) and 7.5 μl of RNA template. The RT-PCR reaction mix was incubated (GeneAmp PCR system 9700, Applied Biosystems, Foster City, USA) at 55ºC for 30 minutes (cDNA synthesis) and 94ºC for 2 minutes (initial denaturation), followed by 35 cycles of 94ºC for 15 seconds (denaturation), 55ºC for 30 seconds (annealing) and 68ºC for 2 minutes (extension), and ended with a 5 minute incubation at 68ºC (final extension).

16Carrier-RNA has two functions. Firstly it ensures maximum binding of RNA to the mini-column membrane thereby concentrating RNA within the column. Secondly, it prevents degradation of viral RNA by RNase molecules, which may have escaped degradation.

17Buffer AVL comprises of detergents and chaotropic salts that function to denature RNase molecules thereby facilitating the isolation of intact viral RNA only.

Reverse transcription-PCR was followed by a second round of PCR using a Takara Ex Taq HS enzyme kit (Takara, Shiga, Japan) and 100-mer forward (5' GAC TCG GCT TGC TGA AGC GCG CAC GGC AAG AGG CGA GGG GCG ACT GGT GAG TAC GCC AAA AAT TTT GAC TAG CGG AGG CTA GAA GGA GAG AGA TGG G 3') and reverse (5' GGC CCA ATT TTT GAA ATT TTT CCT TCC TTT TCC ATT TCT GTA CAA ATT TCT ACT AAT GCT TTT ATT TTT TCT GTC AAT GGC CAT TGT TTA ACT TTT G 3') primers that were complementary to NL43Δ*gag-protease* on either side.

Two 50 µl PCR reactions were prepared per participant sample, each comprising of 37 µl DEPC water, 5 µl of 10X Ex Taq buffer (consisting of 20 mM MgCL2), 4 µl of dNTP's, 0.8 µl of each primer (10 pmol/µl), 0.25 µl of Takara Ex Taq (5 U/µl) and 2 µl of RT-PCR product. The following thermocycler conditions were used: 94ºC for 2 minutes (initial denaturation), 40 cycles of 94ºC for 30 seconds (denaturation), 60ºC for 30 seconds (annealing) and 72ºC for 2 minutes (extension) followed by a 7 minute hold at 72ºC (final extension).

Amplification of the Gag-Protease region was confirmed by gel electrophoresis. A 1 % agarose gel was prepared by adding two agarose tablets (0.5 g) (Bioline, USA) to 100 ml of 1X Tris/Borate/EDTA solution (Sigma-Aldrich, USA) which was then heated to dissolve the agarose and cooled to room temperature prior to casting the gel. Gel loading dye was prepared by mixing 50 µl of gel loading buffer (Sigma-Aldrich, USA) to 1 µl of gel red (Sigma-Aldrich, USA). A total of 2 µl of gel loading dye (Sigma-Aldrich, USA) was mixed with 2 µl of PCR product and loaded onto the gel which was run at 120 V , 500 mA for 40 minutes on an Electrophoresis Power Supply (EPS 301, Amersham Biosciences, Sweden). A low DNA mass ladder (2 µl) (Invitrogen, Carlsbad) (Figure 2.2) was also run on the gel to assist in viewing the approximately 1.7 kb Gag-Protease product (Figure 2.2). The GelVue UV Transilluminator (SynGene, London) was used to view the gel.

A picture of the low mass ladder used in this study is presented in Figure 2.2 below.

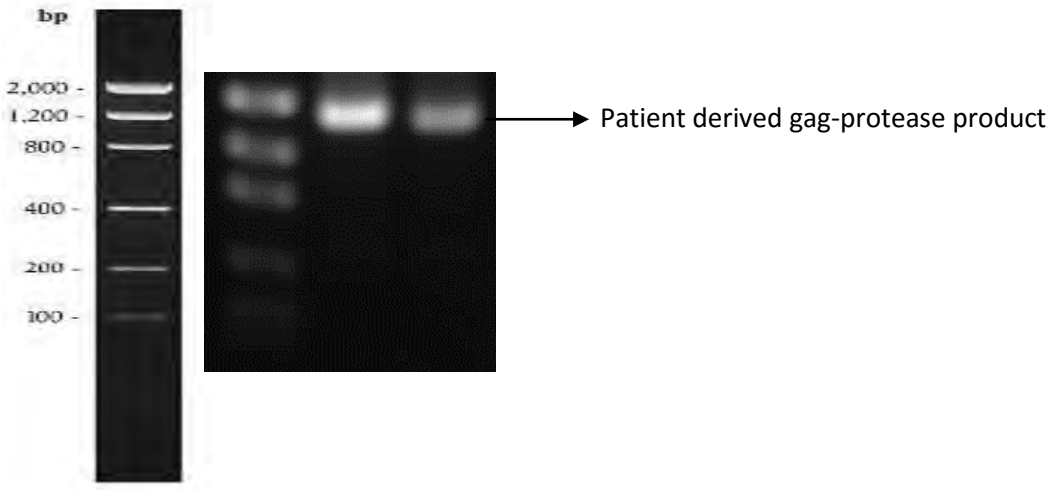


Figure 2-2 Depiction of the Invitrogen low DNA mass ladder used to identify Gag-pro products which were approximately 1.7kb as demonstrated by the gel showing patient-derived gag-protease product on the right.

Once the presence of Gag-Protease product was confirmed, PCR products from the two 50 µl reactions were pooled and stored at -20ºC until used in Sanger sequencing and generation of recombinant viruses (Chapter 3).

2.2.4 Sanger sequencing of HIV-1 Gag-Protease amplicons

Sanger sequencing was conducted using the BigDye terminator kit v3.1 (Applied Biosystems, Foster City, CA). A 1:10 dilution of each PCR product (generated in section 2.2.3) was prepared in PCR grade water. Sequencing primers used included: 5' CTT GTC TAG GGC TTC CTT GGT 3' (HXB2 position: 1,078 – 1,098), 5' CTT CAG ACA GGA ACA GAG GA 3' (991 – 1010), 5' GGT TCT CTC ATC TGG CCT GG 3' (1462 – 1481), 5' CCT TGC CAC AGT TGA AAC ATT T 3' (1960 – 1981), 5' TAG AAG AAA TGA TGA CAG 3' (1817 – 1834), 5' CTA ATA CTG Tat CAT CTG CTC CTG T 3' (2328 – 2353), 5' CCT GGC TTT AAT TTT TAC TGG 3' (2,196 – 2,268). Individual sequencing reactions per primer were prepared in a 96-well Micro Amp plate (Applied Biosystems) with the following components: 1.6 µl PCR grade water, 2 µl 5X sequencing buffer, 1 µl primer (3.2 pmol/µl), 0.4 µl BigDye RR mix and 5 µl of diluted PCR product (approximately 2 ng/µl). Thermal cycling conditions were as follows: 96ºC for 1 minute (initial denaturation) followed by 35 cycles of 96ºC for 10 seconds (denaturation), 50ºC for 30 seconds (annealing) and 60ºC for 4 minutes (extension).

Purification of sequencing products was conducted on the same day. Products were initially diluted in 1 µl ethylenediaminetetraacetic acid (EDTA; 125 mM, pH 8.01, Sigma-Aldrich) and 26 µl of a sodium acetate mixture (comprising of 1 µl sodium acetate [3M, pH 5.2, Sigma-Aldrich] and 25 µl of 99% ethanol). The plate was sealed, vortexed for 5 seconds and centrifuged (Eppendorf centrifuge 5810R, Merck, Germany) at 3,000 x g for 20 minutes. Sequencing products were then dried by inverting plates on a paper towel and centrifuging at 150 x g for 5 minutes. Thereafter pellets were immediately re-suspended in 35 µl of a 70% ethanol solution and centrifuged at 3,000 x g for 5 minutes. Plates were once again inverted and centrifuged at 150 x g for 1 minute. Afterwards, the plate was dried at 50ºC for 5 minutes and stored at -20ºC. Prior to sequencing on the ABI 3130 XL genetic analyser (Applied Biosystems), sequencing products were re-suspended in 10µl HiDi-formamide, mixed and denatured at 95ºC for 3 minutes and cooled to 4ºC for 3 minutes in a thermocycler.

2.2.5 Data analysis

An overview of analysis tools used to analyse Gag-Protease sequences and identify novel Gag mutations is discussed below. Statistical analysis used in this study is also discussed.

2.2.5.1 Sequence analysis

Sequences were edited and visualized in Sequencher version 5.1 (Gene Codes Corporation, Ann Arbor, USA) and aligned to an HXB2 reference strain (GenBank accession number K03455) and a subtype C reference strain (GenBank accession number AY772699) using ClustalX version 2.1 (37) and Bioedit version 7.2.5 (38). For the purpose of quality control, a neighbor joining tree was drawn in Paup (44) and viewed in Figtree (http://tree.bio.ed.ac.uk/software/figtree/). Contamination was evaluated based on branch lengths viewed in the phylogenetic tree. The Rega HIV-1 subtyping tool was used to confirm sample subtypes (40, 41).

Mutations in PR were identified using the Stanford HIVdb resistance interpretation algorithm (45) and the International AIDS Society-USA (IAS-USA) list of resistance mutations (46). The presence of mutations in Gag previously reported to be associated with PI resistance/exposure was identified based on information from literature (Table 1.3) (4, 21, 31).

2.2.5.2 Analysis of novel mutations

Since Gag mutations in HIV-1 subtype C are not well documented, exploratory analysis was conducted, using the RegaDB sequence analysis tool (http://regatools.med.kuleuven.be/sequencetool/sequencetool.wt) and the Viral Epidemiology Signature Pattern Analysis (VESPA) tool (http://www.hiv.lanl.gov/content/sequence/VESPA/vespa.html) in order to identify possible mutations in Gag associated with PI exposure or PI drug resistance.

The RegaDB sequence analysis tool compared each sequence in the PCS cohort and HIV-1 subtype C control groups (i.e. acute and treatment naïve groups) to the reference sequence (HXB2) and scored each amino acid with a 1 (variation present) or a 0 (variation absent). A subtype C reference sequence was also included in the analysis. The program generated a table detailing the presence/absence of variations at each Gag codon. Each codon variation was summed and frequency was calculated by dividing the sum by the total number of sequences in the group. The frequency of variants at each codon was compared amongst sequences from: participants failing a PI inclusive treatment regimen (PCS cohort), HIV-1 subtype C treatment naïve individuals and HIV-1 subtype C acutely infected individuals. Variants with a difference of >5% amongst cohorts were flagged. The frequencies of these variants, in all 3 cohorts, was anlaysed using the Chi square statistic.

The VESPA tool computed the frequency of amino acids at each position for two groups of sequences tested. It selected the position at which the most common amino acid in one group differs from the most common amino acid in the second group and output the frequencies of these amino acids for each group (47). Fishers exact tests were performed for all positions identified as signature patterns. Two sets of VESPA analysis was conducted. The first compared the viral signature patterns between the PCS cohort and the HIV-1 subtype C treatment naïve cohort. The second analysis compared viral signature patterns between the PCS cohort and the acute cohort.

For amino acid substitutions to be considered as being associated with PI resistance/exposure, they were required to appear at a significantly higher frequency in the PCS cohort in comparison to the acute and treatment naïve cohorts.

2.2.6 Statistical analysis

Discrete data (e.g. mutation frequency) was analysed using the Fishers exact tests (for comparisons of 2 groups) or the Chi-square test (for comparisons of more than 2 groups). Multiple comparisons of mutation frequency were subject to Bonferonni's correction (48). Statistical analysis was conducted using Graphpad Prism version 5 (Graphpad Software, California) and R statistics version 3.2.2 (49). Statistical significance was defined as $p < 0.05$. All continuous data (e.g. age, viral load and LPV levels) is presented as the median with an interquartile range (IQR) unless otherwise stated.

2.3 Results

2.3.1 Participant characteristics

Table 2.1 details participant characteristics for 80 participants from the PCS cohort. Of the 80 participants, 40% were male whilst 60% were female. The median age at sampling was 35 years (IQR: 26-42 years). The median viral load was 4.63 $\log_{10}$ copies/ml (IQR: 3.98 – 5.26). Lopinavir levels were detectable in 61 participants (76.25%), a further 10 participants had undetectable LPV levels (12.5%) whilst LPV levels were not available for 9 participants (11.25%).

Table 2-1 Overview of participant characteristics for 80 participants, infected with HIV-1 subtype C, that failed a PI inclusive treatment regimen (i.e. PCS cohort).

Characteristic	Value (IQR) of parameter
Gender (%) **Male** **Female**	 40 60
Median age at sampling (years)	35 (26 – 42)
Median viral load at sampling ($\log_{10}$ copies/ml)	4.63 (3.98 – 5.26)
Lopinavir levels (µg/ml) **Detectable (n=61)** **Undetectable (n=10)** **No data (n=9)**	11.7 (5.9 – 17.8)

All 80 participants were infected with HIV-1 subtype C as established by the RegaDB sequencing tool and depicted in the neighbour joining tree in Figure 2.3 below.

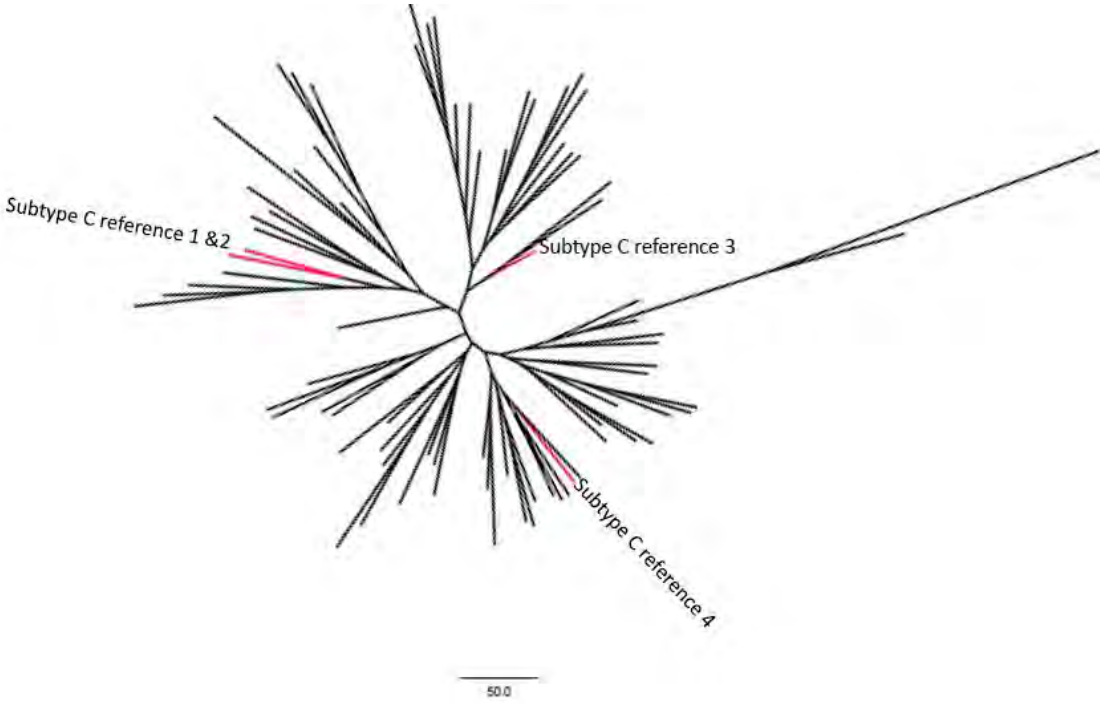


Figure 2-3 Neighbor joining tree (Paup 4.0) for PCS cohort participants. All sequences from the PCS cohort clustered with HIV-1 subtype C reference sequences indicating that all sequences are representative of HIV-1 subtype C infection.

2.3.2 Identification of Protease DRMs in the PCS cohort

As described in an associated study, 34 out of 80 participants from the PCS cohort presented with PR RAMs (i.e. 42.5%) (42).

Figure 2.4 shows each of the 18 PR RAMs identified and lists the number of participants harbouring each RAM next to the mutation name. The I54V PR RAM occurred most frequently (29 out of 34), this was followed by the M46I (28 out of 34), V82A (27 out of 34) and L76V PR RAMs (22 out of 34) respectively. All other PR RAMs occurred in <4 participants (Figure 2.4)

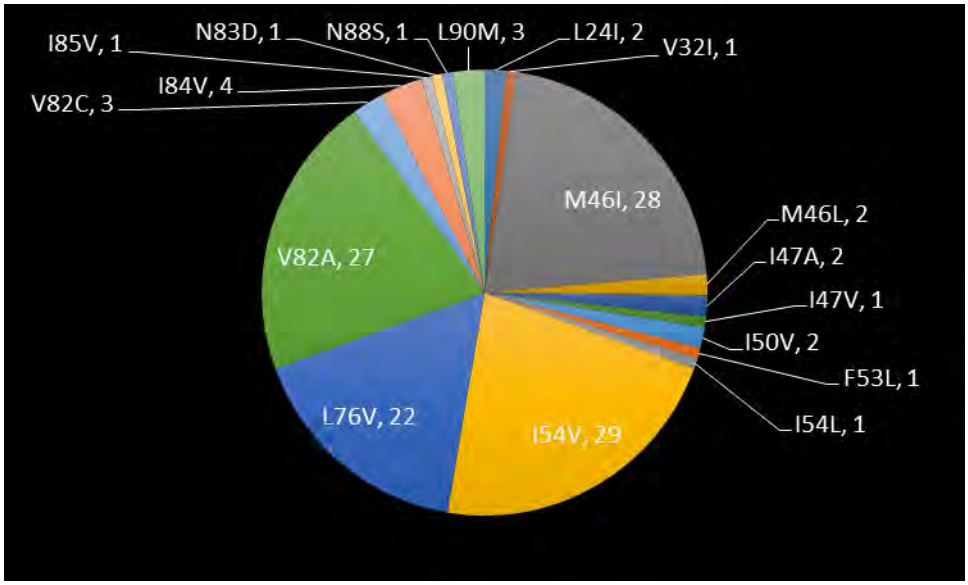


Figure 2-4 Overview of 18 PI associated resistance mutations in PR identified in the PCS cohort and the number of sequences in which each mutation was detected. Mutation names are followed by the number of participants presenting with the particular mutation.

Ninety-four percent of participants with PR RAMs (32 out of 34) harboured more than one PR mutation (Table 2.3). The majority of participants harboured combinations of four PR RAMs (53%; 18 out of 34), 17% of participants harboured combinations of three PR RAMs (6 out of 34), 15% of participants harboured combinations of five PR RAMs (5 out of 34), 6% of participants harboured combinations of either two or six PR RAMs (2 out of 34) whilst 3% of participants harboured one PR RAM only (1 out of 34) (Table 2.2).

The V82A PR RAM was the only PR mutation found in isolation of other PR RAMs in two participants (Table 2.2). The most common pattern of PR RAMs included: M46I+I54V+V82A (highlighted in red in Table 2.2) which occurred in 22 out of 33 participants (67%). Additionally, 16 out of 34 participants presented with the following pattern of PR mutations: M46I+I54V+V82A+L76V (highlighted in blue in Table 2.3) (47%) (Table 2.2). Most other participants presented with varied combinations of PR RAMs, with no specific pattern noted (Table 2.2).

Table 2-2 Illustration showing the combinations of PR RAMs for participants from the PCS cohort.

Number of PR RAMs[a]	Mutation combinations and number of participants with combination							Total Ptnts[b]
1	V82A (n=2)							2
2	I54V, V82A (n=1)							1
3	I54V, L76V, V82A (n=1)	M46I, I47A, I84V (n=1)	M46I, I54V, V82A (n=3)	M46L, I47A, N88S (n=1)				6
4	L24I, M46I, I54V, V82A (n=1)	M46I, I50V, I54V, L90M (n=1)	M46I, I50V, I54V, V82A (n=1)	M46I, I54L, L76V, I84V (n=1)	M46I, I54V, L76V, V82A (n=12)	M46I, I54V, V82A, I85V (n=1)	M46L, I54V, L76V, V82A (n=1)	18
5	L24I, M46I, I54V, L76V, V82A (n=1)	M46I, I47V, I54V, L76V, V82C (n=1)	M46I, I54V, L76V,V82A, N83D (n=1)	M46I, I54V, L76V, V82C, I84V (n=2)				5
6	M46I, F53L, I54V,L76V,V82A, L90M (n=1)	V32I, M46I,I54V,L76V,V82A, L90M (n=1)						2

[a] Represents the number of Protease resistance associated mutations (PR RAMs). [b] Represents the total number of participants with a particular number of PR mutations. Mutations highlighted in red represent the most common mutational pattern (M46I+I54V+V82A). Mutations highlighted in blue represents the addition of a 4th mutation (i.e. L76V) to the initial mutation combination (i.e. M46I+I54V+V82A+L76V), making it the second most commonly occurring mutation combination in the cohort. The number of participants (n) with a particular mutation or group of mutations is presented in brackets within each cell. Abbreviations: PR – Protease; RAMs – resistance associated mutations; ptnts – participants; n – number.

These data suggests that it is uncommon for PR RAMs to occur in isolation of other PR RAMs.

2.3.3 Identification of known Gag mutations associated with PI resistance/exposure in the PCS cohort

A list of 43 Gag mutations collated from existing literature (Chapter 1, Table 1.3) was used as a reference to identify mutations in Gag that have been reported to be associated with PI resistance or PI exposure.

These 43 mutations were divided into 2 groups (i.e. 11 mutations associated with PI resistance [i.e. rGag mutations] and 32 mutations associated with PI exposure) and their frequencies compared amongst sequences from the: PCS cohort, HIV-1 subtype C treatment naïve cohort, HIV-1 subtype C acute cohort and an HIV-1 subtype B treatment naïve cohort (Figure 2.5 and Figure 2.6).

Figure 2.5 details the frequencies of known Gag mutations associated with PI resistance in the: PCS cohort (black bars), HIV-1 subtype C treatment naïve cohort (green bars), HIV-1 subtype C acute cohort (acutes; red bars) and HIV-1 subtype B treatment naïve cohort (blue bars). Eight out of 11 Gag mutations known to be associated with PI resistance were detected in the PCS cohort (Table 2.1 and Figure 2.5). These included: R76K, Y79F, V128I, A431V, K436R, L449F/P, R452K and P453L which will be discussed below.

The R76K and Y79F mutations occurred at high frequencies (>50% and >35% respectively) in all cohorts, suggesting that they are potentially natural polymorphisms in HIV-1 subtype C and possibly in HIV-1 subtype B as well.

The A431V, L449F/P, R452K and P453L resistance associated Gag (rGag) mutations occurred at significantly higher frequencies in the PCS cohort in comparison to all other cohorts ($p<0.0001$ for all). This indicates that these mutations are likely associated with PI resistance in HIV-1 subtype C, as is the case with HIV-1 subtype B (4).

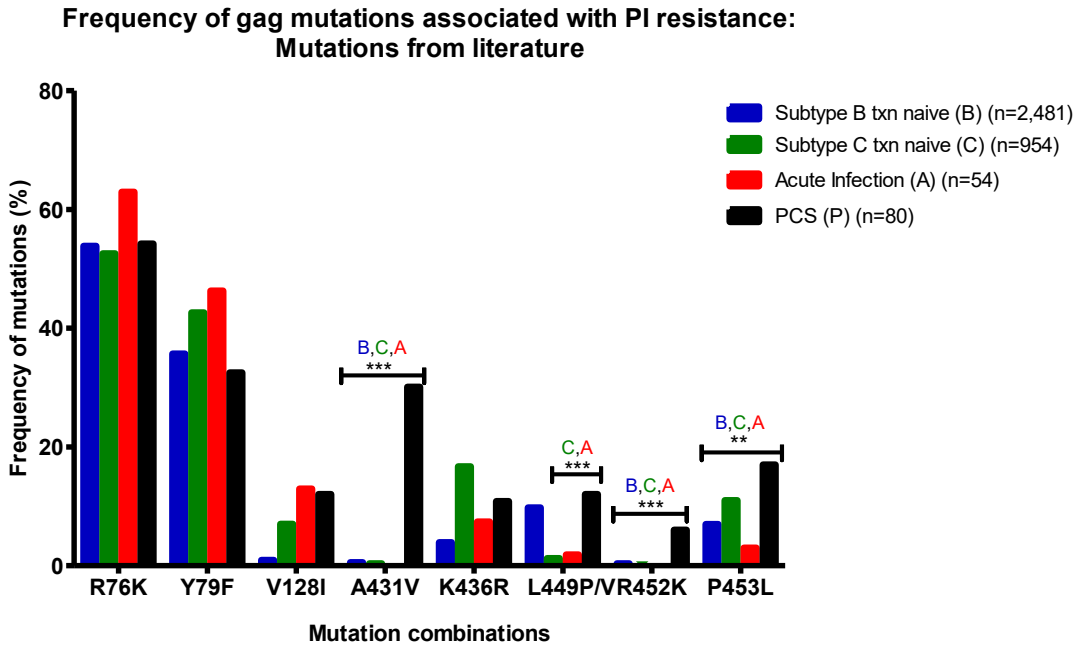


Figure 2-5 Comparison of Gag mutations reported to be associated with PI resistance, amongst sequences from the: PCS cohort, HIV-1 subtype C treatment naïve cohort, HIV-1 subtype C acute cohort and HIV-1 subtype B treatment naive cohort. Significant differences are denoted by asterisks, where *** = p<0.0001 and ** = p<0.001. Color coded alphabets (which match the color of bars) located above the p value are representative of the group which presented with the significantly lower result. E.g. the A431V mutation occurred at a significantly lower frequency in the HIV-1 subtype B treatment naïve cohort (B), HIV-1 subtype C treatment naïve cohort (C) and the HIV-1 subtype C acute cohort (A) in comparison to the PCS cohort.

Figure 2.6 provides an overview of the frequencies of Gag mutations reported to be associated with PI exposure within the: PCS cohort (black bars), HIV-1 subtype C treatment naïve cohort (green bars), HIV-1 subtype C acute cohort (acutes; red bars) and HIV-1 subtype B treatment naïve cohort (blue bars).

Of the 32 PI exposure associated mutations identified in literature (Chapter 1, Table 1.3), 12 were identified in the current study. The majority of these mutations occurred in the p2/NC cleavage site (Figure 2.5). Mutations for which significant differences were observed are discussed below.

The E12K, V370A/M, T375N and G381S RAMs occurred at significantly higher frequencies in subtype C sequences (PCS, treatment naïve and acutes) in comparison to sequences from the HIV-1 subtype B cohort. Their high frequency in the acute and HIV-1

treatment naïve cohort could suggest that these mutations possibly occur as natural polymorphisms in HIV-1 subtype C.

The V35I and I401T mutations were the only mutations to occur at a significantly higher frequency in the PCS cohort in comparison to other cohorts indicating that these mutations are likely to be related to PI exposure/resistance in HIV-1 subtype C.

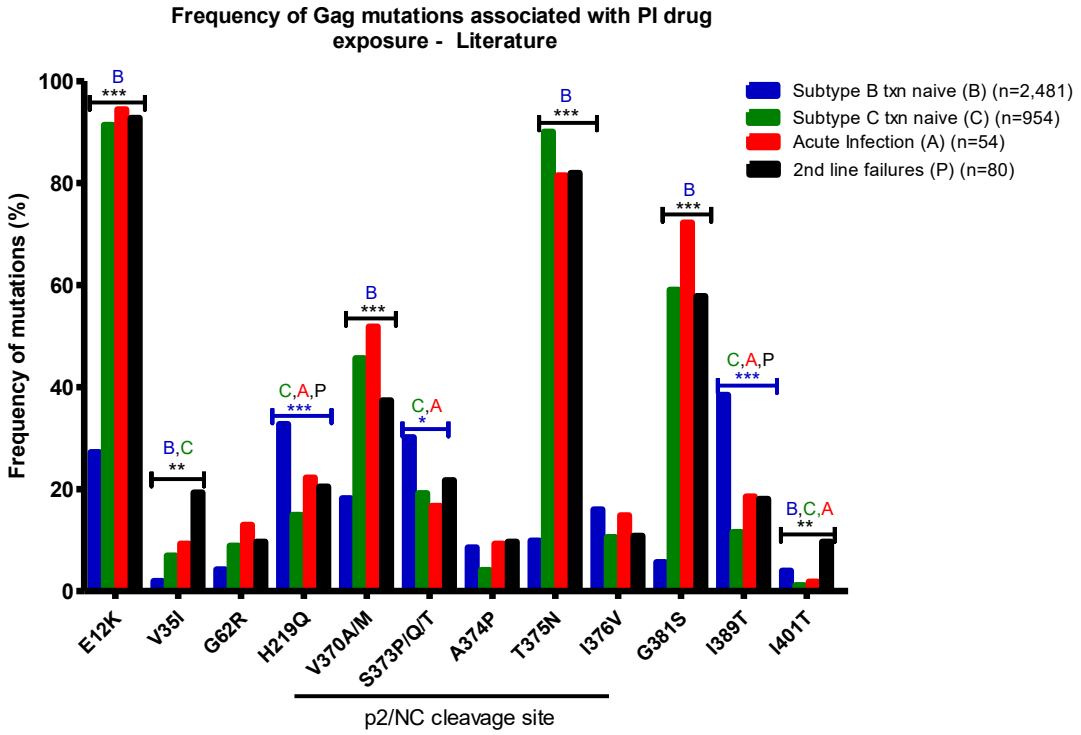


Figure 2-6 Comparison of Gag mutations reported to be associated with PI exposure, amongst sequences from the: PCS cohort, HIV-1 subtype C treatment naïve cohort, HIV-1 subtype C acute cohort and HIV-1 subtype B treatment naive cohort. Significant differences are denoted by asterisks, where *** = p<0.0001, ** = p<0.001 and * = p<0.01. Color coded alphabets (which correspond to the color of bars) located above the p value are representative of the group which presented with the significantly lower result. E.g. the V35I mutation occurred at a significantly lower frequency in the HIV-1 subtype B treatment naïve cohort (B) and the HIV-1 subtype C treatment naïve cohort (C) in comparison to the PCS cohort.

In summary, of the 43 Gag mutations shown to be associated with PI resistance/exposure in literature (Chapter 1, Table 1.3), 20 were detected in the PCS cohort (Figure 2.5 and 2.6). Of these 20 mutations, the A431V, L449F/P/V, R452K, P453L, V35I and I401T rGag mutations were the only mutations to occur at significantly higher frequencies in the PCS cohort in comparison to other cohorts (Figure 2.5 and 2.6), suggesting that these

mutations are associated with PI exposure/resistance in HIV-1 subtype C. The presence of rGag mutations (R76K and Y79F) and PI exposure associated Gag mutations (E12K, V370A/M, T375N and G381S) at frequencies above 50% and 35% respectively, in the acute and HIV-1 subtype C treatment naïve cohorts highlights the possibility of PI resistance/exposure associated mutations occurring as natural polymorphisms in HIV-1 subtype C.

2.3.4 Identification of novel Gag mutations associated with PI resistance/exposure

Exploratory analysis was conducted to identify novel Gag (nGag) mutations associated with PI resistance/exposure in HIV-1 subtype C, using the RegaDB sequence analysis tool and the VESPA tool.

2.3.4.1 RegaDB data analysis

The frequency of amino acids occurring at each codon in Gag was compared between the 80 sequences from participants failing a PI inclusive treatment regimen (PCS) and sequences from: 954 HIV-1 subtype C treatment naïve and 54 HIV-1 subtype C acutely infected individuals using the RegaDB sequence analysis tool (Figure 2.7).

Four novel Gag amino acid substitutions (Q69K, S111C/I, T239A/S and I256V) were identified to occur at a significantly higher frequency in the PCS group than the HIV-1 subtype C treatment naïve and acute infection groups (Q69K, $p<0.0001$; S111C/I, $p<0.0001$; T239A/S, $p<0.0001$ and I256V, $p<0.01$) (Figure 2.7). All four of these mutations occurred in the amino terminal of Gag with two occurring in MA (Q69K and S111C/I) and two occurring in CA (T239A/S and I256V). (Figure 2.7).

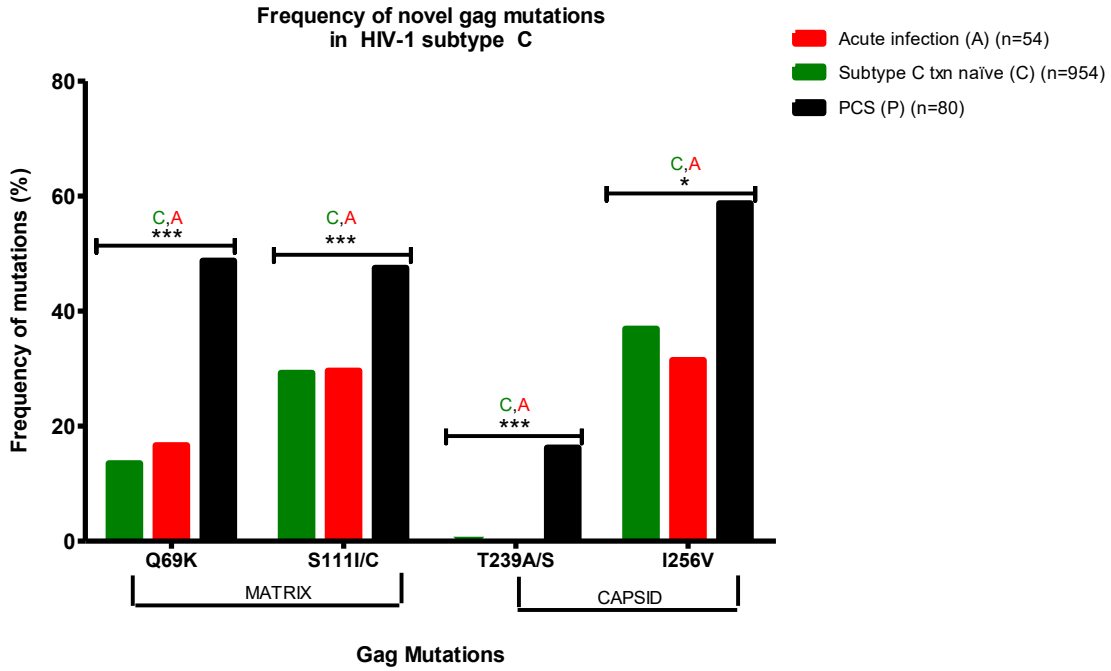


Figure 2-7 Overview of four novel Gag mutations identified using the RegaDB sequence analysis tool. Significant differences are denoted by asterisks, where *** = p<0.0001, ** = p<0.001 and * = p<0.01. Color coded alphabets located above the p value are representative of the group which presented with the significantly lower result. E.g. the S54A mutation occurred at a significantly higher frequency in the PCS cohort and the acute cohort in comparison to the subtype C treatment naïve cohort (C).

2.3.4.2 Vespa analysis

The difference in viral signature patterns amongst sequences from the: PCS cohort, HIV-1 subtype C infected treatment naïve cohort (Subtype C Txn Naïve) and HIV-1 subtype C acute infection cohort (acutes) is presented in Figure 2.8.

Four variations in signature patterns were noted between sequences from the PCS cohort, treatment naïve cohort and the acute infection cohort at position 61, 69, 256 and 451 of Gag. For position 61, the consensus amino acid for subtype B is leucine (L), however subtype C reference sequences generally have an isoleucine (I) at this position. Interestingly sequences from the PCS and acute cohort had predominantly isoleucine at position 61 whilst the majority of sequences in the treatment naïve cohort had a methionine (M) at this position (p<0.010 for both) (Figure 2.8). Similarly, both HIV-1 subtype B and C reference sequences have a Serine (S) at position 451 of Gag. Whilst

most sequences from the PCS cohort had serine at this position, sequences from both the treatment naïve ($p<0.001$) and acute (no significant difference) cohorts presented predominantly with asparagine (N) at this position. The substitution of a serine for asparagine at position 451 is documented to be associated with PI resistance/exposure (24, 31, 50) hence its predominance in the acute and treatment naïve cohort and subsequent lack in the PCS cohort was surprising.

The Q69K and I256V Gag mutations occurred at significantly higher frequencies in the PCS cohort in comparison to the treatment naive ($P<0.001$) and acute ($p<0.0001$) cohorts, suggesting that these mutations may be associated with PI exposure/resistance (Figure 2.8 and Figure 2.7).

Interestingly, the S111I/C and T239A/S mutations identified by RegaDB analysis to occur at a significantly higher frequency in the PCS cohort in comparison to the acute and treatment naïve cohort, did not appear as signature patterns for the PCS cohort in VESPA analysis. This could be attributed to the fact that S111I was combined with S111C and likewise T239A was combined with T239S in Rega analysis thereby increasing the number of participants with these mutations whilst VESPA analysed each amino acid change individually.

Overall, a combination of results from RegaDB and VESPA analysis highlighted four possible nGag mutations which could be associated with PI resistance/exposure in HIV-1 subtype C. These included: Q69K, S111I/C, T239A/S and I256V.

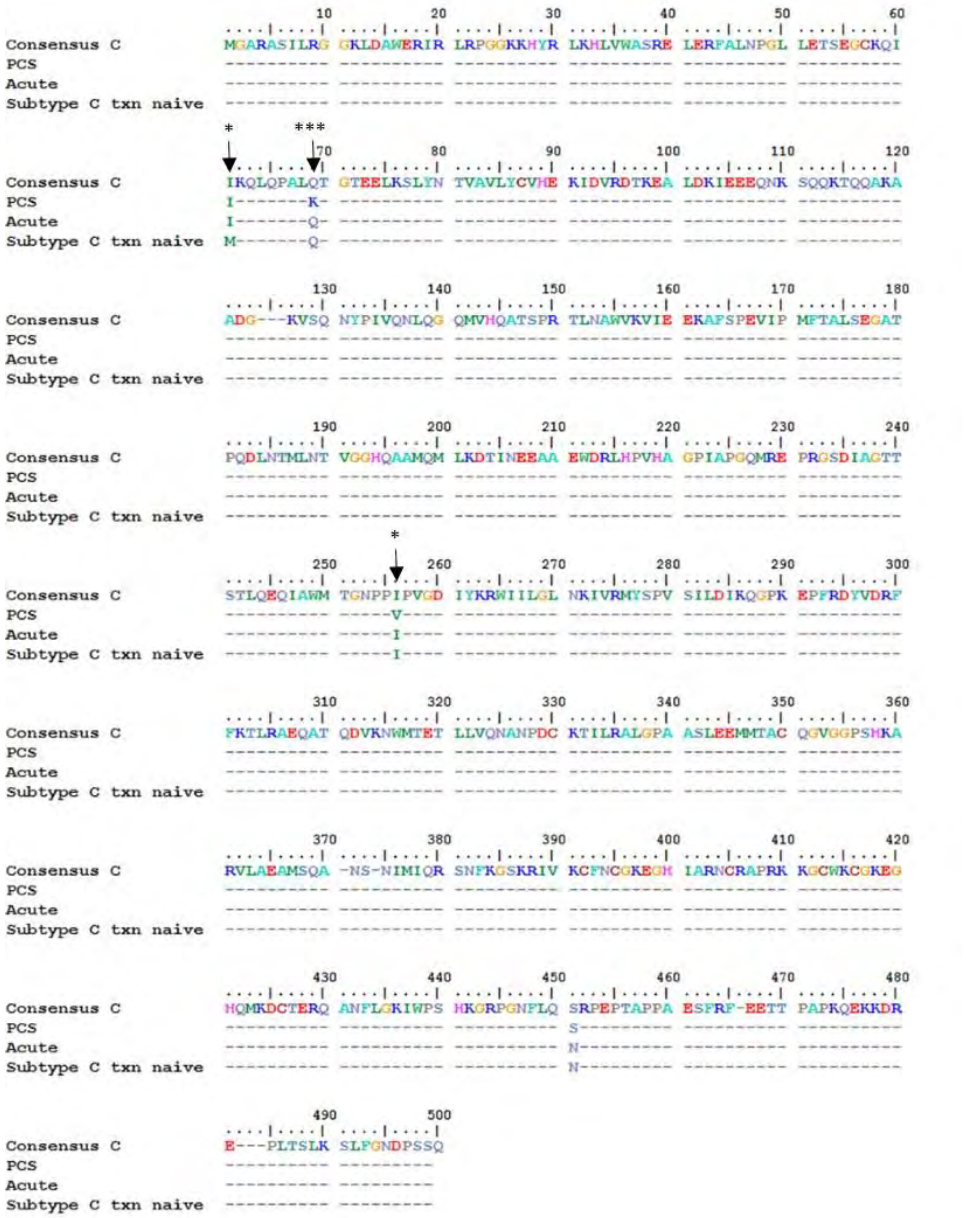


Figure 2-8 Viral signature patterns highlighting sequence variation amongst sequences from the PCS cohort, the HIV-1 subtype C treatment naïve cohort and the HIV-1 subtype C acute cohort. The first line represents amino acids belonging to the subtype C reference sequence. Dashes represent amino acids which correspond to the consensus sequences. Where amino acids do not correspond to the consensus, the variable amino acid is highlighted in colored text. The frequency of amino acid variations was compared using the Fishers exact statistic. Significance is denoted by asterisks, were *= p<0.01, **= p<0.001 and ***=p<0.0001.

2.3.5 Assessing frequency of rGag, nGag and Gag exposure associated mutations in sequences with PR RAMs

An overview of rGag mutations, PI exposure associated Gag mutations and nGag mutations is presented in Table 2.3 for all PCS sequences with PR RAMs. As described in section 2.3.2, the most common pattern of PR RAMs was M46I+I544V+V82A or M46I+I54V+L76V+V82A. The A431V rGag mutation was the most commonly occurring rGag mutation (i.e. present in 25 out of 34 participants, Section 2.3.3). All sequences with PR RAMs had nGag, rGag and Gag exposure associated mutations. The Q69K nGag mutation occurred most commonly (n= 21) followed by the S111I/C mutation (n=19).

Table 2.3 highlights that all participants with PR RAMs harboured at least 1 rGag, nGag and PI exposure associated gag mutation.

Table 2-3 Summary of: PR RAMs, Gag mutations identified in literature to be associated with PI exposure or resistance Gag (rGag) and novel Gag (nGag) mutations identified as part of this study in the PCS cohort.

PID	PR RAMs	Gag mutations associated with PI exposure	Gag mutations associated with PI resistance (rGag)	Novel Gag mutations (nGag)
PCS 002	V32I, M46I, I54V, L76V, V82A, L90M	E12K,H219Q, T375N, G381S	R76K, A431V	S111C, I256V
PCS 003	I54V, L76V, V82A	E12K,V370A, T375N,P497L	Y79F, K436R	S111C
PCS 017	M46I, I54V, L76V, V82A	V370A, T375N,P497L	R76K, A431V	Q69K, I256V
PCS 019	V82A	E12K, V370A, S373T, T375N, I376V,S451N,	L449F, R452K	I256V
PCS 033	M46I, F53L, I54V, L76V, V82A, L90M	E12K,V370A, S373T, T375N, G381S,S451T	R76K, K436R	Q69K, I256V
PCS 040	M46I, I54V, L76V, V82C, I84V	E12K, V370A, T375N, I389T	A431V	I256V,T239S
PCS 053	M46I, I54V, L76V, V82A, N83D	E12K, G62R, T375N, I389T, I401V	A431V, K436R	I256V, T239S
PCS 056	I54V, V82A	E12K	Y79F, V128I, K436R, P453L	Q69K, S111C
PCS 061	M46L, I54V, L76V, V82A	E12K, V370A, S373T, T375N, G381S, I401V,	R76K, V128I, A431V, L449F	Q69K, I256V, T239A
PCS 064	M46I, I54V, V82A, I85V	T375N,S451N	R76K, A431V	S111C
PCS 069	M46I, I54V, L76V, V82A	E12K,T375N, G381S,I401L	R76K, A431V	Q69K, S111C, I256V

PID	PR RAMs	Gag mutations associated with PI exposure	Gag mutations associated with PI resistance (rGag)	Novel Gag mutations (nGag)
PCS 071	M46I, I54V, L76V, V82A	E12K, H219Q, V370A, T375N, G381S	R76K, Y79F, A431V	Q69K, S111C, I256V
PCS 076	V82A	E12K,H219Q, T375N, I401V	R76K	Q69K, I256V
PCS 086	M46I, I47V, I54V, L76V, V82C	E12K, G62R,T375N, G381S, S451N	V128I	T239A
PCS 087	M46I, I54V, L76V, V82A	E12K,T375N, G381S, I389T	R76K, A431V	Q69K
PCS 089	M46I, I54V, L76V, V82A	E12K,S373T, A374P, G381S	R76K, A431V	S111C
PCS 090	M46I, I50V, I54V, V82A	E12K, T375N, I376V, I401L	A431V, P453L	Q69K
PCS 097	M46I, I50V, I54V, L90M	E12K, T375N, G381S,I437V	R76K, V128I, A431V, R452K	Q69K, S111C, I256V
PCS 098	M46I, I54V, L76V, V82C, I84V	E12K,H219Q, T375N, G381S, I389T	Y79F	I256V
PCS 099	M46L, I47A, N88S	E12K,H219Q, T375N, G381S	V128I,A431V,K436R	Q69K, T239A
PCS 100	M46I, I54V, L76V, V82A	E12K, T375N, G381S	Y79F, V128I, A431V	Q69K
PCS 101R	M46I, I54V, L76V, V82A	E12K, T375N, G381S,S451N, I401L	A431V	Q69K, S111C, I256V
PCS 108	M46I, I54V, L76V, V82A	E12K,G381S, I401V	R76K, A431V	Q69K, S111C, T239S
PCS 124	M46I, I54V, V82A	E12K,H219Q, V370A, S373T, T375N,,S451N	R76K, A431V	S111I, T239S
PCS 130	M46I, I54V, V82A	E12K, A374P, T375N,S51N	R76K, Y79F, A431V,P453L	Q69K, S111C
PCS 133	L24I, M46I, I54V, V82A	E12K,V370A, S373P	Y79F, K436R	Q69K, I256V
PCS 134	M46I, I54V, L76V, V82A	E12K, S373Q, A374S, T375N, G381S	R76K, Y79F,A431V	I256V, T239S
PCS 136	M46I, I54V, L76V, V82A	E12K, V370A, T375N, G381S,	R76K, Y79F,A431V, L449F	Q69K, S111C, I256V
PCS 140	M46I, I54V, L76V, V82A	E12K,T375N, I376V, G381S	R76K, Y79F, V128I,A431V	Q69K, S111I, I256V
PCS 152	M46I, I47A, I84V	E12K,V370A	R76K	S111C, I256V
PCS 153	L24I, M46I, I54V, L76V, V82A	E12K, H219Q, V370A, T375N, G381S, S451N	R76K, Y79F	Q69K, I256V
PCM 001	M46I, I54L, L76V, I84V	E12K, V370A, T375N, I376V, G381S	R76K, Y79F,A431V	Q69K, S111C, I256V, T239S
PCM 002	M46I, I54V, V82A	E12K, T375N, I376V, G381S,S451N	R76K, Y79F,A431V	S111C
PCM 007	M46I, I54V, L76V, V82A	E12K, V370A, T375N, G381S	A431V	Q69K, S111C

2.3.6 Assessing frequency of Gag mutations in PCS sequences without PR RAMs

Lastly, the presence of rGag, nGag and Gag PI exposure associated mutations was assessed in sequences without PR RAMs (n=46) (Table 2.4). A total of 33 out of 46 sequences (71.7%) without PR RAMs harboured rGag mutations (Table 2.4). Thirty-seven out of 46 sequences (80%) harboured at least one nGag mutation identified in this study, whilst all 46 sequences (100%) harboured at least one Gag mutation known to be associated with PI exposure.

All rGag mutations, except A431V, were found at similar frequencies in sequences with PR RAMs and sequences without PR RAMs. The A431V rGag mutation was found at a significantly higher frequency in sequences with PR RAMs (n=24) than sequences without PR RAMs (n=1) ($p<0.0001$) (Table 2.4). These data show that rGag, nGag and Gag PI exposure associated mutations also occur independently of PR RAMs.

Table 2-4 Summary of: Gag mutations associated with PI exposure, Gag mutations associated with PI resistance (rGag) and novel Gag mutations (nGag) identified in sequences from the PCS cohort which did not harbor PR RAMs.

PID	Gag mutations identified in literature		Novel Gag mutations (nGag)
	Associated with PI exposure	Associated with PI resistance (rGag)	
PCS 005	E12K, G62R, T375N, H219Q	R76K, Y79F	Q69K
PCS 007	E12K, V35I, T375N, G381S, I389T	None	None
PCS 011	E12K, T375N, G381S	L449P, R452K, P453L	S111I/ I256V
PCS 012	E12K, G381S	R76K	I256V
PCS 014	E12K, V370A, T375N	Y79F	I256V
PCS 016	E12K, G62R, S373Q, T375N, G381S	P453L	None
PCS 018	E12K, A374S, T375N, G381S	R76K, Y79F	Q69K, S111I, I256V
PCS 020	E12K, V370A, G381S	R76K, Y79F, K436R, L449P	I256V
PCS 021	E12K, T375N, H219Q	None	S111C
PCS 022	E12K, V35I, T375N	None	Q69K
PCS 023	S373P, I389T	P453L	None
PCS 024	E12K, T375N, G381S	Y79F	None

PID	Gag mutations identified in literature		Novel Gag mutations (nGag)
	Associated with PI exposure	Associated with PI resistance (rGag)	
PCS 030	E12K, S373T, T375N, G381S, H219Q	R76K	Q69K
PCS 036	E12K, V370A, T375N, H219Q	R76K	T239A
PCS 042	E12K, V35I, V370A	Y79F	None
PCS 047	E12K	None	S111I/ I256V
PCS 049	E12K, T375N, G381S	R76K, Y79F, A431V	S111C
PCS 052	E12K, S373T	R76K, Y79F	S111C
PCS 058	E12K, T375N, G381S, I389T	R76K	S111C
PCS 059	E12K, A374P, T375N, G381S	R76K	Q69K, S111C
PCS 063	E12K, V370A, S373T, T375N	R76K, R452K	Q69K, I256V
PCS 070	E12K, V35I, T375N	R76K, Y79F	I256V
PCS 075	T375N, G381S	V128I	Q69K, I256V
PCS 083	E12K, T375N, G381S	R76K, Y79F	S111C, I256V
PCS 084	E12K, T375N, I376V	Y79F	Q69K
PCS 093	E12K, G62R, A374S, T375N, I376V	R76K, P453L	S111C
PCS 095	E12K, G62R, V370A, S373T, T375N, H219Q	None	T239A, I256V
PCS 096	E12K, A374S, T375N	None	Q69K, S111C
PCS 104	E12K, G62R, V370A, T375N, G381S	R76K, R452K	Q69K
PCS 105	E12K, V370A, S373T, T375N, G381S	K436R	Q69K, S111C
PCS 112	E12K, V35I, S373Q, A374S, I389T	Y79F, V128I	S111C
PCS 114	E12K, V35I, V370A, S373T, T375N, I401V	None	I256V
PCS 115	E12K, G381S, I389T, H219Q	R76K, L449P, P453L	Q69K, T239A, I256V
PCS 116	E12K, V370A, T375N, G381S, H219Q	R76K	None
PCS 118	E12K, T375N, H219Q	R76K, Y79F, V128I, L449P, P453L	Q69K
PCS 120	E12K, V35I, V370A, T375N, I389T	R76K, K436R	Q69K, S111C
PCS 128	E12K, V35I, T375N	None	Q69K, I256V

PID	Gag mutations identified in literature		Novel Gag mutations (nGag)
	Associated with PI exposure	Associated with PI resistance (rGag)	
PCM004	E12K, V35I, S373Q, T375N, G381S	None	Q69K, I256V
PCM009	E12K, V370A, T375N, H219Q	None	S111C
PCSM012	E12K, T375N, G381S	None	Q69K, I256V
PCSM013	E12K, V35I, T375N, I376V, G381S	R76K, P453L	S111I
PCSM020	E12K, V35I, T375N, I401V	R76K, Y79F	None
PCM022	E12K, V370A, T375N, G381S, H219Q	None	S111C
PCM024	E12K, T375N, G381S	None	None
PCM029	E12K, G62R, V370A, T375N, I376V, I389T,	R76K, P453L	None
PCM040	E12K, V370A, T375N, G381S	R76K, L449P	Q69K, S111C, T239S, I256V

Overall, 42.5% (34/80) of participants harboured PR RAMs. Eighty-four percent (67 out of 80) of participants harboured rGag mutations and 90% (72 out of 80) of participants harboured nGag mutations. All participants (i.e. 34 out of 34) with PR RAMs harboured rGag mutations whilst 71% (33 out of 46) of participants without PR RAMs harboured rGag mutations. Similarly, all participants with PR RAMs harboured nGag mutations whilst 76% (38 out of 46) of participants without PR RAMs harboured nGag mutations. This data showed that for the PCS cohort PR RAMs always occurred in conjunction with gag mutations, however gag mutations also occurred without PR RAMs.

2.4 Discussion

The role of Gag in PI resistance is well documented for HIV-1 subtype B, however information is limited for HIV-1 subtype C, the most prevalent subtype globally (28). In this study, PI RAMs, in both Gag and PR, were identified in an HIV-1 subtype C cohort of participants failing a PI inclusive treatment regimen. Both novel and previously identified Gag mutations associated with PI resistance/exposure were identified.

Results showed that 34 out of 80 participants (42.5%) failing a PI inclusive treatment regimen presented with PR RAMs linked to high level resistance to various PI's. The most common pattern for PR RAMs was M46I+I54V+V82A (n=22) followed by M46I+I54V+L76V+V82A (n=16). These patterns were consistent with a recent study which investigated PR RAMs in HIV-1 subtype C sequences from South Africa and Israel (51, 52). Previous studies reported that the M46I+I54V+L76V+V82A combination of PR RAMs confers reduced susceptibility to eight PI's (ATV, FPV, IDV, LPV, NFV, RTV, SQV and TPV) (53-55). Unfortunately all of these studies were performed prior to the introduction of DRV (i.e. from 2003-2006) hence no data was available for DRV in these studies. However, various later studies showed that L76V confers reduced susceptibility to DRV, FPV, IDV and LPV (56-58). These findings indicate that participants with the M46I+I54V+L76V+V82A pattern of PR RAMs would display reduced susceptibility to all PI's. This has implications on future treatment for these patients in SA, since third-line regimens (i.e. the only remaining treatment regimen for the treatment of HIV after second-line regimen failure) comprises of an InSTI, a PI (generally DRV) and an NNRTI.

Analysis of the Gag gene identified six rGag mutations that occurred at significantly higher frequencies in the PCS cohort versus the acute and treatment naive cohorts (Figure 2.5), two of which appeared to be natural polymorphisms (R76K and Y79F) whilst the remaining four (A431V, L449F/P, R452K and P453L), were associated with PI resistance/exposure. Parry et al., (2011) identified R76K and Y79F to be associated with improved replication capacity in PI resistance when they occurred in conjunction with each other, however they only exerted a minor effect on replication capacity when found individually. This could suggest that HIV-1 subtype C viruses may have a replicative advantage under PI drug selection pressure, however further experimental work using a subtype C backbone is required to confirm this.

Of the four rGag mutations identified in this study to be associated with PI resistance (i.e. A431V, L449F/P, R452K and P453L), A431V was the only one to have been previously described to confer resistance to PI's in the absence of PR RAMs. It is considered as both a primary and compensatory mutation in PI resistance (12, 17). This mutation was the most commonly occurring rGag mutation in the PCS cohort, and in 96% (24 out 25) of instances was found in conjunction with PR RAMs which have been previously associated with reduced susceptibility to various PI's (53, 54). Given that A431V can confer reduced susceptibility to PI's alone, combining it with PR RAMs would likely increase drug

susceptibility scores thereby providing a more accurate measure of PI drug susceptibility. The inclusion of Gag in PI resistance algorithms should thus be considered.

Unlike A431V; L449F/P, R452K and P453L are not considered as primary resistance mutations, since they work to enhance PI resistance in the presence of PR RAMs (23, 24, 59). The L449F/P, R452K and P453L mutations, all located at the p1/p6 CS of Gag (4) are often seen in PI resistance (9, 19, 22, 24, 28, 38, 40, 44, 46, 47,48, 49,51, 52,54, 55) and have been shown to occur together with the following PR mutations: D30N, N88D, I50V and/or I84V (24, 59, 60). In contrast to previous studies which showed that the L449F Gag mutation occurred in conjunction with the following PR mutations: D30N, I50V, I84V and N88D (4, 59, 60), the current study showed it to occur predominantly in isolation of PR RAMs. Thereafter it was commonly found with V82A alone and a combination of A431V+M46I+I54V+L76V+V82A. The R452K mutation was found in combination with the following PR RAMs: M46I, I50V, I54V, V82A and L90M, with the L90M PR RAM being the only PR mutation previously reported to occur with the R452K Gag mutation (24). The P453L Gag mutation was found to occur together with the following PR RAMs in the current study: M46I, 50V, I54V and V82A. It has however been previously reported to occur with the I84V, N88D and L90M PR RAMs (24, 59, 60). Overall, we highlight variation in PR RAMs associated with rGag mutations between this study and previous studies, which could be attributed to subtype variation since most of the previous studies have been conducted on HIV-1 subtype B. This highlights that subtype B studies may not always translate to subtype C and thus supports the need for more research on HIV-1 subtype C.

Whilst the current study identified A431V, L449F/P and R452K in participants failing a PI inclusive treatment regimen, a recent study by Li et al., (2014) in which 137 HIV-1 subtype C PI resistant Gag-Protease sequences were compared to 1,786 HIV-1 subtype C PI susceptible Gag-Protease sequences, did not show the occurrence of these mutations in subtype C sequences (31). The difference in data between the studies could potentially represent treatment specific variations where LPV could possibly be associated with the development of A431V, L449F and R452K in subtype C. This however could not be confirmed since the Li study did not have matched treatment data for the sequences used in their study.

Several Gag mutations associated with PI exposure were also identified in this study. Four of these exposure associated mutations (E12K, V370A/M, T375N and G381S) were found at frequencies >40% in the PCS cohort and control groups (i.e. treatment naïve groups) (Figure 2.6) suggesting that they are natural polymorphisms in HIV-1 subtype C. The E12K mutation, located in matrix, was shown previously to occur in response to ATV treatment in HIV-1 subtype B. It has been suggested that it may have a role in improving viral fitness in the presence of PI's (21). The V370A mutation, located in the p2 region of Gag, has been shown to accumulate during PI treatment, with its role in PI resistance remaining unclear (22, 61). Codon 375 and 381 are located in the p2/NC Gag CS which is known to control the rate and order of polyprotein cleavage (62). Although the role of these mutations in PI resistance is unclear, their presence as natural polymorphisms highlights that Gag in HIV-1 subtype C differs from Gag in HIV-1 subtype B and further supports the need for subtype specific studies.

The V35I and I401T exposure associated mutations occurred at significantly higher frequencies in the PCS cohort in comparison to the control groups indicating that both of these mutations may have a role in LPV resistance/exposure in HIV-1 subtype C. Gatanaga et al., (2012) demonstrated the development of V35I in response to ATV and suggested that it may have a compensatory role in viral replication in the presence of PI's (21). Stray et al., (2013) showed the development of I401T several months after exposure to a test PI (GS-8374). Although clear that both these non-CS mutations occur under PI selection pressure, their exact roles in PI resistance are unknown, with further research required.

The current study represents the first to identify four novel non-CS mutations in Gag (i.e. Q69K, S111I/C, T239A/S and I256V) associated with PI resistance/exposure (Figure 2.7). Interestingly, these four Gag mutations were found at similar frequencies in participants with and without PR RAMs indicating that their development is not dependent upon PR RAMs but is likely dependent on PI exposure. Their function in PI resistance is unclear and will be explored further in Chapter 3 and 4.

Interestingly, this study showed that PR RAMs always occurred in conjunction with rGag, nGag and PI exposure associated Gag mutations, however these Gag mutations (i.e. rGag, nGag and PI exposure associated Gag mutations) were found to also occur without PR RAMs. This suggests that Gag mutations may develop before PR RAMs, and could

potentially facilitate the development of PR RAMs. Such a scenario has been suggested for rGag mutations in two previous studies. In the first, Bally et al., (2000) showed that the L449F Gag resistance mutation favours the development of the I84V PR RAM. In the second study, Gatanaga et al., (2002) reported that the H219Q Gag mutation preceded the development of PR RAMs and suggested that it may facilitate the development of PR RAMs (2002). There is however much debate on whether Gag or PR mutations appear first with further investigations required to elucidate the mechanisms of Gag-Protease co-evolution in PI resistance.

Furthermore, 89% of all participants with PR RAMs harboured a rGag mutation in either the NC/p1 or p1/p6 cleavage site. A study by Kolli et al., (2009) displayed a similar result, were all viruses with PR mutations within the active site, also harboured at least one Gag mutation in either the NC/p1 or p1/p6 Gag cleavage site (42). Mutations in the NC/p1 and p1/p6 CS have been shown to alter the structure of the CS's which enables sustained cleavage by the mutant PR in the presence of PI's (42). The high prevalence of rGag CS mutations occurring in conjunction with PR RAMs could thus suggest that Gag and Protease co-evolve during PI resistance.

Overall, 58% of participants failing a PI inclusive treatment regimen did not harbour mutations in PR. All these participants however harboured either rGag and/or PI exposure associated Gag mutations, which have been previously reported to enhance PI resistance or viral replication in the presence of PR RAMs. This could suggest that Gag mutations in HIV-1 subtype C contribute to PI resistance in the absence of PR RAMs. Further work is however required to determine the role of these Gag mutations on viral replication capacity and PI drug susceptibility.

2.5 Conclusion

Variability in Gag and PR between HIV-1 subtype B and HIV-1 subtype C, highlights that research on subtype B cannot always be translated to subtype C. Protease RAMs always occurred in conjunction with Gag mutations, however Gag mutations were found to occur without PR RAMs, suggesting that Gag mutations may occur first and could have a role in facilitating the development of PR RAMs. Not all participants failing a PI inclusive treatment regimen harboured PR RAMs, a large proportion of these participants did however harbour rGag, nGag and PI exposure associated Gag mutations, suggesting a

possible role for Gag in PI resistance. Lastly four novel Gag mutations associated with PI resistance were identified, their role in PI resistance requires further investigation which will be addressed in Chapter 3.

2.6 References

1. Lefebvre E, Schiffer CA. 2008. Resilience to resistance of HIV-1 Protease inhibitors: profile of darunavir. AIDS Rev 10:131-142.
2. Volberding PA, Deeks SG. 2010. Antiretroviral therapy and management of HIV infection. Lancet 376:49-62.
3. Thompson MA, Aberg JA, Cahn P, Montaner JS, Rizzardini G, Telenti A, Gatell JM, Gunthard HF, Hammer SM, Hirsch MS, Jacobsen DM, Reiss P, Richman DD, Volberding PA, Yeni P, Schooley RT, International AS-USA. 2010. Antiretroviral treatment of adult HIV infection: 2010 recommendations of the International AIDS Society-USA panel. JAMA 304:321-333.
4. Fun A, Wensing AM, Verheyen J, Nijhuis M. 2012. Human Immunodeficiency Virus Gag and Protease: partners in resistance. Retrovirology 9:63.
5. Clavel F, Mammano F. 2010. Role of Gag in HIV Resistance to Protease Inhibitors. Viruses 2:1411-1426.
6. Sutherland KA, Parry CM, McCormick A, Kapaata A, Lyagoba F, Kaleebu P, Gilks CF, Goodall R, Spyer M, Kityo C, Pillay D, Gupta RK, Group DV. 2015. Evidence for Reduced Drug Susceptibility without Emergence of Major Protease Mutations following Protease Inhibitor Monotherapy Failure in the SARA Trial. PLoS One 10:e0137834.
7. Larrouy L, Chazallon C, Landman R, Capitant C, Peytavin G, Collin G, Charpentier C, Storto A, Pialoux G, Katlama C, Girard PM, Yeni P, Aboulker JP, Brun-Vezinet F, Descamps D, Group AS. 2010. Gag mutations can impact virological response to dual-boosted Protease inhibitor combinations in antiretroviral-naive HIV-infected patients. Antimicrob Agents Chemother 54:2910-2919.
8. Gupta RK, Kohli A, McCormick AL, Towers GJ, Pillay D, Parry CM. 2010. Full-length HIV-1 Gag determines Protease inhibitor susceptibility within in vitro assays. AIDS 24:1651-1655.
9. Kolli M, Lastere S, Schiffer CA. 2006. Co-evolution of nelfinavir-resistant HIV-1 Protease and the p1-p6 substrate. Virology 347:405-409.
10. Prabu-Jeyabalan M, Nalivaika EA, Romano K, Schiffer CA. 2006. Mechanism of substrate recognition by drug-resistant human immunodeficiency virus type 1 Protease variants revealed by a novel structural intermediate. J Virol 80:3607-3616.

11. Parry CM, Kohli A, Boinett CJ, Towers GJ, McCormick AL, Pillay D. 2009. Gag determinants of fitness and drug susceptibility in Protease inhibitor-resistant human immunodeficiency virus type 1. J Virol 83:9094-9101.
12. Nijhuis M, van Maarseveen NM, Lastere S, Schipper P, Coakley E, Glass B, Rovenska M, de Jong D, Chappey C, Goedegebuure IW, Heilek-Snyder G, Dulude D, Cammack N, Brakier-Gingras L, Konvalinka J, Parkin N, Krausslich HG, Brun-Vezinet F, Boucher CA. 2007. A novel substrate-based HIV-1 Protease inhibitor drug resistance mechanism. PLOS Medicine 4:e36.
13. Iyidogan P, Anderson KS. 2014. Current perspectives on HIV-1 antiretroviral drug resistance. Viruses 6:4095-4139.
14. Mahalingam B, Louis JM, Reed CC, Adomat JM, Krouse J, Wang YF, Harrison RW, Weber IT. 1999. Structural and kinetic analysis of drug resistant mutants of HIV-1 Protease. Eur J Biochem 263:238-245.
15. Croteau G, Doyon L, Thibeault D, McKercher G, Pilote L, Lamarre D. 1997. Impaired fitness of human immunodeficiency virus type 1 variants with high-level resistance to Protease inhibitors. J Virol 71:1089-1096.
16. Gulnik SV, Suvorov LI, Liu B, Yu B, Anderson B, Mitsuya H, Erickson JW. 1995. Kinetic characterization and cross-resistance patterns of HIV-1 Protease mutants selected under drug pressure. Biochemistry 34:9282-9287.
17. Dam E, Quercia R, Glass B, Descamps D, Launay O, Duval X, Krausslich HG, Hance AJ, Clavel F, Group AS. 2009. Gag mutations strongly contribute to HIV-1 resistance to Protease inhibitors in highly drug-experienced patients besides compensating for fitness loss. PLoS Pathog 5:e1000345.
18. Borman AM, Paulous S, Clavel F. 1996. Resistance of human immunodeficiency virus type 1 to Protease inhibitors: selection of resistance mutations in the presence and absence of the drug. J Gen Virol 77 (Pt 3):419-426.
19. Prado JG, Wrin T, Beauchaine J, Ruiz L, Petropoulos CJ, Frost SD, Clotet B, D'Aquila RT, Martinez-Picado J. 2002. Amprenavir-resistant HIV-1 exhibits lopinavir cross-resistance and reduced replication capacity. AIDS 16:1009-1017.
20. Doyon L, Payant C, Brakier-Gingras L, Lamarre D. 1998. Novel Gag-Pol frameshift site in human immunodeficiency virus type 1 variants resistant to Protease inhibitors. J Virol 72:6146-6150.
21. Gatanaga H, Suzuki Y, Tsang H, Yoshimura K, Kavlick MF, Nagashima K, Gorelick RJ, Mardy S, Tang C, Summers MF, Mitsuya H. 2002. Amino acid substitutions in Gag protein at non-cleavage sites are indispensable for the development of a high

multitude of HIV-1 resistance against Protease inhibitors. J Biol Chem 277:5952-5961.

22. Myint L, Matsuda M, Matsuda Z, Yokomaku Y, Chiba T, Okano A, Yamada K, Sugiura W. 2004. Gag non-cleavage site mutations contribute to full recovery of viral fitness in Protease inhibitor-resistant human immunodeficiency virus type 1. Antimicrob Agents Chemother 48:444-452.
23. Yates PJ, Hazen R, St Clair M, Boone L, Tisdale M, Elston RC. 2006. In vitro development of resistance to human immunodeficiency virus Protease inhibitor GW640385. Antimicrob Agents Chemother 50:1092-1095.
24. Kolli M, Stawiski E, Chappey C, Schiffer CA. 2009. Human immunodeficiency virus type 1 Protease-correlated cleavage site mutations enhance inhibitor resistance. J Virol 83:11027-11042.
25. Parry CM, Kolli M, Myers RE, Cane PA, Schiffer C, Pillay D. 2011. Three residues in HIV-1 matrix contribute to Protease inhibitor susceptibility and replication capacity. Antimicrob Agents Chemother 55:1106-1113.
26. van Maarseveen NM, Andersson D, Lepsik M, Fun A, Schipper PJ, de Jong D, Boucher CA, Nijhuis M. 2012. Modulation of HIV-1 Gag NC/p1 cleavage efficiency affects Protease inhibitor resistance and viral replicative capacity. Retrovirology 9:29.
27. Cote HC, Brumme ZL, Harrigan PR. 2001. Human immunodeficiency virus type 1 Protease cleavage site mutations associated with Protease inhibitor cross-resistance selected by indinavir, ritonavir, and/or saquinavir. J Virol 75:589-594.
28. Alcalde R, Guimaraes ML, Duarte AJ, Casseb J. 2012. Clinical, epidemiological and molecular features of the HIV-1 subtype C and recombinant forms that are circulating in the city of Sao Paulo, Brazil. Virol J 9:156.
29. Flynn WF, Chang MW, Tan Z, Oliveira G, Yuan J, Okulicz JF, Torbett BE, Levy RM. 2015. Deep sequencing of Protease inhibitor resistant HIV patient isolates reveals patterns of correlated mutations in Gag and Protease. PLoS Comput Biol 11:e1004249.
30. Sutherland KA, Mbisa JL, Cane PA, Pillay D, Parry CM. 2014. Contribution of Gag and Protease to variation in susceptibility to Protease inhibitors between different strains of subtype B human immunodeficiency virus type 1. J Gen Virol 95:190-200.
31. Li G, Verheyen J, Theys K, Piampongsant S, Van Laethem K, Vandamme AM. 2014. HIV-1 Gag C-terminal amino acid substitutions emerging under selective

pressure of Protease inhibitors in patient populations infected with different HIV-1 subtypes. Retrovirology 11:79.

32. Giandhari J, Basson AE, Coovadia A, Kuhn L, Abrams EJ, Strehlau R, Morris L, Hunt GM. 2015. Genetic Changes in HIV-1 Gag-Protease Associated with Protease Inhibitor-Based Therapy Failure in Pediatric Patients. AIDS Res Hum Retroviruses 31:776-782.
33. Evans D. 2013. Ten years on ART where to now. S Afr Med J 103:229-230.
34. Moholisa RR, Schomaker M, Kuhn L, Meredith S, Wiesner L, Coovadia A, Strehlau R, Martens L, Abrams EJ, Maartens G, McIlleron H. 2014. Plasma lopinavir concentrations predict virological failure in a cohort of South African children initiating a Protease-inhibitor-based regimen. Antivir Ther 19:399-406.
35. Wright JK, Novitsky V, Brockman MA, Brumme ZL, Brumme CJ, Carlson JM, Heckerman D, Wang B, Losina E, Leshwedi M, van der Stok M, Maphumulo L, Mkhwanazi N, Chonco F, Goulder PJ, Essex M, Walker BD, Ndung'u T. 2011. Influence of Gag-Protease-mediated replication capacity on disease progression in individuals recently infected with HIV-1 subtype C. J Virol 85:3996-4006.
36. Los Alamos National Laboratory. HIV sequence database. http://www.hiv.lanl.gov/content/sequence/HIV/MAP/landmark.html. Accessed 26 September.
37. Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673-4680.
38. Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 95-98.
39. Los Alamos National Laboratory. 2015. ElimDupes. http://www.hiv.lanl.gov/content/sequence/ELIMDUPES/elimdupes.html. Accessed April 2015.
40. de Oliveira T, Deforche K, Cassol S, Salminen M, Paraskevis D, Seebregts C, Snoeck J, van Rensburg EJ, Wensing AM, van de Vijver DA, Boucher CA, Camacho R, Vandamme AM. 2005. An automated genotyping system for analysis of HIV-1 and other microbial sequences. Bioinformatics 21:3797-3800.
41. Alcantara LC, Cassol S, Libin P, Deforche K, Pybus OG, Van Ranst M, Galvao-Castro B, Vandamme AM, de Oliveira T. 2009. A standardized framework for

accurate, high-throughput genotyping of recombinant and non-recombinant viral sequences. Nucleic Acids Res 37:W634-642.

42. Pillay K. 2015. The impact of the p7/p1 cleavage site mutations on replication capacity and drug resistance in HIV-1 subtype C. Masters of Science. University o Kwa-Zulu Natal, South Africa.
43. Wright JK, Brumme ZL, Carlson JM, Heckerman D, Kadie CM, Brumme CJ, Wang B, Losina E, Miura T, Chonco F, van der Stok M, Mncube Z, Bishop K, Goulder PJ, Walker BD, Brockman MA, Ndung'u T. 2010. Gag-Protease-mediated replication capacity in HIV-1 subtype C chronic infection: associations with HLA type and clinical parameters. J Virol 84:10820-10831.
44. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307-321.
45. Liu TF, Shafer RW. 2006. Web resources for HIV type 1 genotypic-resistance test interpretation. Clin Infect Dis 42:1608-1618.
46. Wensing AM, Calvez V, Gunthard HF, Johnson VA, Paredes R, Pillay D, Shafer RW, Richman DD. 2014. 2014 Update of the drug resistance mutations in HIV-1. Top Antivir Med 22:642-650.
47. Korber B, Myers G. 1992. Signature pattern analysis: a method for assessing viral sequence relatedness. AIDS Res Hum Retroviruses 8:1549-1560.
48. Storey JD. 2002. A direct approach to false discovery rates. . Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64:479-498.
49. R Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.
50. Kaufmann GR, Suzuki K, Cunningham P, Mukaide M, Kondo M, Imai M, Zaunders J, Cooper DA. 2001. Impact of HIV type 1 Protease, reverse transcriptase, cleavage site, and p6 mutations on the virological response to quadruple therapy with saquinavir, ritonavir, and two nucleoside analogs. AIDS Res Hum Retroviruses 17:487-497.
51. Van Zyl GU, Liu TF, Claassen M, Engelbrecht S, de Oliveira T, Preiser W, Wood NT, Travers S, Shafer RW. 2013. Trends in Genotypic HIV-1 Antiretroviral Resistance between 2006 and 2012 in South African Patients Receiving First- and Second-Line Antiretroviral Treatment Regimens. PLoS One 8:e67188.
52. Grossman Z, Schapiro JM, Levy I, Elbirt D, Chowers M, Riesenberg K, Olstein-Pops K, Shahar E, Istomin V, Asher I, Gottessman BS, Shemer Y, Elinav H,

Hassoun G, Rosenberg S, Averbuch D, Machleb-Guri K, Kra-Oz Z, Radian-Sade S, Rudich H, Ram D, Maayan S, Agmon-Levin N, Sthoeger Z. 2014. Comparable long-term efficacy of Lopinavir/Ritonavir and similar drug-resistance profiles in different HIV-1 subtypes. PLoS One 9:e86239.

53. Mo H, King MS, King K, Molla A, Brun S, Kempf DJ. 2005. Selection of resistance in Protease inhibitor-experienced, human immunodeficiency virus type 1-infected subjects failing lopinavir- and ritonavir-based therapy: mutation patterns and baseline correlates. J Virol 79:3329-3338.
54. Baxter JD, Schapiro JM, Boucher CA, Kohlbrenner VM, Hall DB, Scherer JR, Mayers DL. 2006. Genotypic changes in human immunodeficiency virus type 1 Protease associated with reduced susceptibility and virologic response to the Protease inhibitor tipranavir. J Virol 80:10794-10801.
55. Gonzales MJ, Belitskaya I, Dupnik KM, Rhee SY, Shafer RW. 2003. Protease and reverse transcriptase mutation patterns in HIV type 1 isolates from heavily treated persons: comparison of isolates from Northern California with isolates from other regions. AIDS Res Hum Retroviruses 19:909-915.
56. Rhee SY, Taylor J, Fessel WJ, Kaufman D, Towner W, Troia P, Ruane P, Hellinger J, Shirvani V, Zolopa A, Shafer RW. 2010. HIV-1 Protease mutations and Protease inhibitor cross-resistance. Antimicrob Agents Chemother 54:4253-4261.
57. Vermeiren H, Van Craenenbroeck E, Alen P, Bacheler L, Picchio G, Lecocq P, Virco Clinical Response Collaborative T. 2007. Prediction of HIV-1 drug susceptibility phenotype from the viral genotype using linear regression modeling. J Virol Methods 145:47-55.
58. De Meyer S, Hill A, Picchio G, DeMasi R, De Paepe E, de Bethune MP. 2008. Influence of baseline Protease inhibitor resistance on the efficacy of darunavir/ritonavir or lopinavir/ritonavir in the TITAN trial. J Acquir Immune Defic Syndr 49:563-564.
59. Maguire MF, Guinea R, Griffin P, Macmanus S, Elston RC, Wolfram J, Richards N, Hanlon MH, Porter DJ, Wrin T, Parkin N, Tisdale M, Furfine E, Petropoulos C, Snowden BW, Kleim JP. 2002. Changes in human immunodeficiency virus type 1 Gag at positions L449 and P453 are linked to I50V Protease mutants in vivo and cause reduction of sensitivity to amprenavir and improved viral fitness in vitro. J Virol 76:7398-7406.

60. Bally F, Martinez R, Peters S, Sudre P, Telenti A. 2000. Polymorphism of HIV type 1 Gag p7/p1 and p1/p6 cleavage sites: clinical significance and implications for resistance to Protease inhibitors. AIDS Res Hum Retroviruses 16:1209-1213.
61. Verheyen J, Verhofstede C, Knops E, Vandekerckhove L, Fun A, Brunen D, Dauwe K, Wensing AM, Pfister H, Kaiser R, Nijhuis M. 2010. High prevalence of bevirimat resistance mutations in Protease inhibitor-resistant HIV isolates. AIDS 24:669-673.
62. de Oliveira T, Engelbrecht S, Janse van Rensburg E, Gordon M, Bishop K, zur Megede J, Barnett SW, Cassol S. 2003. Variability at Human Immunodeficiency Virus Type 1 Subtype C Protease Cleavage Sites: an Indication of Viral Fitness? Journal of Virology 77:9422-9430.
63. Rabi SA, Laird GM, Durand CM, Laskey S, Shan L, Bailey JR, Chioma S, Moore RD, Siliciano RF. 2013. Multi-step inhibition explains HIV-1 Protease inhibitor pharmacodynamics and resistance. J Clin Invest 123:3848-3860.

CHAPTER 3

Effect of Gag-Protease mutations on replication capacity and drug susceptibility

3 CHAPTER 3: EFFECT OF GAG AND PR MUTATIONS ON REPLICATION CAPACITY AND DRUG SUSCEPTIBILITY

3.1 Introduction

The development and implementation of cART has enabled HIV-1 infected individuals to control viral replication thereby improving morbidity and mortality and reducing the risk of HIV-1 transmission. The development of drug resistance mutations however threatens the success of treatment.

Drug resistance mutations occur as a result of: low fidelity viral DNA polymerase, high error rate of viral RT, treatment interruptions (caused mainly by poor adherence and poor access to treatment), suboptimal drug concentrations and viral recombination (discussed in Chapter 1, section 1.9) (1). Studies have shown that the development of drug resistance mutations in functional enzymes (i.e. PR, RT and INT) and their substrates may be associated with altered viral replication and drug susceptibility (1-5).

This has been particularly well demonstrated for PR and Gag (1-4, 6-8), were several studies have shown that PR mutations are associated with reduced replication capacity and drug susceptibility to PI's, however replication capacity can be rescued by mutations in Gag, with varying effects on drug susceptibility (1-4, 9-13).

The majority of these studies have focused on the carboxy terminal of Gag (9, 12-18) with limited information available on the amino terminal of Gag (2, 4). A recent study, using site directed mutagenesis to insert mutations of interest into a subtype B backbone, showed that under PI drug selection pressure, three residues in the Gag amino terminal (i.e. R76K, Y79F and T81A) were associated with increased replication capacity when found in conjunction with PR RAMs. Additionally, these three residues caused a 5-7 fold reduction in drug susceptibility when found in the absence of PR RAMs (2). An additional study, employing six patient derived Gag-Protease viruses cloned into a subtype B backbone, showed that one participant with six Gag mutations, spanning the amino and carboxy terminal of Gag (V7I, G49D, R69Q, A120D, Q127K, N375S, I462S), displayed a reduced susceptibility to PI's in the absence of PR RAMs (4). These data indicates that the amino terminal of Gag plays a role in PI resistance, as with the carboxy terminal and thus

encourages the use of full-length Gag in studies pertaining to PI resistance. To date there is only one study that has investigated the role of the amino and carboxy terminal of Gag in PI-resistance in HIV-1 subtype C (50). This study used a paediatric cohort of 20 patients to identify Gag-Protease mutations and measure their impact on PI drug susceptibility. it however did not investigate the impact of these mutations on repllication capacity.

As part of the current study we sought to identify Gag codon substitutions significantly associated with altered replication capacity in HIV-1 subtype C infected patients failing a PI inclusive treatment regimen. Additionally, we investigated the impact of these codons and four novel Gag mutations (nGag; identified in Chapter 2) on replication capacity and drug susceptibility.

This was achieved by constructing 80 recombinant viruses using patient-derived full length Gag-Protease amplicons. Thereafter the replication capacity of all 80 viruses was measured. A codon by codon analysis tool was used to identify mutations in Gag associated with significantly altered replication capacity. Additionally, the replication capacity of viruses with nGag mutations, occurring with or without PR RAMs, was compared to investigate the role of nGag mutations in PI resistance. Based on these results, 18 selected viruses, containing Gag mutations of interest, were subject to phenotypic drug susceptibility analysis in order to determine the impact of identified mutations on drug susceptibility. An overview of the approach of this study is provided in Figure 3.1 below.

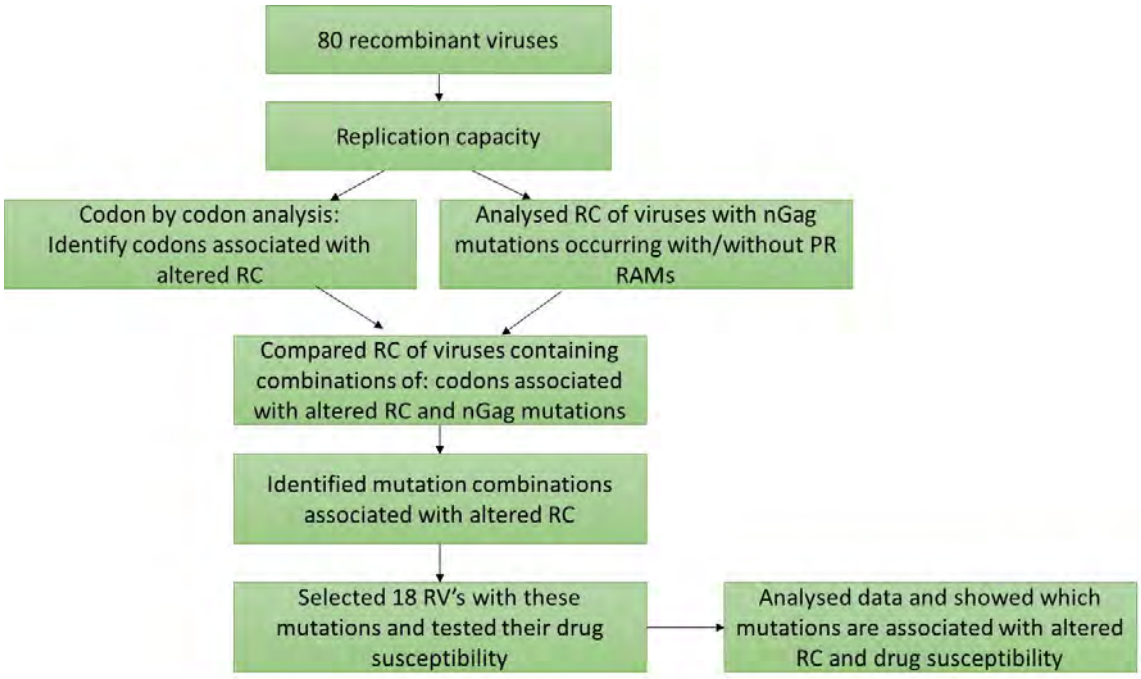


Figure 3-1 Summary of study approach. Abbreviations: RC - replication capacity; nGag - novel Gag; RAMs - resistance associated mutations; RV's - recombinant viruses.

3.2 Methods

3.2.1 Generation of Gag-Protease recombinant virus stocks

Recombinant viruses for 80 participant derived samples were generated by co-transfection of a CEM derived T-cell line (i.e. GXR cells) with an NL43-deleted-*gag-protease* (NL43Δ*gag-protease*) backbone and patient derived Gag-Protease amplicons (generated in section 2.2.3) using a method previously described and validated (19-22).

3.2.1.1 Generation of pNL43ΔGag-*pro* backbone

A pNL43Δ*gag-protease* plasmid, supplied in *E. coli* STBL3 cells by Dr Jaclyn Mann (University of KwaZulu-Natal, HIV Pathogenesis Programme, Durban, South Africa), was used as a backbone for generation of recombinant viruses. The plasmid contained a BstE II restriction enzyme recognition site in the Gag-Protease coding region, enabling linearization of the plasmid as required for recombination with patient-derived Gag-Protease amplicons (19-21).

Large quantities of pNL43Δ*gag-pro* plasmid stocks were generated by adding 30 µl of the pNL43Δ*gag-pro* plasmid (in STBL3 cells) to a mixture of Luria-Bertani (LB) (Sigma Aldrich, St Louis, USA) broth and ampicillin (100 ml of LB broth and 100 µl ampicillin). This was incubated in a shaking incubator (Infors HT, Bottmingen, Switzerland) at 37ºC and 230 rpm for 16 hours. Thereafter, the plasmid was extracted and purified using the Qiagen Plasmid Maxi kit (Qiagen) as per manufacturer's instructions, and quantified using a nanodrop spectrophotometer (Thermo Scientific, Delaware, USA). Quantified plasmid was aliquoted and stored at -80ºC.

3.2.1.2 Preparation of CEM-GXR25 cells

A CEM-GXR25 green fluorescent protein (GFP) reporter T-cell line (i.e. GXR cells) (23) was supplied by Dr Mark Brockman (Simon Fraser University, Vancouver, Canada). GXR cells are replication competent as they express the CD4 receptor and both the CXCR4 and CCR5 co-receptors. Additionally, they encode a Tat-inducible HIV-1 LTR- GFP expression cassette, which is responsible for the expression of GFP during infection thereby facilitating detection of infected cells using flow cytometry (23).

A frozen aliquot (i.e. 1 ml) of approximately 1 million GXR cells (stored in dimethylsulfoxide [DMSO], Sigma) was transferred from a liquid nitrogen freezer (Custom Biogenics Systems, Romeo, USA) directly into a preheated 37ºC water bath. The tube of cells was gently agitated in the water bath until the contents were completely thawed. Thereafter the tube of cells was transferred into a T25 flask (Corning-Costar, New York, USA) containing 4 ml of pre-warmed R10 culture medium, and incubated at 37ºC and 5% CO_2 for 24 hours in a humidified Heraeus incubator (Thermo Scientific). R10 media comprised of RPMI-1640 (Sigma), supplemented with 50 U/ml penicillin streptomycin (Gibco, New York, USA), 10 mM N-2-hydroxyethylpiperazine-N'-2ethanesulfonic acid (HEPES; Gibco), 2 mM L-glutamine (Sigma) and 10% foetal bovine serum (FBS; Gibco).

After 24 hours, the contents of the T25 flask was transferred into a 15 ml falcon tube and centrifuged at 1,500 rpm for 10 minutes (Heraeus multifuge 3SR+, Thermo Scientific) in order to remove DMSO. Cells were then re-suspended in 1 ml of pre-warmed R10 and transferred into a T25 flask containing 9 ml of pre-warmed R10. The flask was then incubated at 37ºC and 5% CO_2 for a further 24 hours.

Following incubation, cells were counted by adding 10 µl of thoroughly mixed cell culture to 10 µl of trypan blue (Bio-Rad, Hercules, USA). A total of 10 µl of this mixture was inserted into a TC20 cell counting slide (Bio-Rad) which was subsequently loaded into a TC20 automated cell counter (Bio-Rad). The output of the TC20 cell counter was the cell concentration (i.e. cells/ml). The volume of cell culture used to obtain a required number of cells was calculated as follows:

Volume of cell culture (ml) = number of cells required ÷ cell concentration (cells/ml)

Cells were maintained at a concentration of 250 000 cells/ml in a final volume of 30 ml in a T75 flask (Corning). Cell growth was monitored every second day and if not used for experiments, 80% of the cell culture was removed and replaced with fresh pre-warmed R10. Cells were maintained for a maximum of two months, after which time a new aliquot of GXR cells were thawed and prepared for use.

Aliquots of GXR cells were also stored in liquid nitrogen for use in future experiments, within two weeks of thawing. Briefly, 5 million GXR cells re-suspended in 900 µl of R10 was added to a cryovial (Greiner Bio-One, Germany). Immediately thereafter, 100 µl of DMSO was added drop by drop with simultaneous gentle agitation. The DMSO functioned to prevent the formation of water crystals during cryopreservation (24). The cryovials were then stored in a strata-cooler (Agilent Technologies, Waldbronn, Germany) at -80ºC overnight. The Strata cooler functioned to gradually reduce the temperature of the cells to -80ºC (at a controlled rate of 0.4-0.6ºC per minute) prior to transfer into a liquid nitrogen freezer (25).

3.2.1.3 Co-transfection

Recombinant viruses were prepared as previously described (26). Briefly, 2 hours prior to co-transfection, 10 µg of pNL43Δ*gag-pro* plasmid (prepared in 3.2.1.1) was digested with 2 µl of a 10 U/µl stock of BstE II enzyme (Promega, Madison, USA) for 2 hours in a water bath set to 60ºC. This functioned to linearize the plasmid. During this time, GXR cells were prepared (described in section 3.2.1.2). A total of 5 million cells were required for each co-transfection. One million of which was re-suspended in 1 ml of R10 and 4 µl of polybrene (10 µg/µl)[18] (Sigma) and subsequently added to an additional 8 ml of R10 in a T25 flask.

[18] Polybrene functions to make the cell wall porous, in order to facilitate virus entry into the cell.

These flasks were incubated at 37ºC and 5% CO_2. The remaining 4 million cells were re-suspended in 300 µl of R10 and added to a 4 mm electroporation cuvette (Bio-Rad).

A total of 80-90 µl of patient-derived Gag-Protease amplicons (i.e. 5-20 ng/µl) (generated in section 2.2.3), with either end of Gag and Protease exactly complementary to that of NL43, was thawed and added to the respective cuvette.

Thereafter, approximately 15–20 µl of digested pNL43Δ*gag-pro* plasmid (i.e. 10 µg of plasmid) was added to the cuvette. The contents of the cuvette was mixed using a 200 µl pipette. This was followed by electroporation in a Gene Pulser II electroporator (Bio-Rad), with conditions set to 250 V and 950 µF. Electroporation served to temporarily disrupt the cell wall enabling virus and plasmid to enter cells (27). Cuvettes were left at room temperature for 5 minutes, to allow for cell recovery. Thereafter the contents of cuvettes were transferred into the T25 flask containing the 1 million GXR cells in 9 ml R10 and 4 µl polybrene.

Flasks were incubated at 37ºC and 5% CO_2 for five days. On the fifth day an additional 5 ml of R10 was added to each flask followed by further incubation for five more days. On the tenth day, and every two days thereafter, 2 ml of culture was removed and replaced with 2 ml fresh pre-warmed R10. The 2 ml of culture removed was used to monitor viral growth.

No more than 10 recombinant viruses were generated at a time. Additionally, each experiment included a negative control (GXR cells only) and a positive control (NL43-WT virus). The NL43-WT virus was generated as part of this study by transfection of GXR cells with 10 µg of NL43-WT plasmid.

3.2.1.4 Monitoring viral growth by flow cytometry

Viral growth was monitored from the tenth day since the linearized plasmid and Gag-Protease amplicons required time to recombine and replicate (26).

From the 2 ml of culture removed from each flask (see section 3.2.1.3), a total of 1 ml was added to matrix cluster tubes (Corning Costar) and centrifuged at 1,500 rpm for 10 minutes, in order to pellet the cells. The majority of supernatant was discarded, leaving

behind very little residual supernatant to re-suspended the pelleted cells. Thereafter, 200 µl of 2% paraformaldehyde (PFA) was added to each tube and vortexed to mix thoroughly. Paraformaldehyde was prepared by adding 10 g of paraformaldehyde powder (Sigma) to 400 ml of Dulbecco's phosphate buffered saline (PBS; Gibco) and incubating the mixture at 55ºC for 45 minutes in a water bath. Thereafter the pH was adjusted to 7–7.4 and 100 ml more of Dulbecco's PBS was added. The 2% PFA solution was aliquoted and stored at -20ºC.

Following the 10 minute incubation of cells with PFA, matrix cluster tubes were inserted into FACs tubes (BD Biosciences, San Jose, USA), vortexed and flowed using the BD FACS Calibur (BD Biosciences). A negative control (i.e. GXR cells only) was always flowed first in order to gate live cells and distinguish between infected and uninfected cells (Figure 3.2). A total of 15,000 cells were gated. All data from the flow cytometer was further analysed using FlowJo (28). Figure 3.2 shows the gating of live cells; expression of >0.05% GFP indicated cell infection.

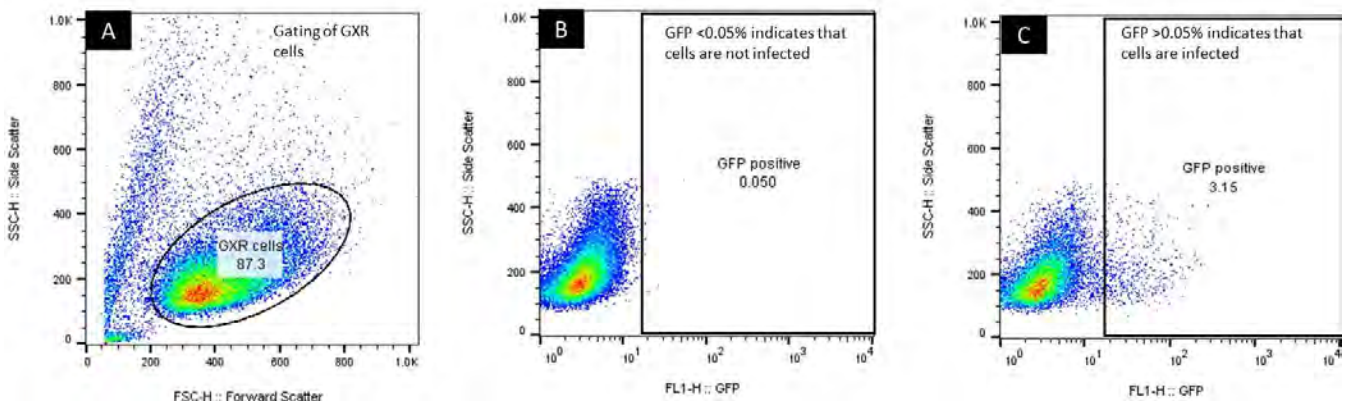


Figure 3-2 Example of the gating strategy used to distinguish between live and dead GXR cells and infected versus uninfected GXR cells. (a) Gating of a negative control identified that 87.3% of cells were alive. (b) The threshold to distinguish between infected and uninfected cells was set at 0.05% (c) Flow cytometry of a sample showed that a total of 3.15% of cells in one of the test samples was infected with HIV.

3.2.1.5 Harvesting recombinant viruses

Viruses were harvested once 25–30% of GXR cells were infected, as previously described (26). For harvesting, the contents of the T25 flask, containing the viral culture, was emptied into a falcon tube and centrifuged at 1,700 rpm and 4ºC for 5 minutes. The

supernatant was then aliquoted into cryovials (1 ml aliquots) and stored at -80ºC for use in titre, replication capacity and phenotyping assays.

3.2.1.6 Validation of recombinant viruses

A total of eight randomly selected recombinant viruses were subject to RNA extraction, amplification and sequencing of the Gag-Protease region (Chapter 2, section 2.2.2, 2.2.3 and 2.2.4). Virus sequences were analysed in conjunction with matched HIV-1 RNA plasma sequences in order to determine if sequences of recombinant viruses were representative of the matched plasma sequences. This was achieved by using the highlighter tool in Los Alamos to quantify similarity between pairs of sequences (i.e. recombinant virus sequence and plasma sequence) (29). Sequences were also assessed for contamination using a neighbour joining tree drawn in Paup version 4.0 which was edited in Figtree version 1.4.2 (30).

3.2.1.7 Validating consistency of the replication capacity assay

The consistency of the replication capacity assay was also validated. Stored RT-PCR products from 10 samples which were analysed for replication capacity in a previous study conducted in 2008, at the HIV Pathogenesis Programme laboratory, (26) were used to generate recombinant viruses for the current study. These recombinant viruses were subject to replication capacity assays, conducted in duplicate. Replication capacity data from 2008 was then correlated to the replication capacity data for the same samples generated in 2014.

3.2.2 Virus titration and replication capacity assays

Methods used to titre viruses and measure their replication capacities were based on those previously described (21, 23, 26), and are described below.

3.2.2.1 Virus titration

Virus titres were conducted in duplicate for each of the 80 recombinant viruses generated. Titres served to determine the amount of virus required to achieve a multiplicity of infection (MOI) of 0.3% on day 2 of a replication capacity assay (21, 26).

Briefly, a total of 1 million GXR cells in 100 µl of R10 was added to each well of a 24 well culture plate (Corning Costar). This was followed by the addition of 400 µl of virus (harvested in 3.2.1.5 above). The culture plate was incubated at 37ºC and 5% CO_2 for 24 hours. Following incubation 1 ml of pre-warmed R10 was added to each well and the plate was incubated at 37ºC and 5% CO_2 for a further 24 hours. Thereafter, the contents of each well was thoroughly mixed and 500 µl of culture was removed and prepared for flow cytometry (described in section 3.2.1.4) in order to determine the percent of cells infected. The percent of cells infected was used to calculate the amount of virus required to achieve an MOI of 0.3% on day two in subsequent replication capacity assays. The calculation employed was as follows:

Volume of virus required (µl) = (0.3% ÷ % of cells infected)*400 µl

3.2.2.2 Replication capacity assay

Replication capacity assays (RCA's) were conducted to measure the exponential growth of viruses in GXR cells over a period of seven days. Each assay was conducted at least in duplicate, independently, for the 80 viruses generated in this study. Each assay included a negative control (GXR cells only) and a positive control (NL43-WT virus).

For the assay, a total of 1 million GXR cells in 100 µl of R10 was added to each well of a 24 well culture plate. One aliquot of stored virus was thawed. The relevant amount of virus (i.e. volume calculated as part of the titre step in 3.2.2.1) was diluted in R10 to achieve a final volume of 400 µl. The diluted virus was added to the respective wells of the 24 well culture plate, which was incubated at 37ºC and 5% CO_2 for 24 hours. After 24 hours, 1 ml of pre-warmed R10 was added to each well and incubated for a further 24 hours. Following incubation, the contents of each well was mixed thoroughly and 500 µl of culture was removed for use in flow cytometry, to determine the percent of infected cells (described in section 3.2.2.1). The removed culture was replaced with 500 µl of fresh pre-warmed R10.

The removal of culture and subsequent replacement with R10 media was repeated for the next four days, with the removed culture used for flow cytometry on each day in order to determine the percent of infected cells within the culture over a period of five days.

In order to measure viral replication capacity, the exponential increase in percent infected cells from day 3–6 post infection was calculated using the semi-log method in Microsoft Excel. Thereafter the "logest" function in Excel was used to calculate a $\log_{10}$ exponential curve, with the best fit to the data. This returned a value representative of the slope of the curve, which was then converted to its' natural log, since the natural log represents data which is commonly used in studying exponential growth in biology. All replication values for patient-derived viruses were normalised by dividing the slope of exponential growth of each virus by the slope of exponential growth of the NL43-WT reference strain. Duplicate slope values were averaged.

3.2.3 Virus titration and phenotypic drug susceptibility testing

Phenotypic drug susceptibility assays were conducted to determine the drug susceptibility profiles of 18 selected viruses. Two drugs were used; LPV and DRV (both obtained from the AIDS Research and Reference Reagent Programme, Division of AIDS, NIAID, NIH). Viruses employed in this assay were selected based on data from replication capacity assays and sequence analysis (described further in the results section).

For the current study, a TZMBL cell based two-cycle phenotypic drug susceptibility assay was employed (31). A two-cycle assay was selected, since the inhibitory effect of PIs cannot be established in a single cycle. Viral titres were also conducted using a two-cycle assay (31).

3.2.3.1 Preparation of TZMBL cells

TZMBL cells are a luciferase reporter based HeLa cell line. These cells are adherent, they express CD4 and CCR5 and have been engineered to include a luciferase gene which is under the control of the HIV-1 promoter. As such they are replication competent and allow for detection of infected cells by luminescence. TZMBL cells, were obtained from the AIDS Research and Reference Reagent Programme, Division of AIDS, NIAID, NIH.

An aliquot of TZMBL cells (1 ml) was removed from a liquid nitrogen freezer and immediately placed in a water bath at 37ºC. The cells were transferred to a 15 ml falcon tube containing 10 ml of pre-warmed Dulbecco's Modified Eagle's Medium (DMEM; Sigma), supplemented with 10% FBS (Gibco), 50 U/ml penicillin streptomycin and 10 mM

HEPES. The 15 ml falcon tube was centrifuged at 1,200 rpm for 10 minutes. Supernatant was removed, pellets were re-suspended in fresh DMEM (5 ml) and transferred to a T25 flask containing an additional 5 ml of DMEM. The T25 flask was incubated at 37ºC and 5% CO_2 for 48 hours. After 48 hours, cells were visualised using a Zoe® fluorescent imager (Bio-Rad). If cells were not 100% confluent, media was removed and replaced with fresh pre-warmed DMEM. If cells were 100% confluent, they were counted and used in experiments or seeded accordingly.

Counting of cells required that they be dislodged from the monolayer. This was achieved by removing DMEM from the flask, rinsing the monolayer with PBS and adding 2.5 ml of 0.25% trypsin-EDTA (Sigma) to the cell monolayer. The flask was incubated at room temperature for 2 minutes, followed by removal of trypsin-EDTA and subsequent incubation at 37ºC and 5% CO^2 for 4 minutes. Thereafter, 10 ml of pre-warmed DMEM was added and the wall of the T25 flask, containing the cell monolayer, was repeatedly rinsed in order to dislodge cells. The contents of the flask was thoroughly mixed. Cells were then counted as described for GXR cells in section 3.2.1.2.

The required number of cells were removed and used in experiments, whilst the remaining cells were maintained at 250,000 cells/ml in DMEM in a T25 culture flask, incubated at 37ºC and 5% CO_2. Cells were monitored by microscopy, fed and split every 48 hours. Cells were maintained for a maximum of one month, after which time a new aliquot of cells was prepared for use in experiments.

3.2.3.2 Two-cycle virus titration

As per Puertas et al., (2012), a two-cycle infection assay was conducted in the absence of drug, in order to determine the amount of virus required to obtain the mean 50% tissue culture infective dose ($TCID_{50}$) of each virus. The main aim of this approach was to standardise the virus input thereby attaining a similar rate of infection for each virus stock in order to exclude failure of the experiment due to fitness costs inherently associated with PR RAMs (31).

Virus titrations were set up in a 96 well tissue culture plate (Corning Costar). A total of 100 µl of DMEM was added to all wells of the plate. Thereafter, 25 µl of virus (generated in 3.2.1) was added to the first 3 wells (i.e. in triplicate). A 5-fold serial dilution was then

performed. All wells of column 12 were virus free; this served as the cell control. A total of 10,000 TZMBL cells in 100 µl of DMEM and 0.05 g/µl of diethylaminoethyl-dextran hydrochloride (DEAE dextran; Sigma) was added to all wells. The plate was then incubated at 37ºC and 5% CO_2 for 48 hours. Following incubation, the contents of each well was mixed and 100 µl of supernatant was removed from each well and added to the corresponding wells of a new 96 well culture plate, containing 10,000 TZMBL cells in 100 µl of DMEM and 0.05 g/µl of DEAE-dextran. The new culture plate was incubated at 37ºC and 5% CO_2 for a further 48 hours. After the second incubation (i.e. second-cycle of infection), 100 µl of culture medium was removed from each well and replaced with 100 µl of Bright Glo luciferase reagent (Promega). The plate was incubated at room temperature for 2 minutes, the contents of wells were thoroughly mixed and 100 µl of culture from each well was transferred to corresponding wells of a 96 well black solid bottom microplate (Promega). Luminescence (i.e. indicator of infectivity) was measured using a Glomax®-Multi Microplate Multimode reader (Promega).

Data expressed as relative light units (RLUs) was analysed according to the method by Reed and Muench (32). Positive infection was quantified using a cut-off of 2.5 times that of the background RLU.

3.2.3.3 Two-cycle phenotypic drug susceptibility assay

The two-cycle phenotypic assay was performed as previously described (31). Drug concentrations for LPV and DRV used in this study ranged from 1 µM to 0.00032 µM. These drug concentrations were within the range of that used in a previous susceptibility assay (33). Each assay comprised of a cell control (cells only, no virus or drug), a virus control (virus and cells only, no drug) and virus experiments (virus, cells and drug).

Briefly, in the first round of infection 10,000 TZMBL cells (in 100 µl of DMEM, 0.05 g/µl of DEAE dextran) and 3-fold serial dilutions of the respective PI was infected with the relevant amount of virus that yielded 50 $TCID_{50}$s (as calculated in 3.2.3.2), in a 96 well culture plate. The plate was incubated for 48 hours at 37ºC and 5% CO_2. Following incubation, the contents of each well was mixed and 100 µl of culture was transferred to corresponding wells of a 96 well plate containing 10,000 TZMBL cells (in 100 µl of DMEM and 0.05 g/µl of DEAE dextran). The plate was once again incubated at 37ºC and 5% CO_2 for an additional 48 hours, this constituted the second cycle of infection. After 48 hours,

150 µl of supernatant was removed and replaced with 100 µl of Bright Glo reagent. The plate was incubated for 2 minutes at room temperature. Thereafter, the contents of each well was mixed and transferred to corresponding wells of a black solid bottom microplate. HIV-1 infection was measured by luciferase expression using the Glomax luminometer

The extent to which the drug inhibited viral replication (i.e. percent inhibition) was calculated by determining the difference in RLUs between the test wells and the cell control wells (i.e. negative/cells only) and dividing this value by the difference between the virus control wells (i.e. virus without exposure to drugs) and the cell control well and multiplying the result by 100 for each virus. The calculation is given below:

%inhibition = (test wells – cell control wells) ÷ (virus control wells – cell control wells)

The concentration of drug required to inhibit viral replication by 50% (i.e. IC_{50}) was calculated by fitting the percent inhibition data to a sigmoidal dose-response curve (with a variable slope) in GraphPad Prism.

Fold change in drug susceptibility (i.e. variations in amount of drug required to achieve the IC_{50} between the virus of interest and a reference virus) was calculated by dividing the IC_{50} of each virus by the IC_{50} of the NL43-WT reference strain, which was known to be susceptible to LPV and DRV.

For this study, two classifications of drug susceptibility was used to categorize viruses these included: susceptible (S) and reduced susceptibility (RS). For viruses to be considered as susceptible, their FC was required to be below the lower FC cut-off established as part of this study. To be categorized as having reduced susceptibility the FC values were required to be greater than the lower FC cut-off level.

The lower FC cut-off level for LPV and DRV was calculated using the 99^{th} percentile of the average IC_{50} for the NL43-WT reference virus, which was known to be susceptible to LPV and DRV (50). This data was obtained from multiple (i.e. 20) independently repeated phenotypic susceptibility assays for both LPV and DRV. Table 3.1 shows the calculated lower FC cut-off values for LPV and DRV used in this study. The lower FC cut-off for LPV was 2.42 whilst DRV was 1.49. A previous study reported the lower biological cut-off for

LPV to be 1.6 and DRV to be 2, using the Antivirogram-Virco kit (33) whilst the clinical cut-off for both LPV and DRV has been reported to be 10 (54). The cut-off values used in the current study are thus similar to those which have been previously reported.

The drug susceptibility of each virus was measured in triplicate in each assay, with each virus being phenotyped in at least duplicate, in independently repeated experiments.

Table 3-1Overview of data used in the calculation of the lower FC cut-off levels for lopinavir and darunavir

	Lopinavir	Darunavir
Average	0.013	0.040
Standard deviation	0.004	0.004
95% confidence interval	0.011	0.038
Number (n=)	20	20
Minimum	0.010	0.033
Median	0.012	0.040
Maximum	0.024	0.048
(Q1)	0.010	0.040
(Q3)	0.013	0.042
1st Percentile	0.010	0.033
99th Percentile	0.024	0.048
Lower FC cut-off	**2.42**	**1.49**

Table 3-2 FC cut-off levels used to categorize viruses as being susceptible or conferring reduced susceptibility to LPV and DRV.

	Susceptible (S)	Reduced susceptibility (RS)
Lopinavir	<2.42	>2.42
Darunavir	<1.49	>1.49

3.2.4 Data analysis

3.2.4.1 Replication capacity assay: Codon by codon analysis

In order to assess the association of replication capacity and codon variations, at each codon in Gag-Protease, a codon by codon analysis tool, developed by the Brockman group at the Simon Fraser University, was used (34). As part of the analysis several automated Mann Whitney U tests were conducted and p values were generated for each comparison of amino acid variation and replication capacity. Correction for multiple comparisons, which yielded a q value (i.e. the equivalent of the p value for false discovery

rate) was conducted. All analysis with n<5 (i.e. occurrence of an amino acid variant less than five times) were excluded from further investigation in order to remove uncommon amino acid substitutions from the dataset. All variations with a p<0.05 and q<0.02 were considered to be significantly associated with replication capacity. These criteria were in line with two previous studies in which similar analysis was conducted (26, 35).

3.2.4.2 Statistical analysis

To measure concordance between data either a Pearson's correlation test (for parametric data) or a Spearman's Rank correlation test (for non-parametric data) was used. Students T-tests (for parametric data) or Mann Whitney U tests (for non-parametric data) were used to compare continuous data between two groups. For comparisons of continuous data between more than two groups, One-way Analysis of Variance (ANOVA) with Tukey post hoc testing (for parametric data) or Kruskall Wallis with the Dunns test for multiple comparisons (for non-parametric data) was used. All continuous data is presented as the median with an interquartile range unless otherwise stated. Statistical analysis was conducted using Graphpad Prism version 5 (Graphpad Software) and R statistics version 3.2.2 (36). Statistical significance was defined as p<0.05.

3.3 Results

3.3.1 Replication capacity assay

3.3.1.1 Validation of Recombinant viruses vs plasma derived HIV-1 RNA sequences

Gag-Protease sequences from eight randomly selected recombinant viruses were compared to matched sequences from patient derived HIV-1 RNA plasma, in order to determine if sequences of recombinant viruses were representative of plasma sequences. Sequences had a median nucleotide similarity of 99.6% (IQR: 99.1% - 99.75%), as such plasma derived sequences were used for all further analysis. The neighbour joining tree showed clustering of recombinant viruses (red) and plasma sequences (green) for matched patient samples (Figure 3.3). No cross contamination occurred.

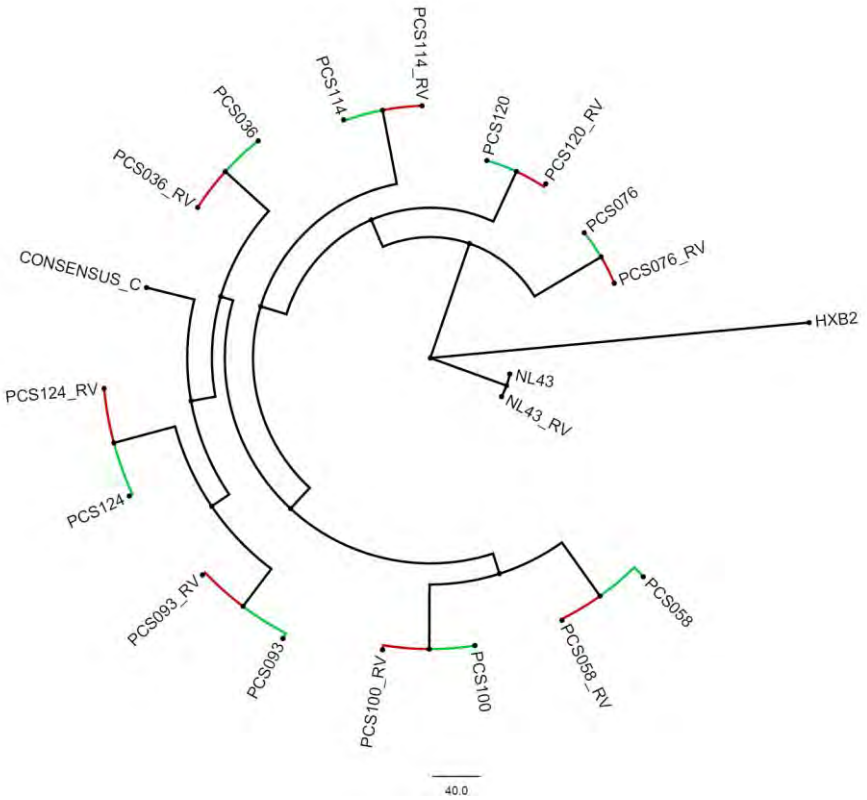


Figure 3-3 Neighbour joining phylogenetic tree (Paup version 4.0) representing matched sequences from plasma and recombinant viruses. Data shows clustering of plasma sequences (green) and recombinant virus sequences (red) for matched patient samples.

3.3.1.2 Validating reproducibility of the replication assay

Each replication assay was conducted in duplicate. Replicate assays were performed at least three days apart. All replication capacity values were normalised to the growth of NL43-WT. Pearson's correlation of the duplicate assays showed high concordance of replication capacity between replicates (Pearson's correlation: $r=0.93$, $p<0.0001$) (Figure 3.4). This highlights the reproducibility of the assay.

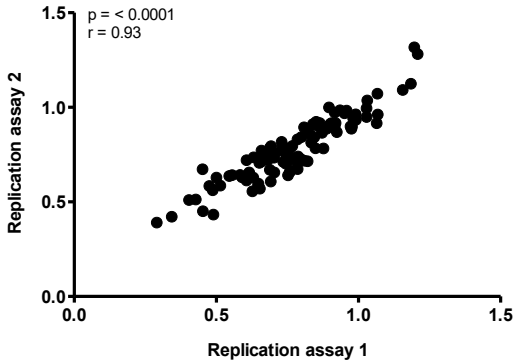


Figure 3-4 Comparison of duplicated replication assays. Pearson's correlation showed a high concordance between the two data sets indicating reproducibility of the assay.

3.3.1.3 Validating replication assay consistency

Replication capacity measured for samples in 2008 was compared to replication capacity of the same samples measured in 2014. Results showed a high concordance between data (Spearman's correlation: r_s=0.92, p<0.0001) indicating consistency in the performance of the assay (Figure 3.5).

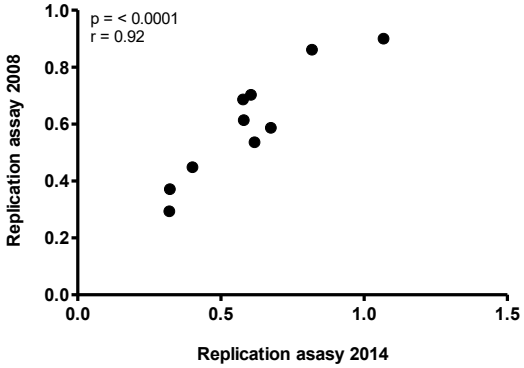


Figure 3-5 Comparison between replication data generated for ten samples in 2008 and 2014. Spearman's correlation showed a high concordance between datasets indicating consistency in performance of the assay**.**

3.3.1.4 Frequency distribution of replication capacity

The Kolmogorov test for normality showed that RCA data approximated a normal distribution (KS = 0.08, p>0.05) (Figure 3.6). The mean for RCA data was 0.78 with a standard deviation of 0.17.

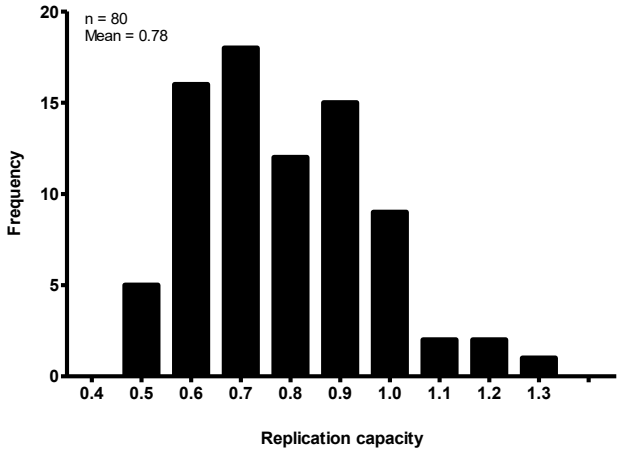


Figure 3-6 Frequency distribution of replication capacity in a cohort of 80 participants failing a PI inclusive treatment regimen. Data approximated a normal distribution with a mean of 0.78 and a standard deviation of 0.17**.**

3.3.1.5 Correlation of viral load and replication capacity

The viral load and age of all 80 participants from the PCS cohort was correlated to replication capacity. Viral load showed no correlation to replication capacity (Pearson's correlation: r=-0.02, p=0.8065) (Figure 3.7). This shows that viral load did not influence replication capacity.

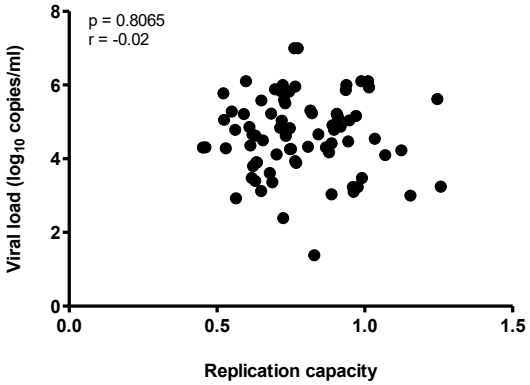


Figure 3-7 Associations between replication capacity and viral load. Pearson's correlation showed no concordance between data indicating that viral load did not influence replication capacity.

3.3.1.6 Percent similarity to Gag-Protease consensus C vs replication capacity

The percent similarity between a Gag and Protease consensus subtype C sequence and 80 patient-derived Gag and Protease sequences was calculated using the similarity matrix tool in Bioedit. This data was correlated with replication capacity to determine if variations in Gag and Protease sequences were associated with replication capacity.

There was no relationship between percent similarity of Gag sequences and replication capacity (Spearman's correlation: r_s= -0.03, p=0.7865) (Figure 3.8a). To determine if this was true for all regions in Gag, data was stratified to represent each of six regions in Gag (i.e. MA, CA, P2, NC, P1 and P6) (Figure 3.8c). The similarity of each stratified region to the consensus subtype C sequence was established and used in correlation analysis. Data showed no correlation between sequence similarity and replication capacity in each of the six regions analysed (Figure 3.8c).

Interestingly, a borderline significant trend towards increased replication capacity for viruses which were more similar to the consensus C PR sequences was observed (Spearman's correlation: r_s=0.25, p=0.058) (Figure 3.8b). This indicates that viruses which diverge from the consensus are likely to be associated with reduced replication capacity (i.e. viruses with polymorphisms in PR are likely to be associated with lower replication capacity). Therefore for PR, the fittest viruses were the most similar to the consensus C PR sequence.

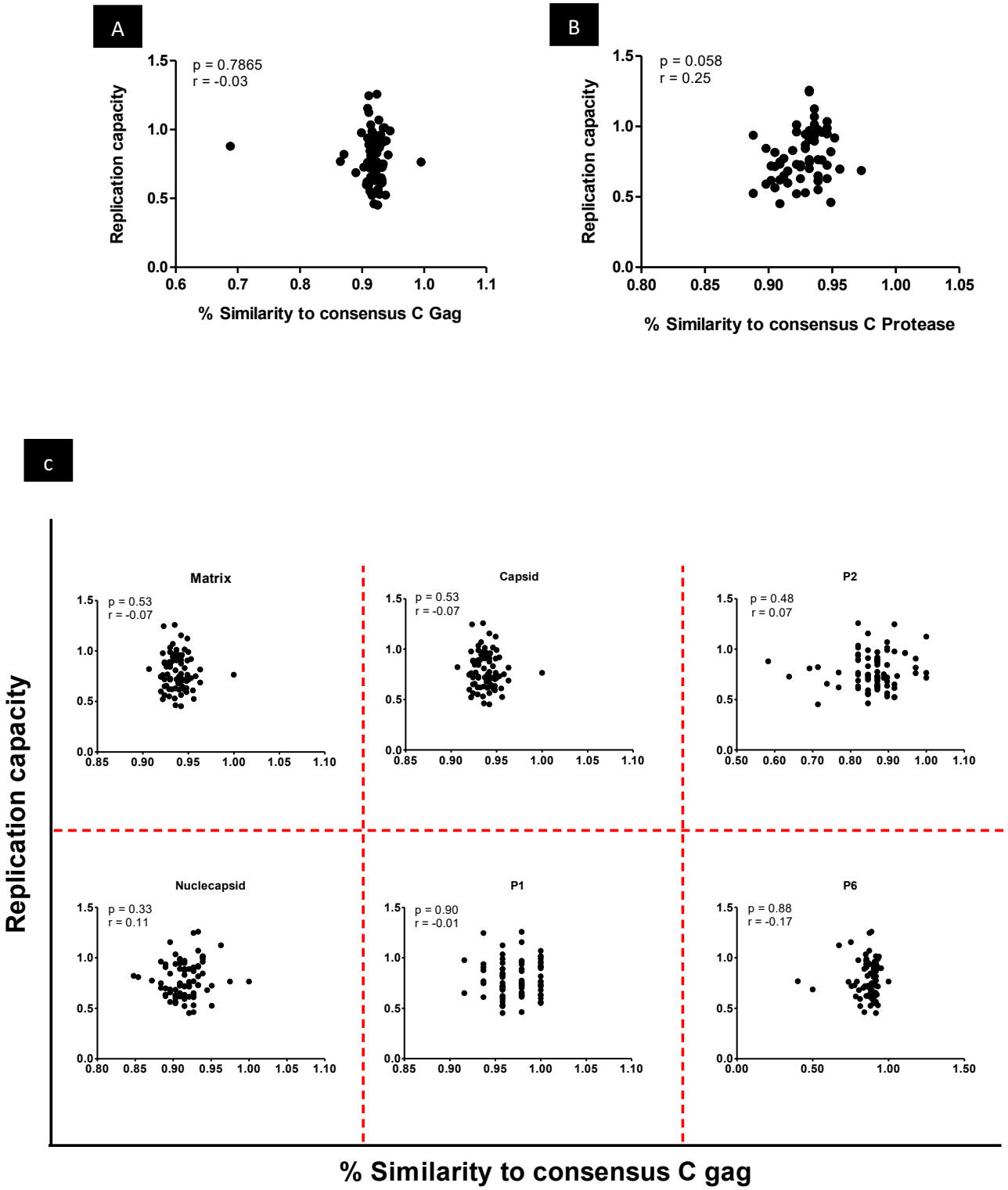


Figure 3-8 Association between polymorphisms in Gag-Protease and replication capacity. (a) No relationship was noted between percent similarity of Gag sequences and replication capacity (Spearman's correlation). (b) A borderline significant positive correlation was noted for viruses with polymorphic PR, i.e. more fit viruses had the most similar sequences to the consensus sequence. (c) No relationship between polymorphisms within the 6 regions of Gag (i.e. Matrix, capsid, p2, nucleocapsid, p1 and p6) and replication capacity was observed.

3.3.1.7 Number of mutations in Gag and Protease versus replication capacity

The relationship between number of mutations in Gag and PR and replication capacity was assessed. The number of mutations in Gag (Figure 3.9a) and PR (Figure 3.9b) did not show an association with replication capacity (Spearman's correlation: r_s=-0.09, p=0.53 and r_s=-0.258 and p=0.14 respectively). This was unexpected for PR and could be attributed to a small sample size.

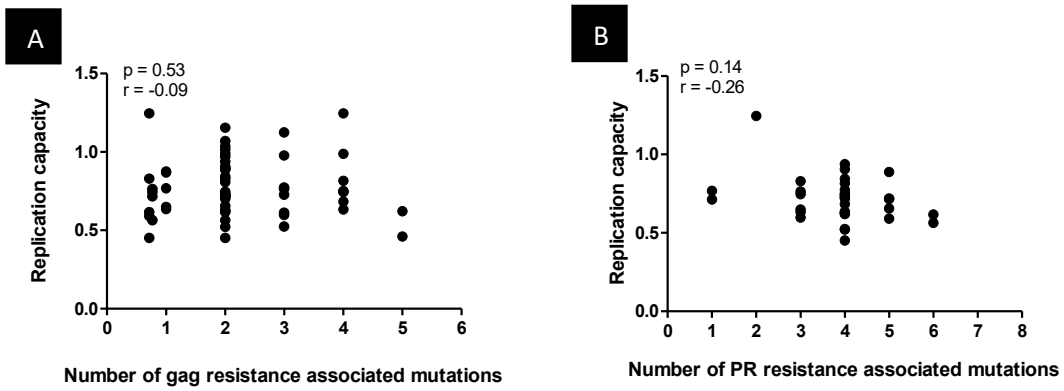


Figure 3-9 Overview of relationship between the number of mutations in Gag (a) and Protease (b) on replication capacity. The number of mutations in Gag and Protease did not correlate with replication capacity.

3.3.1.8 Replication capacity: Gag-Protease mutations

Replication capacity was compared between recombinant viruses with PR RAMs (n=34) and without PR RAMs (n=46). As expected, viruses with PR RAMs had a significantly lower replication capacity than viruses without PR RAMs (Student T-test: p<0.001) (Figure 3.10a).

Data was further stratified to determine the impact of combinations of PR RAMS and rGag mutations on replication capacity. For this analysis replication capacity was compared amongst: viruses harbouring PR RAMs and rGag mutations (n=34), viruses with rGag mutations only (n=33) and viruses without PR RAMs and rGag mutations (n=13). Viruses with PR RAMs and rGag mutations had a significantly lower replication capacity than viruses with rGag mutations only and viruses devoid of Gag and PR mutations (One-way ANOVA with Tukey test for multiple comparisons: p<0.01) (Figure 3.10b).

Collectively this data is in line with several previous studies, showing that PR RAMs occur at a fitness cost (37-39).

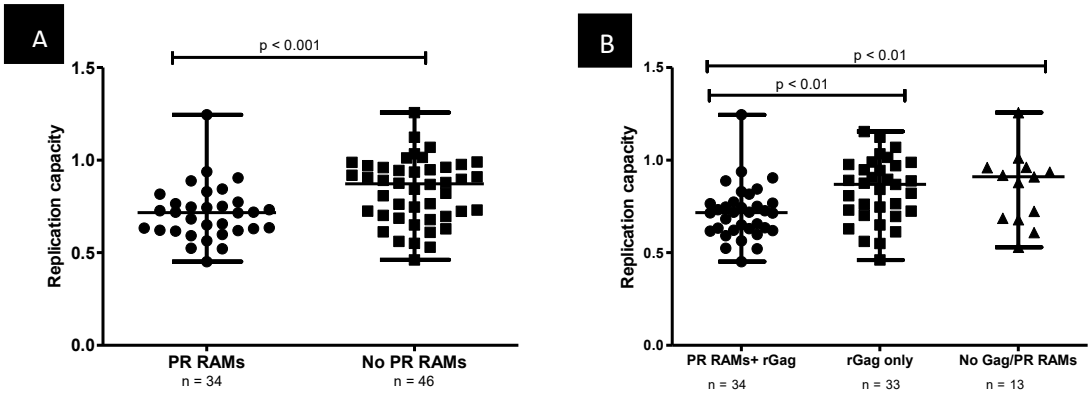


Figure 3-10 Comparison of the effect of combinations of Gag and PR RAMs on replication capacity. (a) Viruses with PR RAMs had a significantly lower replication capacity in comparison to viruses without PR RAMs ($p<0.001$). (b) Viruses with PR RAMs and and rGag mutations had a significantly lower replication capacity than viruses with rGag mutations only ($p<0.01$) and viruses without rGag or PR mutations ($p<0.01$). Abbreviations: PR – Protease, RAMs – resistance associated mutations.

The effectof rGag and nGag mutations on RC when found in conjunction with PR RAMs could not be assessed since all viruses with PR RAMs harboured both rGag and nGag mutations. The impact of rGag and nGag mutations on replication capacity, in the absence of PR RAMs, was however assessed. Comparisons were made amongst: viruses with nGag mutations only (n=7), viruses with rGag mutations only (n=17), viruses with both rGag and nGag mutations (n=16) and viruses without Gag mutations (n=6) (Figure 3.11). Results showed no significant difference in replication capacity amongst these groups indicating that viral replication capacity may be unaffected by rGag or nGag mutations in the absence of PR RAMs.

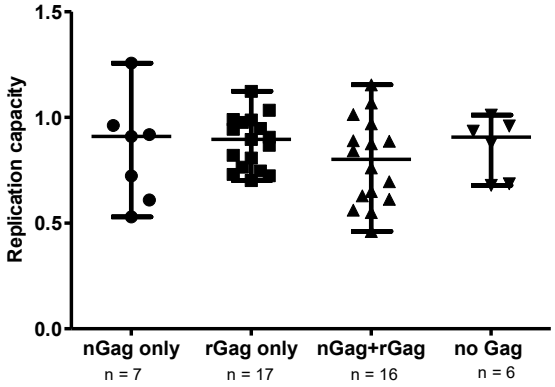


Figure 3-11 Comparison of replication capacity amongst viruses harbouring novel Gag (nGag) mutations only, viruses harbouring resistance Gag (rGag) mutations only, viruses harbouring a combination of resistance and novel Gag mutations and viruses without Gag mutations. No significant difference in replication capacity was observed between all four groups that were compared. Abbreviations: rGag – resistance Gag mutations; nGag – novel Gag mutations associated with PI resistance/exposure.

Overall, PR RAMs appear to be associated with significantly reduced replication capacity. In the absence of PR RAMs, rGag mutations were found to have no significant impact on RC indicating that wild-type PR can still cleave mutant Gag. This may suggest that un-mutated PR recognises both mutated Gag and wild-type Gag.

3.3.1.9 Replication capacity: amino acid variants

The relationship between replication capacity and variations at each codon in Gag and PR was assessed next using the codon by codon analysis tool.

A total of five amino acid variations were significantly associated with replication capacity prior to corrections for multiple comparisons (i.e. $p<0.05$; Table 3.3). These included: K28H, G62R, R91N, K335R and A431V (Table 3.1). The K28H, G62R and R91N amino acid substitutions were all associated with significantly increased replication capacity, whilst the K335R and A431V substitutions were associated with significantly reduced replication capacity ($p<0.05$). Following corrections for multiple comparisons, only two amino acid variations (i.e. K335R and A431V) remained to be significantly associated with replication capacity (i.e. $p<0.05$ and $q<0.02$) (Table 3.3).

Table 3-3 Amino acids associated with altered replication capacity in Gag-pro recombinant viruses derived from patients

Gag protein	Codon	AA+	Consensus	Median Replication Capacity		Number of viruses		p value	q value
				AA+	Consensus	AA+	Consensus		
Matrix	28	Q	K	0.8605	0.7455	10	70	0.04969	0.884918
Matrix	62	R	G	0.9175	0.7485	8	72	0.04894	0.884918
Matrix	91	N	R	0.914	0.747	8	72	0.02983	0.871699
Capsid	335	R	K	0.727	0.907	63	17	0.000258	0.022609
Nucleocapsid	431	V	A	0.683	0.856	24	56	8.63E-05	0.022609

Literature has shown that the A431V mutation is associated with PI resistance/exposure (discussed in Chapter 2), however no data is available in this regard for K335R. In order to determine if K335R is associated with PI resistance/exposure, comparisons in frequency of K335R amongst the PCS, HIV-1 subtype C treatment naïve and acute cohort were made (Figure 3.12). Results showed that the K335R substitution occurred at similar frequencies in all three groups, indicating that it is not associated with PI exposure, but is rather a natural polymorphism in HIV-1 subtype C.

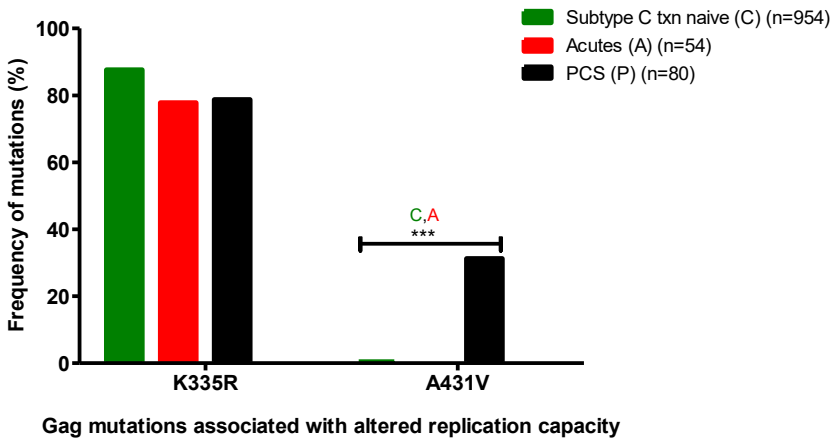


Figure 3-12 Comparison of frequency of K335R and A431V amongst sequences from the: PCS cohort, HIV-1 subtype C treatment naïve cohort and the HIV-1 subtype C acute cohort. The A431V mutation occurred at a significantly higher frequency in the PCS cohort in comparison to subtype C treatment naïve (C) and acute (A) cohort. In contrast, the K335R mutation occurred at similar frequencies in all three HIV-1 subtype C cohorts analysed. Significant differences are denoted by asterisks, were *** represents p<0.0001. Alphabets (A and C) are used to denote the groups (i.e. acute [A] and HIV-1 subtype C treatment naïve [C]) with significant results.

In order to establish the impact of combinations of PR RAMs, K335R and A431V on replication capacity, one way ANOVA with Tukey post hoc testing was performed. Results showed that viruses with PR RAMs and K335 (consensus) had significantly higher replication capacity than viruses with PR RAMs and 335R (substitution) ($p<0.01$). Similarly, viruses with K335 in the absence of PR RAMs had a significantly higher replication capacity than viruses with 335R in the absence of PR RAMs ($p<0.01$). The 335R substitution is thus associated with reduced replication capacity both in the absence and presence of PR RAMs (Figure 3.13a).

For A431V, the 431V substitution seldom occurred without PR RAMs (n=1), therefore its impact on replication capacity in the presence/absence of PR RAMs could not be established. Viruses harbouring A431 (i.e. consensus) in the absence of PR RAMs had a significantly higher replication capacity than viruses harbouring 431V in the presence of PR RAMs ($p<0.001$). Viruses with 431V and PR RAMs showed no difference in replication capacity when compared to viruses harbouring A431 and PR RAMs.

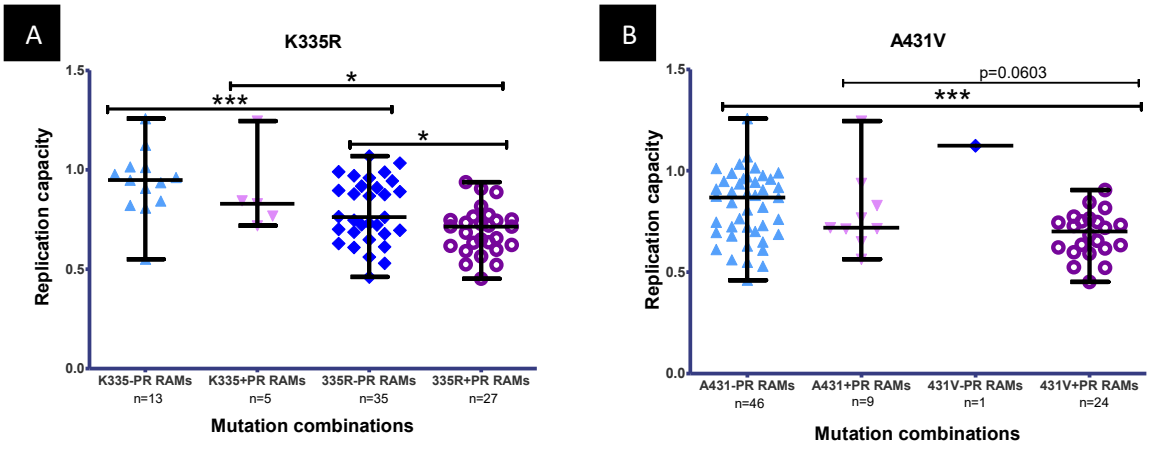


Figure 3-13 Effect of K335R and A431V on replication capacity when found in the presence (+PR RAMs) or absence (-PR RAMs) of Protease resistance associated mutations (PR RAMs). (a) Viruses with the 335R substitution were associated with significantly reduced replication capacity when found in the presence and absence of PR RAMs. (b) Viruses with the 431V substitution and PR RAMs showed no variation in replication capacity when compared to viruses with A431 and PR RAMs. Significance is denoted by asterisks, were * represents $p<0.01$, ** represents $p<0.001$ and *** represents $p<0.0001$. The number of viruses with each mutation combination is denoted by n.

In summary, this data shows that the 335R substitution is associated with reduced replication in both the presence and absence of PR RAMs, however it occurs as a natural

polymorphism in HIV-1 subtype C and is thus not associated with PI exposure. The 431V substitution seldom occurs in the absence of PR RAMs and thus its effect on replication capacity when found without PR RAMs could not be established. Interestingly though, no significant difference in replication capacity was noted for viruses with A431+PR RAMs versus viruses with 431V+PR RAMs, suggesting that both an A or a V at position 431 have the same impact on replication capacity in the presence of PR RAMs.

3.3.1.10 Replication capacity: Novel Gag mutations in the presence/absence of PR mutations

Replication capacity of recombinant viruses harbouring nGag mutations (i.e. 69K, 111I/C, 239A/S and 256V), with or without PR RAMs, was analysed next. This analysis aimed to determine if nGag mutations play a role in viral replication.

Viruses with Q69 and T239 (i.e. consensus amino acids) were associated with significantly reduced replication capacity in the presence of PR RAMs in comparison to when they occurred without PR RAMs ($p=0.0004$ and $p<0.001$ respectively) (Figure 3.14). Substituting the amino acids at these codons with the mutant variant (i.e. 69K, or 239A/S), had no impact on replication capacity when the substitution occurred with/without PR RAMs. Interestingly though, viruses with 69K-PR RAMs had the same median replication capacity as viruses with Q69-PR RAMs therefore one would have expected to see a significantly reduced replication capacity when 69K (substitution) occurred together with PR RAMs, as did occur when Q69 was found with PR RAMs. Furthermore, it was observed that viruses with 69K+PR RAMs had a slightly higher median than viruses with Q69+PR RAMs. This could suggest that 69K helps to maintain or slightly improve replication capacity in the presence of PR RAMs (Figure 3.14).

Viruses with 111I/C or 256V (i.e. substitution) displayed significantly lower replication capacity when found with PR RAMs in comparison to when found without PR RAMs ($p=0.0189$ and $p=0.0016$ respectively) (Figure 3.14).

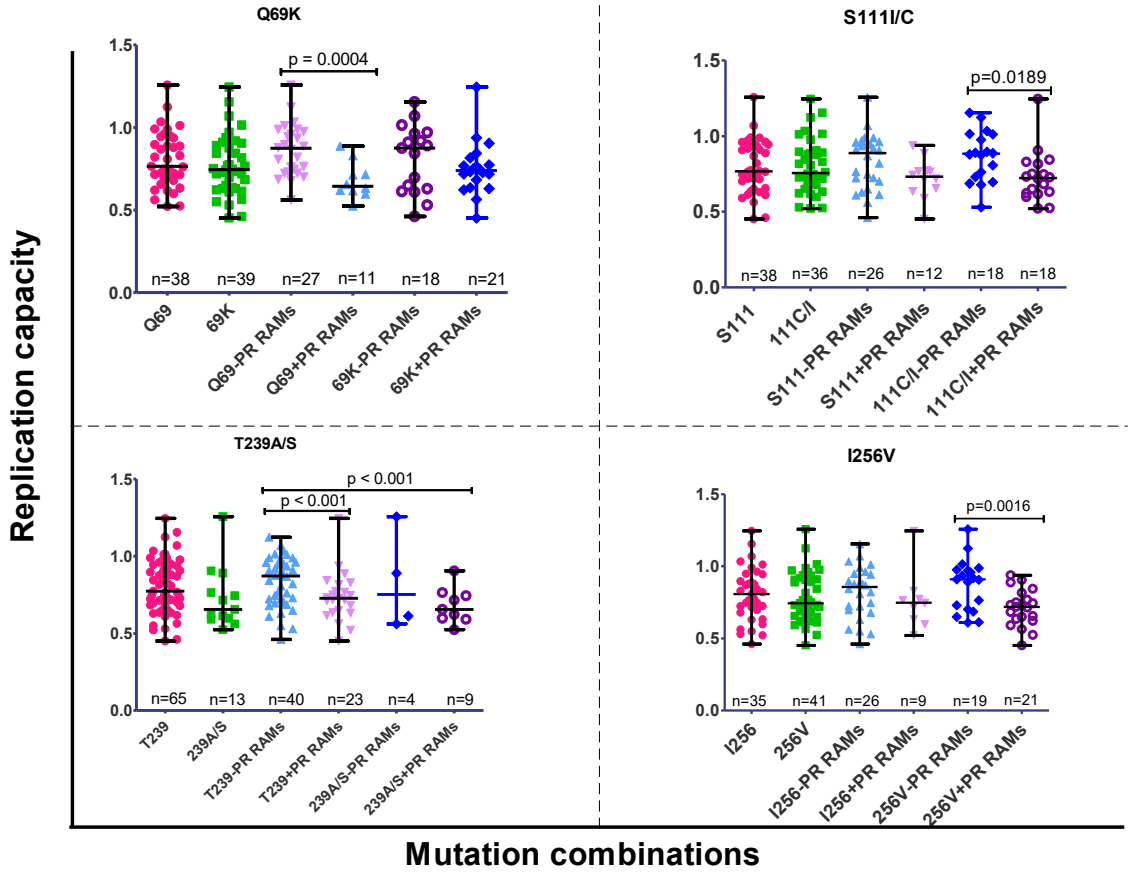


Figure 3-14 Overview of the impact of novel Gag mutations found in combination with/without Protease resistance associated mutations (PR RAMs) on replication capacity within 80 recombinant viruses. The presence of a PR RAM is indicated as +PR RAM whilst the absence of a PR RAM is indicated as –PR RAM.

3.3.1.11 Replication capacity: A431V and novel Gag mutations versus PR RAMs

The next analysis performed aimed to determine if novel Gag mutations (69K, S111I/C, 239A/S and 256V) played a role in replication capacity when found in conjunction with 431V/335R and PR RAMs. 431V and 335R were selected for this analysis since they were both found to be significantly associated with reduced replication capacity (see 3.3.1.9). For these analysis Student's T-tests were used to compare the replication capacity of viruses with 431V or 335R occurring in conjunction with PR RAMs in either the presence or absence of each nGag mutation.

Viruses with 431V and PR RAMs (n=24) occurring in conjunction with the 69K (n=16) nGag mutation had a significantly higher replication capacity than viruses with 431V and

PR RAMs occurring without 69K (n=8) (p=0.0039) (Figure 3.15). This suggests that the 69K mutation could have a compensatory role in replication capacity when found in the presence of PR RAMs and 431V (Figure 3.15).

All other nGag amino acid substitutions (i.e. 111I/C, 239A/S and 256V) showed no association with altering replication capacity in the presence of PR RAMs and 431V, suggesting an alternate role for these amino acid substitutions in PI resistance/exposure (Figure 3.15).

Whilst 69K appeared to play a role in improving replication capacity when occurring in conjunction with 431V and PR RAMs, no such role was evident for viruses with 335R and PR RAMs with any of these four nGag amino acid substitutions (data not shown). This data could indicate that compensatory mutations only function in the presence of rGag mutations (such as 431V) or that compensatory mutations only function in specific combinations of mutations.

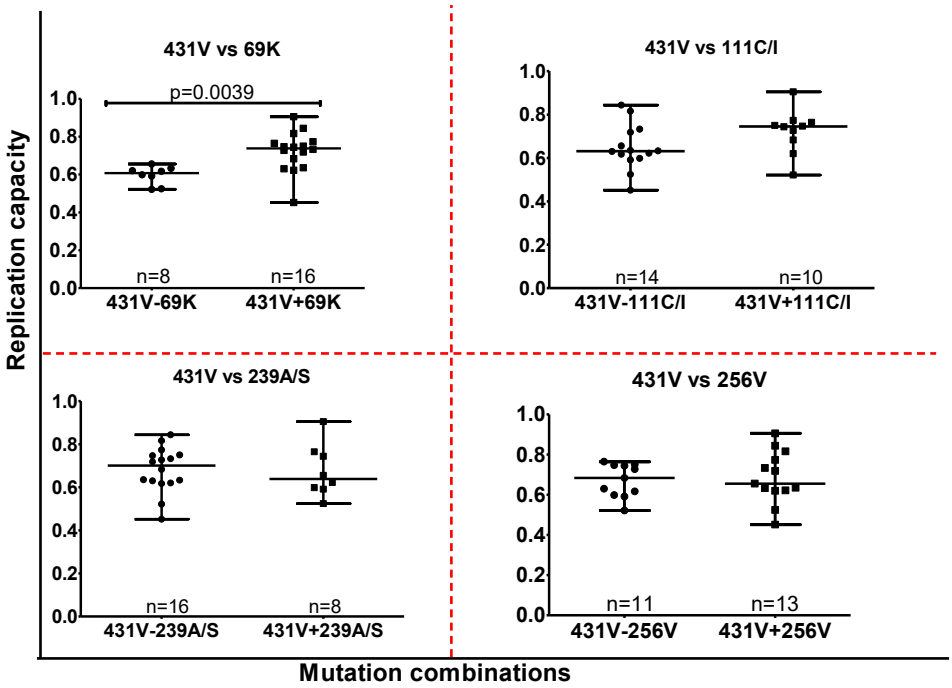


Figure 3-15 Overview of the effect of novel Gag mutations (i.e. 69K, 111I/C, 239A/S and 256V) on replication capacity when occurring in conjunction with 431V and Protease resistance associated mutations (PR RAMs). Note all viruses used in these analysis harboured PR RAMs. Mutation combinations are represented with a (-) and (+) which denotes the absence or presence of a mutation respectively.

3.3.2 Phenotypic drug susceptibility

Since the 69K nGag mutation and the 431V rGag RAM were found to be associated with altered replication capacity (Figure 3.15), their effect on susceptibility to LPV and DRV was assessed using 18 representative recombinant viruses.

3.3.2.1 Quality control of the phenotypic drug resistance assay

Drug susceptibility was measured in duplicate, for each of 18 recombinant viruses using DRV and LPV. Replicate assays were performed at least two days apart. The IC_{50} values for replicate one and replicate two, for both LPV and DRV based assays, showed good concordance indicating that the assay was reproducible (Spearman's correlation: r_s=0.8673 and p<0.0001; r_s=0.9022 and p<0.0001 respectively) (Figure 3.16).

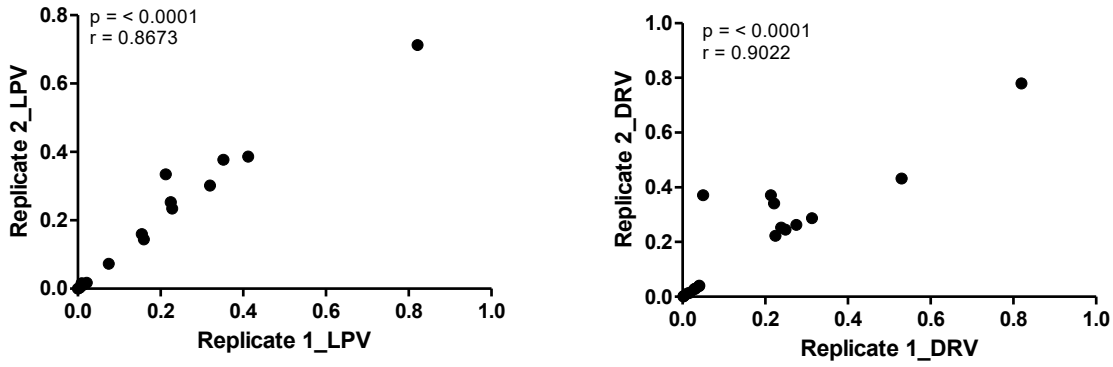


Figure 3-16 Comparison of replicate data for viruses phenotyped using lopinavir (LPV) and darunavir (DRV). Spearman's correlation showed good concordance between data indicating that the assay was reproducible.

3.3.2.2 Correlation of viral load, replication capacity and IC50

No correlation between viral load and IC_{50}, and replication capacity and IC_{50} was observed for viruses phenotyped with LPV or DRV. This indicated that IC_{50} was independent of viral load and replication capacity (Figure 3.17a-b).

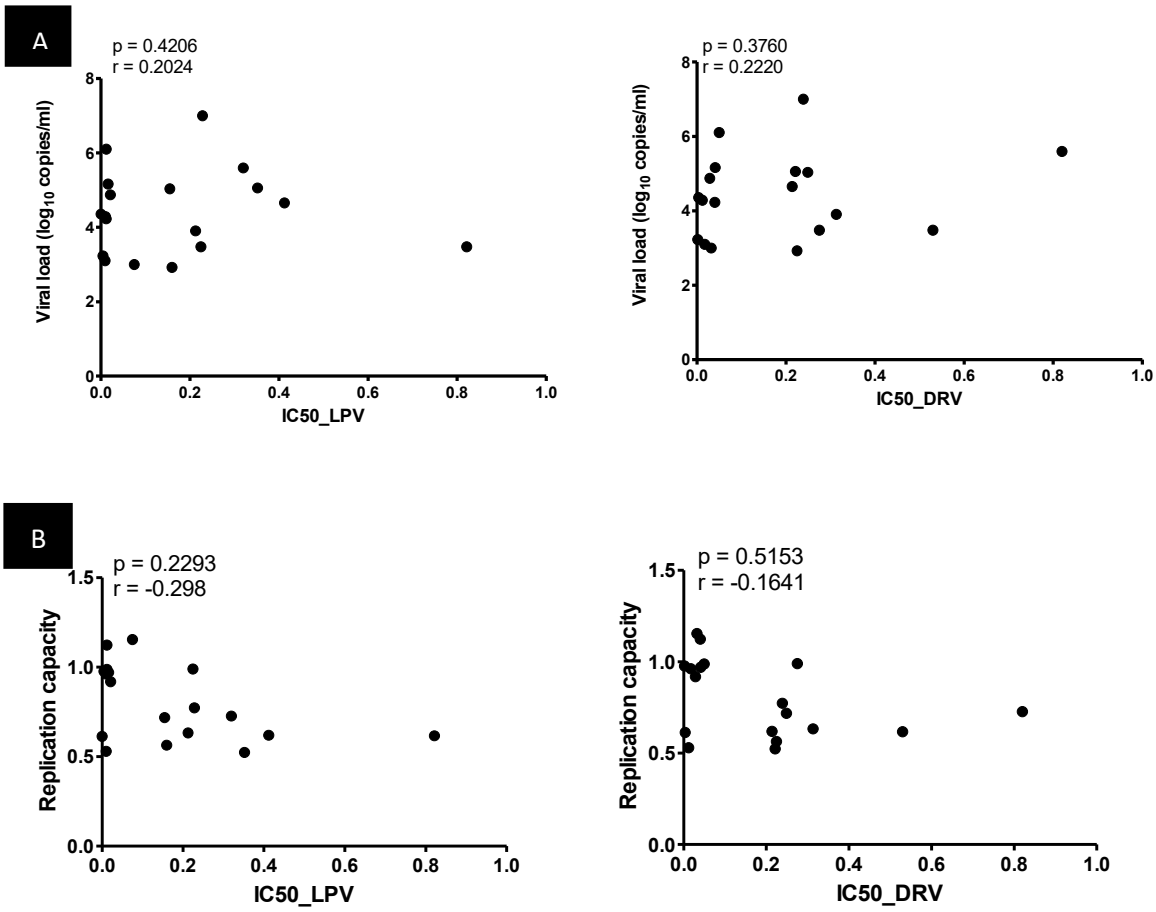


Figure 3-17 Correlation of IC_{50} with viral load and replication capacity for viruses treated with (a) lopinavir (LPV) or (b) darunavir (DRV). No correlation between IC_{50} and viral load or replication capacity was demonstrated for both LPV and DRV.

3.3.2.3 Number of mutations in Gag and Protease versus drug susceptibility

The relationship between number of mutations in Gag and Protease and drug susceptibility was assessed. The number of PR RAMs correlated positively with LPV and DRV IC_{50} (i.e. the higher the number of PR RAMs, the higher the IC_{50} for LPV and DRV) (Spearman's correlation: r_s=0.7404, p=0.0004, and r_s=0.7255, p=0.0007, respectively) (Figure 3.18a). Conversely, the number of mutations in Gag did not correlate significantly with the IC_{50} of LPV or DRV (Figure 3.18b).

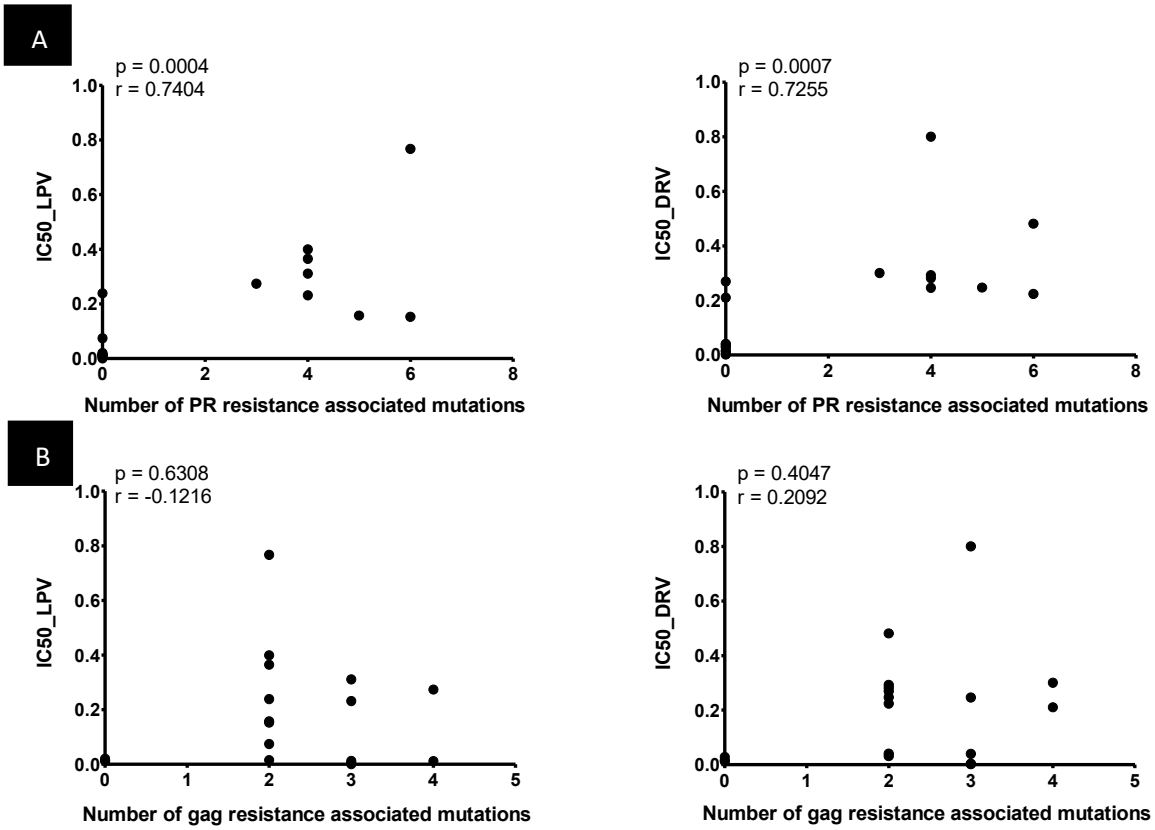


Figure 3-18 Comparison between number of Protease (a) and Gag (b) resistance associated mutations and IC_{50} for lopinavir (LPV) and darunavir (DRV). (a) Spearman's correlation showed that the number of PR RAMs correlated positively with IC_{50} for both LPV and DRV. (b) No significant correlation was noted between the number of Gag RAMs and IC_{50} for LPV or DRV.

3.3.2.4 Percent similarity to consensus vs drug susceptibility

The percent similarity of recombinant viruses to a consensus C sequence was computed. The values generated (as described in section 3.3.1.6) were correlated with IC_{50} of LPV and DRV treated viruses in order to establish if sequence variation was associated with IC_{50}.

There was a strong negative correlation between percent similarity of PR and IC_{50} for viruses treated with both LPV and DRV (Spearman's correlation: r_s=-0.7657, p=0.0009, and r_s=-0.6762, p=0.0056 respectively) (Figure 3.19a). This showed, as expected, that viruses with similar PR sequences to the consensus C sequence had lower IC_{50} for both LPV and DRV. No significant relationship between sequence variation in Gag and IC_{50} for both LPV and DRV was observed (Figure 3.19b). This indicated that variations in Gag

were not significantly associated with altered IC_{50} for viruses treated with either LPV or DRV.

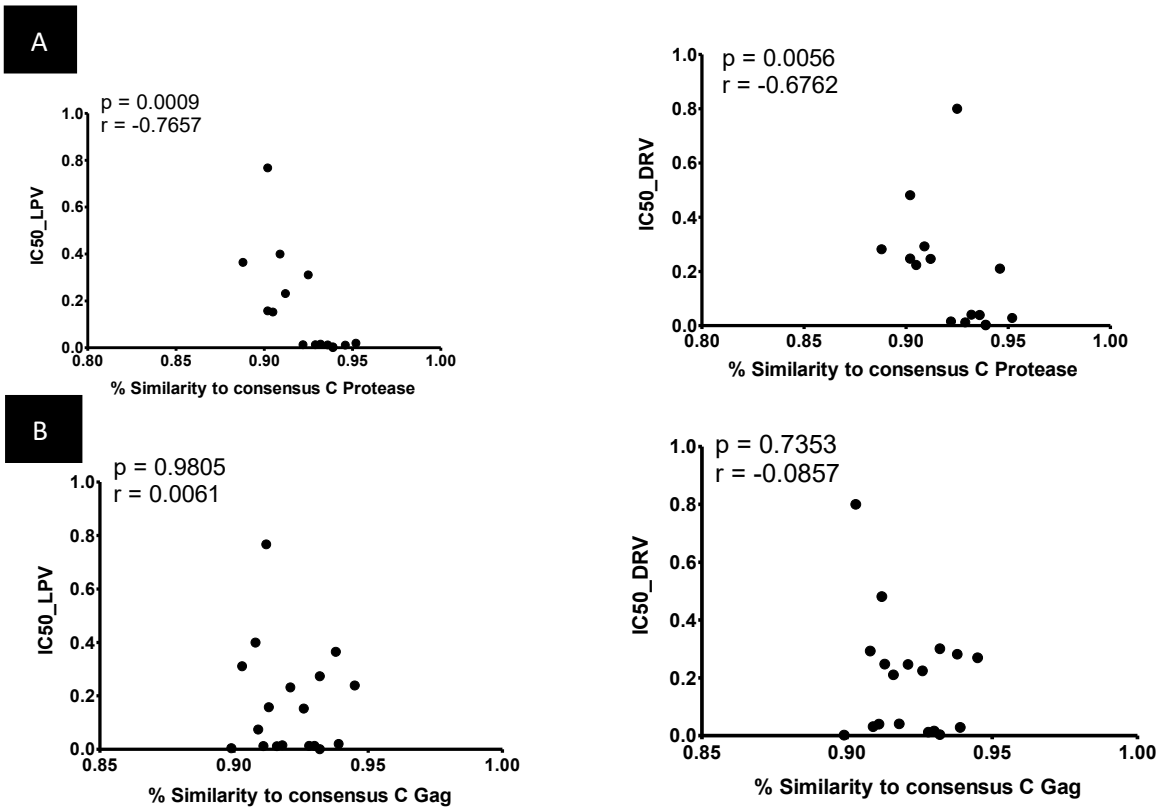


Figure 3-19 Correlation between percent similarity of sequences for both Protease (a) and Gag (b) and IC_{50} of viruses treated with LPV or DRV. (a) There was a strong negative correlation between percent similarity of PR and IC_{50} for LPV and DRV. (b) No correlation between percent similarity of Gag and IC_{50} for viruses treated with LPV or DRV was shown.

3.3.2.5 Drug susceptibility of patient-derived Gag-Protease recombinant viruses

The 18 recombinant viruses assessed were stratified according to the presence of PR RAMs, rGag mutations (including 431V) and nGag mutations (including 69K). Viruses were divided into five groups for analysis as follows:

- Four viruses with PR and rGag mutations (all viruses included harbored the 431V rGag mutation. Referred to as PR + rGag in Table 3.4);
- Four viruses with PR RAMs, rGag mutations and nGag mutations (all viruses included harbored 69K. Referred to as PR RAMs + rGag + nGag in Table 3.4);

- Four viruses with rGag RAMs only (all viruses included harbored any rGag mutation, since 431V only occurred in the absence of PR RAMs in 1 virus. Referred to as rGag in Table 3.4);
- Three viruses with nGag mutations only (all viruses included harbored the 69K mutation. Referred to as nGag in Table 3.4) and
- Three viruses with rGag and nGag mutations only (all viruses in this group harbored 69K. Referred to as rGag **+** nGag in Table 3.4).

Table 3.4 provides an overview of the 18 viruses used in the phenotypic drug susceptibility assay. Information is provided on the PR RAMs, rGag mutations and nGag mutations harboured by each virus. The FC in IC_{50} and associated standard deviation, for LPV and DRV, is provided for each virus. Additionally, the resistance profile/classification of each virus is provided, where S represents susceptible viruses and RS represents viruses with reduced susceptibility (cut-off values are described in section 3.2.3.3, Table 3.2).

Of the 18 viruses, ten (i.e. PCS002, PCSM002, PCS089, PCS134, PCS100, PCS071, PCS033, PCS153 and PCM029, PCS120) had reduced susceptibility to LPV and eight viruses (i.e. PCS011, PCS020, PCS049, PCS022, PCS096, PCS0128, PCS115 and PCS63) were susceptible to LPV (Table 3.4). Nine viruses that had reduced susceptibility to LPV also displayed reduced susceptibility to DRV, however the FC in IC_{50} was higher for LPV than DRV, suggesting that LPV IC_{50} is affected more so than DRV IC_{50} (Table 3.4). All eight viruses with PR RAMs displayed reduced susceptibility to LPV and DRV, indicating that viruses with PR RAMs require higher concentrations of PI's to inhibit viral replication than that required by wild-type viruses.

Of the four viruses with rGag associated mutations only, one displayed reduced susceptibility to LPV and DRV (PCM029) whilst three were susceptible to LPV (PCS011, PCS020 and PCS049). Interestingly PCS049 harboured the 431V Gag resistance associated mutation which has been previously shown to confer resistance to all PI's except DRV in the absence of PR RAMs. Whilst this virus was susceptible to DRV, it did not show a reduced susceptibility to LPV which was surprising (6, 7).

All three viruses harbouring the Q69K nGag amino acid substitution (PCS022, PCS096 and PCS128), were susceptible to LPV and DRV.

Two out of three viruses with rGag associated mutations and nGag amino acid substitutions (PCS115 and PCS063), were susceptible to both LPV and DRV, however one virus harbouring R76K, K436R and Q69K (i.e. PCS120) showed an increase of 7.40 fold in IC50 for LPV. This suggests that combinations of Gag mutations may contribute to reduced susceptibility of PI's in the absence of PR RAMs, as shown in previous studies (4, 6, 7).

Overall the data shows that viruses with PR RAMs require a higher concentration of LPV/DRV to inhibit viral replication by 50% when compared to viruses without PR RAMs.

Table 3-4 Overview of FC in IC50 of lopinavir and darunavir for viruses with combinations of novel Gag (nGag), resistance Gag (rGag) or Protease resistance associated mutations (RAMs).

Mutation combination	PID	PR RAMs	rGag RAMs	nGag RAMs	FC LPV	STDEV	Res profile[a]	FC DRV	STDEV	Res profile
PR RAMs + rGag	PCS002	V32I, M46I, I54V, L76V, V82A, L90M	A431V, K436R	S111C, I256V	76.71	0.07715	RS	12.03	0.0689	RS
	PCM002	M46I, I54V, V82A	R76K, Y79F, A431V, P453L	S111C	27.33	0.08605	RS	7.51	0.0182	RS
	PCS089	M46I, I54V, L76V, V82A	R76K, A431V	S111C	39.91	0.01846	RS	7.31	0.1113	RS
	PCS134	M46I, I54V, L76V, V82A	R76K, A431V	T239A/S, I256V	36.44	0.01782	RS	7.04	0.0849	RS
PR RAMs + rGag +nGag	PCS100	M46I, I54V, L76V, V82A	Y79F, V128I, A431V	Q69K	23.13	0.00417	RS	6.15	0.0101	RS
	PCS071	M46I, I54V, L76V, V82A	R76K, Y79F, A431V	Q69K, S111C, I256V	31.08	0.0128	RS	20.00	0.0283	RS
	PCS033	M46I, F53L, I54V, L76V, V82A, L90M	R76K, K436R	Q69K, I256V	15.22	0.01068	RS	5.60	0.0011	RS
	PCS153	L24I, M46I, I54V, L76V, V82A	R76K, Y79F	Q69K, I256V	15.73	0.00346	RS	6.18	0.0023	RS
rGag	PCS011	None	L449P, R452K, P453L	S111C, I256V	0.39	0.00104	S	0.04	6E-05	S
	PCS020	None	R76K, Y79F, K436R, L449P	I256V	1.16	0.00072	S	5.26	0.2273	RS
	PCM029	None	R76K, P453L	None	23.88	0.02015	RS	6.73	0.0085	RS
	PCS049	None	R76K, Y79F, A431V	S111C	1.18	7.1E-05	S	0.99	0.0007	S
nGag	PCS022	None	None	Q69K	1.26	0.00443	S	0.39	0.0022	S
	PCS096	None	None	Q69K, S111C	1.26	0.00354	S	0.30	0.0001	S
	PCS128	None	None	Q69K, I256V	1.95	0.00243	S	0.71	0.0005	S
rGag + nGag	PCS115	None	R76K, L449P, P453L	Q69K, T239A/S, I256V	0.02	3.1E-05	S	0.08	0.0002	S
	PCS120	None	R76K, K436R	Q69K, S111C	7.40	0.00128	RS	0.78	0.0013	S
	PCS063	None	R76K, R452K	Q69K, I256V	1.52	0.00127	S	1.02	0.0006	S

[a] Represents the resistance classification/profile of a virus. Abbreviations: PR – Protease, rGag – resistance Gag; nGag – novel Gag; RAMs – resistance associated mutations; PID – patient ID, LPV – lopinavir; DRV – darunavir; FC – fold change, Res profile – resistance profile; RS – reduced susceptibility; S – Susceptible; STDEV – Standard deviation

3.4 Discussion

3.4.1 Replication capacity

Various studies have investigated the impact of Gag-Protease mutations on replication capacity and drug susceptibility in HIV-1 subtype B however no such study exists for HIV-1 subtype C.

In the current study, replication capacity was measured for 80 recombinant viruses generated using patient derived full-length Gag-Protease amplicons. Analysis of replication capacity data and associated genotypic data (Chapter 2) resulted in the selection of 18 viruses which were used in phenotypic drug susceptibility assays. Collectively these data was used to investigate the impact of Gag-Protease mutations on replication capacity and drug susceptibility in an HIV-1 subtype C cohort of participants failing a PI inclusive treatment regimen.

Results showed a direct association between variability in PR and reduced Gag-Protease replication capacity. This was supported by analysis which demonstrated that an increase in number of PR mutations was associated with reduced replication capacity. Moreover, viruses with PR mutations had a significantly lower replication capacity than viruses without PR RAMs. This result could be attributed to amino acid variations in PR that alter its conformational structure thereby reducing cleavage of the natural substrate (i.e. Gag), which manifests as reduced viral replication (40).

In contrast to PR, Gag mutations displayed no association with replication capacity. No difference in replication capacity was observed amongst viruses with rGag, nGag or no Gag mutations. Furthermore, no association between an increasing number of Gag mutations and replication capacity was observed. This suggests that cleavage of Gag can still occur despite variations in peptide sequences in cleavage and non-cleavage sites. Prabu-Jeyabalan and colleagues have demonstrated that PR recognizes the 3D structure of Gag rather than specific peptide sequences at CS (41). The variations in Gag noted in the current study may not have affected the 3D structure of the Gag CS to an extent where it becomes unrecognizable to PR and thus replication capacity may have remained unaltered. Bearing in mind that 42.5% of participants in this cohort had PR RAMs, it is also possible that variations in peptide sequences in Gag altered the CS 3D structure making it more recognizable to the mutant PR. Such a relationship has been previously demonstrated for the A431V rGag mutation and the V82A PR RAM (6, 42). However the study of Gag-Protease interactions on a structural level is suggested to elucidate the relationship between these two genes.

Interestingly, in the current study 89% of all viruses with PR RAMs harboured at least one rGag mutation in the NC/p1 (i.e. A431V) or p1/p6 (i.e. L449F, R452K or P453L) Gag CS (Chapter 2, Table 2.3). The A431V, L449F/P and P453L rGag mutations have all been previously identified to rescue viral replication capacity, which had been decreased by mutations in viral PR (9, 13, 15, 42, 43). Although these rGag mutations are associated with rescuing/restoring replication capacity, the level to which they rescue/restore it is variable. This has been demonstrated by previous studies, one of which showed that viral replication was increased by 16% when the Protease mutations I50V and M46I occurred in conjunction with L449F, however this virus still had a replication capacity that was less than 20% of NL43-WT (43). Similarly Zhang et al., (1997) reported an increase from 32% to 73% in viral replication (relative to NL43-WT) when M46L+V82A PR RAMs occurred in conjunction with a Gag mutation at the p1/p7 Gag cleavage site (13). These studies highlight that the level to which replication capacity is rescued/restored by Gag mutations varies drastically and is dependent upon mutation combinations (i.e. high or low).

It was thus not surprising to find that A431V was associated with significantly lower replication capacity in the current study, since 24/25 (96%) viruses with A431V also harboured at least three major PR RAMs which in itself would significantly impair Gag polypeptide processing. Interestingly we observed that the addition of the Q69K nGag mutation to viruses harbouring A431V and PR RAMs increased viral replication by 13%. This could suggest a potential compensatory role for the novel Q69K Gag mutation, which will be explored further by site-directed mutagenesis in Chapter 4. The potential compensatory role of Q69K also highlights that non-CS mutations in the amino terminal of Gag could also be important in PI resistance.

3.4.2 Drug susceptibility

As with replication capacity, the number of mutations in Gag and the variability of Gag sequences, showed no association with IC_{50} for LPV or DRV indicating that Gag mutations, in this cohort, may not contribute directly toward reducing drug susceptibility.

Variability in PR however, was significantly associated with increased IC_{50} for both LPV and DRV. This study demonstrated at least a 2.2 fold increase in LPV IC_{50} for every PR RAM found in the viruses analysed. This result was similar to a previous study which reported a 1.74 fold increase in LPV IC_{50} for every PR RAM added to a virus already harbouring three PR RAMs (44). The increase in IC_{50} associated with variability in PR could be attributed to a reduction in affinity of viral PR for PI's. This is caused by PR RAMs which alter the conformational structure

of the PR active site, to which PI's bind (7, 45-47). As a result the amount of PI required to inhibit 50% of viral replication would increase.

All viruses with PR RAMs, displayed reduced susceptibility to LPV and DRV, however the FC in IC_{50} for LPV was higher than that for DRV. This could be attributed to the higher genetic barrier of DRV in comparison to LPV or it could possibly be due to variations in mutations which confer resistance to LPV and DRV. The only PR mutation in the 18 viruses analysed in this study, known to impact DRV susceptibility was L76V. Other key DRV associated PR RAMs including I47V, I50V and I84V were absent from viruses analysed (48, 49). In contrast all PR RAMs present in the viruses analysed in this study have been associated with reduced susceptibility to LPV (49). As such it would be expected for LPV to have higher FC in IC_{50} than DRV for the viruses selected in this study.

Of the four viruses with rGag mutations only (i.e. no PR RAMs and no nGag mutations), two displayed reduced susceptibility to LPV and/or DRV. The first virus (PCM029) harboured the P453L Gag mutation whilst the second virus (PCS020) harboured the K436R and L449F Gag resistance associated mutations (Table 3.4). All three of these rGag mutations have been shown to be associated with reduced susceptibility to PI's in the absence of mutations in PR (17, 50-52).

Interestingly, all viruses harbouring the Q69K mutation (i.e. no PR RAMs or rGag mutations) were susceptible to LPV and DRV. This suggests that the Q69K mutation is not associated with reduced drug susceptibility. However further investigation using site-directed mutagenesis will follow in Chapter 4. Lastly, all viruses except one, which harboured rGag mutations and Q69K were susceptible to LPV and DRV. The one virus in this group with reduced susceptibility to LPV (i.e. PCS120) harboured the K436R rGag mutation, which has been previously associated with reduced susceptibility to PI's (17, 51, 52).

According to a previous study, the clinical cut-off for LPV and DRV was a FC in IC50 of ≥10 (54). A total of nine viruses in the current study displayed a FC in IC50 of ≥10 to LPV whilst two viruses displayed a FC in IC50 of ≥10 to DRV (Table 3.4). This could imply that these patients would be resistant to LPV or DRV respectively.

Overall 7 out of 10 viruses with rGag and nGag mutations only (i.e. no PR RAMs) remained fully susceptible to LPV and DRV. This would suggest that participants with these viruses experienced PI treatment failuer as a result of either poor adherence or an alternate path of PI resistance. With all 7 of these viruses belonging to participants with detectable LPV levels, it is

more likely that failure ccured as a result of an alternate pathway. Interestingly, Rabi et al., (2013) reported that mutations in Env could cause resistance to PI's in the absence of mtuations in PR. Future studies should therefore consider inclusion of the viral Env when investigating PI resistance (53).

No correlation was observed between drug susceptibility and replication capacity showing that replication capacity cannot be used as an indicator of treatment outcomes (Figure 3.17b).

3.5 Conclusion

Protease RAMs are responsible for reduced replication capacity in viruses harbouring PR RAMs and rGag mutations. These mtuations are also associated with significantly reducing drug susceptibility to LPV and DRV. Combinations of rGag mutations in the absence of PR RAMs can reduce drug susceptibility to LPV/DRV without altering replication capacity. The Q69K mutation is possibly a compensatory mutation in PI resistance, however its role in rescuing viral replication requires validation using site-directed mutagenesis.

3.6 References

1. Parry CM, Kohli A, Boinett CJ, Towers GJ, McCormick AL, Pillay D. 2009. Gag determinants of fitness and drug susceptibility in Protease inhibitor-resistant human immunodeficiency virus type 1. J Virol 83:9094-9101.
2. Parry CM, Kolli M, Myers RE, Cane PA, Schiffer C, Pillay D. 2011. Three residues in HIV-1 matrix contribute to Protease inhibitor susceptibility and replication capacity. Antimicrob Agents Chemother 55:1106-1113.
3. Sutherland KA, Mbisa JL, Cane PA, Pillay D, Parry CM. 2014. Contribution of Gag and Protease to variation in susceptibility to Protease inhibitors between different strains of subtype B human immunodeficiency virus type 1. J Gen Virol 95:190-200.
4. Sutherland KA, Parry CM, McCormick A, Kapaata A, Lyagoba F, Kaleebu P, Gilks CF, Goodall R, Spyer M, Kityo C, Pillay D, Gupta RK, Group DV. 2015. Evidence for Reduced Drug Susceptibility without Emergence of Major Protease Mutations following Protease Inhibitor Monotherapy Failure in the SARA Trial. PLoS One 10:e0137834.
5. Deeks SG, Wrin T, Liegler T, Hoh R, Hayden M, Barbour JD, Hellmann NS, Petropoulos CJ, McCune JM, Hellerstein MK, Grant RM. 2001. Virologic and immunologic consequences of discontinuing combination antiretroviral-drug therapy in HIV-infected patients with detectable viremia. N Engl J Med 344:472-480.
6. Clavel F, Mammano F. 2010. Role of Gag in HIV Resistance to Protease Inhibitors. Viruses 2:1411-1426.
7. Fun A, Wensing AM, Verheyen J, Nijhuis M. 2012. Human Immunodeficiency Virus Gag and Protease: partners in resistance. Retrovirology 9:63.
8. Gupta RK, Kohli A, McCormick AL, Towers GJ, Pillay D, Parry CM. 2010. Full-length HIV-1 Gag determines Protease inhibitor susceptibility within in vitro assays. AIDS 24:1651-1655.
9. Doyon L, Croteau G, Thibeault D, Poulin F, Pilote L, Lamarre D. 1996. Second locus involved in human immunodeficiency virus type 1 resistance to Protease inhibitors. J Virol 70:3763-3769.
10. Gatanaga H, Suzuki Y, Tsang H, Yoshimura K, Kavlick MF, Nagashima K, Gorelick RJ, Mardy S, Tang C, Summers MF, Mitsuya H. 2002. Amino acid substitutions in Gag protein at non-cleavage sites are indispensable for the development of a high multitude of HIV-1 resistance against Protease inhibitors. J Biol Chem 277:5952-5961.
11. Myint L, Matsuda M, Matsuda Z, Yokomaku Y, Chiba T, Okano A, Yamada K, Sugiura W. 2004. Gag non-cleavage site mutations contribute to full recovery of viral fitness in Protease inhibitor-resistant human immunodeficiency virus type 1. Antimicrob Agents Chemother 48:444-452.

12. Tamiya S, Mardy S, Kavlick MF, 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 78:12030-12040.
13. Zhang YM, Imamichi H, Imamichi T, Lane HC, Falloon J, Vasudevachari MB, Salzman NP. 1997. Drug resistance during indinavir therapy is caused by mutations in the Protease gene and in its Gag substrate cleavage sites. J Virol 71:6662-6670.
14. Ho SK, Coman RM, Bunger JC, Rose SL, O'Brien P, Munoz I, Dunn BM, Sleasman JW, Goodenow MM. 2008. Drug-associated changes in amino acid residues in Gag p2, p7(NC), and p6(Gag)/p6(Pol) in human immunodeficiency virus type 1 (HIV-1) display a dominant effect on replicative fitness and drug response. Virology 378:272-281.
15. Mammano F, Petit C, Clavel F. 1998. Resistance-associated loss of viral fitness in human immunodeficiency virus type 1: phenotypic analysis of Protease and Gag coevolution in Protease inhibitor-treated patients. J Virol 72:7632-7637.
16. Mammano F, Trouplin V, Zennou V, Clavel F. 2000. Retracing the evolutionary pathways of human immunodeficiency virus type 1 resistance to Protease inhibitors: virus fitness in the absence and in the presence of drug. J Virol 74:8524-8531.
17. Nijhuis M, van Maarseveen NM, Lastere S, Schipper P, Coakley E, Glass B, Rovenska M, de Jong D, Chappey C, Goedegebuure IW, Heilek-Snyder G, Dulude D, Cammack N, Brakier-Gingras L, Konvalinka J, Parkin N, Krausslich HG, Brun-Vezinet F, Boucher CA. 2007. A novel substrate-based HIV-1 Protease inhibitor drug resistance mechanism. PLOS Medicine 4:e36.
18. Robinson LH, Myers RE, Snowden BW, Tisdale M, Blair ED. 2000. HIV type 1 Protease cleavage site mutations and viral fitness: implications for drug susceptibility phenotyping assays. AIDS Res Hum Retroviruses 16:1149-1156.
19. Wright JK, Novitsky V, Brockman MA, Brumme ZL, Brumme CJ, Carlson JM, Heckerman D, Wang B, Losina E, Leshwedi M, van der Stok M, Maphumulo L, Mkhwanazi N, Chonco F, Goulder PJ, Essex M, Walker BD, Ndung'u T. 2011. Influence of Gag-Protease-mediated replication capacity on disease progression in individuals recently infected with HIV-1 subtype C. J Virol 85:3996-4006.
20. Jaclyn K. Wright VLN, Zabrina L. Brumme, Jessica L. Prince, Daniel T. Claiborne, Philip J.R. Goulder, Mark A. Brockman, Eric Hunter, Thumbi Ndung'u. 2012. Impact of HLA-B*81-Associated Mutations in HIV-1 Gag on Viral Replication Capacity. Journal of Virology 86:3193-3199.
21. Miura T, Brockman MA, Brumme ZL, Brumme CJ, Pereyra F, Trocha A, Block BL, Schneidewind A, Allen TM, Heckerman D, Walker BD. 2009. HLA-associated alterations

in replication capacity of chimeric NL4-3 viruses carrying Gag-Protease from elite controllers of human immunodeficiency virus type 1. J Virol 83:140-149.

22. Jaclyn W. 2011. Impact of immune-driven sequence variation in HIV-1 subtype C Gag-Protease on viral fitness and disease progression. PhD. University of KwaZulu-Natal, Durban.
23. Brockman MA, Tanzi GO, Walker BD, Allen TM. 2006. Use of a novel GFP reporter cell line to examine replication capacity of CXCR4- and CCR5-tropic HIV-1 by flow cytometry. J Virol Methods 131:134-142.
24. Meryman HT. 1974. Freezing injury and its prevention in living cells. Annu Rev Biophys Bioeng 3:341-363.
25. Agilent Technologies. 2013. StrataCooler cryo preservation module https://www.agilent.com/cs/library/usermanuals/Public/7008100.pdf. Accessed January 2014.
26. Wright JK, Brumme ZL, Carlson JM, Heckerman D, Kadie CM, Brumme CJ, Wang B, Losina E, Miura T, Chonco F, van der Stok M, Mncube Z, Bishop K, Goulder PJ, Walker BD, Brockman MA, Ndung'u T. 2010. Gag-Protease-mediated replication capacity in HIV-1 subtype C chronic infection: associations with HLA type and clinical parameters. J Virol 84:10820-10831.
27. Chu G, Hayakawa H, Berg P. 1987. Electroporation for the efficient transfection of mammalian cells with DNA. Nucleic Acids Res 15:1311-1326.
28. FlowJo. FlowJo in Africa. http://www.flowjo.com/home/africa/. Accessed June 2013.
29. Los Alamos National Laboratory. Highlighter. http://www.hiv.lanl.gov/content/sequence/HIGHLIGHT/highlighter.html. Accessed June 2014.
30. Rambaut A. 2015. Figtree. http://tree.bio.ed.ac.uk/software/figtree. Accessed 12 November 2014.
31. Puertas MC, Buzon MJ, Ballestero M, Van Den Eede P, Clotet B, Prado JG, Martinez-Picado J. 2012. Novel two-round phenotypic assay for Protease inhibitor susceptibility testing of recombinant and primary HIV-1 isolates. J Clin Microbiol 50:3909-3916.
32. Reed LA, Muench H. 1938. A simple method of estimating fifty percent endpoints. Am J Hyg 27:493-497.
33. Haidara A, Chamberland A, Sylla M, Aboubacrine SA, Cissé M, Traore HA, Maiga MY, Tounkara A, Nguyen VK, Tremblay C. 2012. Drug Resistance Pathways and Impact of Protease Mutation L10I/V in HIV-1 Non-B Subtypes. J Antivir Antiretrovir 4:043-050.
34. Codon by codon analysis. 2014. Codon by codon analysis tool. http://brockman-srv.mbb.sfu.ca/~B_Team_iMac/Codon_by_codon/. Accessed June 2014.

35. Mann JK, Byakwaga H, Kuang XT, Le AQ, Brumme CJ, Mwimanzi P, Omarjee S, Martin E, Lee GQ, Baraki B, Danroth R, McCloskey R, Muzoora C, Bangsberg DR, Hunt PW, Goulder PJ, Walker BD, Harrigan PR, Martin JN, Ndung'u T, Brockman MA, Brumme ZL. 2013. Ability of HIV-1 Nef to downregulate CD4 and HLA class I differs among viral subtypes. Retrovirology 10:100.
36. R Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.
37. Hsieh SM, Pan SC, Chang SY, Hung CC, Sheng WH, Chen MY, Chang SC. 2013. Differential impact of resistance-associated mutations to Protease inhibitors and nonnucleoside reverse transcriptase inhibitors on HIV-1 replication capacity. AIDS Res Hum Retroviruses 29:1117-1122.
38. Martinez-Picado J, Savara AV, Sutton L, D'Aquila RT. 1999. Replicative fitness of Protease inhibitor-resistant mutants of human immunodeficiency virus type 1. J Virol 73:3744-3752.
39. Barbour JD, Wrin T, Grant RM, Martin JN, Segal MR, Petropoulos CJ, Deeks SG. 2002. Evolution of phenotypic drug susceptibility and viral replication capacity during long-term virologic failure of Protease inhibitor therapy in human immunodeficiency virus-infected adults. J Virol 76:11104-11112.
40. Kozisek M, Henke S, Saskova KG, Jacobs GB, Schuch A, Buchholz B, Muller V, Krausslich HG, Rezacova P, Konvalinka J, Bodem J. 2012. Mutations in HIV-1 Gag and pol compensate for the loss of viral fitness caused by a highly mutated Protease. Antimicrob Agents Chemother 56:4320-4330.
41. Prabu-Jeyabalan M, Nalivaika E, Schiffer CA. 2002. Substrate shape determines specificity of recognition for HIV-1 Protease: analysis of crystal structures of six substrate complexes. Structure 10:369-381.
42. Prabu-Jeyabalan M, Nalivaika EA, King NM, Schiffer CA. 2004. Structural basis for coevolution of a human immunodeficiency virus type 1 nucleocapsid-p1 cleavage site with a V82A drug-resistant mutation in viral Protease. J Virol 78:12446-12454.
43. Maguire MF, Guinea R, Griffin P, Macmanus S, Elston RC, Wolfram J, Richards N, Hanlon MH, Porter DJ, Wrin T, Parkin N, Tisdale M, Furfine E, Petropoulos C, Snowden BW, Kleim JP. 2002. Changes in human immunodeficiency virus type 1 Gag at positions L449 and P453 are linked to I50V Protease mutants in vivo and cause reduction of sensitivity to amprenavir and improved viral fitness in vitro. J Virol 76:7398-7406.
44. Kempf DJ, Isaacson JD, King MS, Brun SC, Xu Y, Real K, Bernstein BM, Japour AJ, Sun E, Rode RA. 2001. Identification of genotypic changes in human immunodeficiency virus Protease that correlate with reduced susceptibility to the Protease inhibitor

lopinavir among viral isolates from Protease inhibitor-experienced patients. J Virol 75:7462-7469.

45. Prabu-Jeyabalan M, Nalivaika EA, Romano K, Schiffer CA. 2006. Mechanism of substrate recognition by drug-resistant human immunodeficiency virus type 1 Protease variants revealed by a novel structural intermediate. J Virol 80:3607-3616.
46. Kolli M, Lastere S, Schiffer CA. 2006. Co-evolution of nelfinavir-resistant HIV-1 Protease and the p1-p6 substrate. Virology 347:405-409.
47. Nalam MN, Ali A, Altman MD, Reddy GS, Chellappan S, Kairys V, Ozen A, Cao H, Gilson MK, Tidor B, Rana TM, Schiffer CA. 2010. Evaluating the substrate-envelope hypothesis: structural analysis of novel HIV-1 Protease inhibitors designed to be robust against drug resistance. J Virol 84:5368-5378.
48. Poveda E, de Mendoza C, Martin-Carbonero L, Corral A, Briz V, Gonzalez-Lahoz J, Soriano V. 2007. Prevalence of darunavir resistance mutations in HIV-1-infected patients failing other Protease inhibitors. J Antimicrob Chemother 60:885-888.
49. Wensing AM, Calvez V, Gunthard HF, Johnson VA, Paredes R, Pillay D, Shafer RW, Richman DD. 2014. 2014 Update of the drug resistance mutations in HIV-1. Top Antivir Med 22:642-650.
50. Giandhari J, Basson AE, Coovadia A, Kuhn L, Abrams EJ, Strehlau R, Morris L, Hunt GM. 2015. Genetic Changes in HIV-1 Gag-Protease Associated with Protease Inhibitor-Based Therapy Failure in Pediatric Patients. AIDS Res Hum Retroviruses 31:776-782.
51. Shibata J, Sugiura W, Ode H, Iwatani Y, Sato H, Tsang H, Matsuda M, Hasegawa N, Ren F, Tanaka H. 2011. Within-host co-evolution of Gag P453L and Protease D30N/N88D demonstrates virological advantage in a highly Protease inhibitor-exposed HIV-1 case. Antiviral Res 90:33-41.
52. Kolli M, Stawiski E, Chappey C, Schiffer CA. 2009. Human immunodeficiency virus type 1 Protease-correlated cleavage site mutations enhance inhibitor resistance. J Virol 83:11027-11042.
53. Rabi SA, Laird GM, Durand CM, Laskey S, Shan L, Bailey JR, Chioma S, Moore RD, Siliciano RF. 2013. Multi-step inhibition explains HIV-1 Protease inhibitor pharmacodynamics and resistance. J Clin Invest 123:3848-3860.
54. Lathouwers E, De Meyer S, Dierynck I, Van de Casteele T, Lavreys L, de Béthune P, Picchio G. 2011. Virological characterization of patients failing darunavir/ritonavir or lopinavir/ritonavir treatment in the ARTEMIS study: 96-week analysis. Antivir Ther 16: 99-108.

CHAPTER 4:

Impact of Q69K and A431V on replication capacity, drug susceptibility and cleavage

4 CHAPTER 4: IMPACT OF Q69K AND A431V ON REPLICATION CAPACITY, DRUG SUSCEPTIBILITY AND CLEAVAGE

4.1 Introduction

The production of mature infectious viral particles is dependent upon the ordered cleavage of the Gag polyprotein into its core structural proteins by the viral Protease (detailed in Chapter 1, section 1.6.9) (1-3). Protease inhibitors function to inhibit this cleavage thereby reducing virus production (4). However mutations in PR (e.g. M46I/L, I54V, L76V, V82A, D30N and G48V) have been shown to inhibit the effect of PI's (4-7). These mutations can cause reduced viral cleavage as a result of alterations to the conformational structure of the PR substrate binding cleft (4, 8). In order to compensate for this reduced cleavage, secondary/compensatory mutations occur in Gag and/or PR (9-11). The L10I/F and A71V mutations in PR are amongst the most well documented compensatory mutations in PR, whilst the A431V, L449F and P453L mutations in Gag are well known compensatory mutations in PI resistance (12-15). Interestingly, the A431V Gag mutation has also been identified as a primary PI resistance associated mutation, in that it can confer resistance to PI's in the absence of PR RAMs (16, 17).

Most compensatory and primary mutations in Gag have been identified through the use of SDM (16-19). Site-directed mutagenesis allows for the insertion of a mutation of interest into a double stranded DNA plasmid (20-22). The result is a mutant plasmid with the insert of interest which can be propagated as a recombinant virus and used in functional assays to assess the impact of a particular mutation on viral replication, drug susceptibility or cleavage.

The role of Q69K, a novel Gag mutation found to occur at significantly higher frequencies in participants failing a PI inclusive treatment regimen than in treatment naïve individuals in Chapter 2, and thereafter found to be associated with improving replication capacity when found in combination with A431V+PR RAMs in patient samples in Chapter 3, is investigated in the current chapter. Site-directed mutants containing variations of Q69K, A431V and the following PR RAMs: M46I, I54V, L76V and V82A were prepared and subject to replication capacity, drug susceptibility and polyprotein cleavage assays. The purpose of which was to determine if Q69K has a compensatory role in PI resistance, when found in conjunction with A431V and PR RAMs.

4.2 Methods

4.2.1 Generation of mutant viruses

The QuikChange Multi site-directed mutagenesis kit (Agilent) was employed in this study. It allows for the insertion of up to five point mutations into a plasmid. The process of SDM involves amplification of a double stranded DNA vector (i.e. plasmid) by PCR, using primers containing the desired mutations. The resultant product is a mutant plasmid comprising of several staggered nicks. Parental DNA is then removed by Dpn1 digestion, which enzymatically digests methylated DNA. In this instance plasmid DNA isolated from *E.coli* is methylated and is thus digested whilst DNA generated by PCR remains intact. Thereafter the nicked mutant plasmid is transformed into bacterial cells and plated onto LB-agar plates containing ampicillin. Following propagation, colonies can be processed and sequenced to determine if the mutant plasmid contains the mutations of interest. Furthermore stocks of the mutant plasmid can be prepared by culturing a colony in LB broth followed by a mini-prep or maxi-prep to purify the plasmid.

Since Q69K was found to increase the replication capacity of virsues harbouring A431V+PR RAMs, site directed mutatnts harbouring Q69K, A431V and/or PR RAMs (i.e .M46I, I54V, L76V and V82A) were preapred individually and in combinations (Table 4.1) Of the 20 mutants produced, one was generated by MSC student J. Giandhari (Mutant 1) (23) and eight were generated by MSC student K. Pillay (Mutant 2-9) (24). Eleven mutants were produced as part of the current study (Mutant 10-20).

Table 4-1 Overview of mutants generated for the current study. Mutants comprised of either Gag mutations only, Protease mutations only or combinations of Gag and Protease mutations.

Mutant number	Gag mutations	Protease mutations
Mutant 1	A431V	None
Mutant 2	A431V	V82A
Mutant 3	A431V	V82A + I54V
Mutant 4	A431V	V82A + I54V + M46I
Mutant 5	A431V	V82A + I54V + M46I + L76V
Mutant 6	None	V82A
Mutant 7	None	V82A + I54V
Mutant 8	None	V82A + I54V + M46I
Mutant 9	None	V82A + I54V + M46I + L76V
Mutant 10	Q69K	None
Mutant 11	Q69K	V82A
Mutant 12	Q69K	V82A + I54V

Mutant number	Gag mutations	Protease mutations
Mutant 13	Q69K	V82A + I54V + M46I
Mutant 14	Q69K	V82A + I54V + M46I + L76V
Mutant 15	Q69K + A431V	None
Mutant 16	Q69K + A431V	V82A
Mutant 17	Q69K + A431V	V82A + I54V
Mutant 18	Q69K + A431V	V82A + I54V + M46I
Mutant 19	Q69K + A431V	V82A + I54V + M46I + L76V
Mutant 20	Q69K + A431V	None

4.2.1.1 Primer design

Primers were designed to be complementary to the Gag-Protease gene of a patient designated SK254 (GenBank accession number: HM593258). The Gag-Protease sequence of this patient was found to most closely resemble that of the consensus subtype C Gag-Protease sequence generated from several treatment naïve patient-derived sequences obtained in our laboratory (HIV Pathogenesis Programme, University of KwaZulu-Natal).

Primers were designed using the web based QuikChange primer design program (www.genomics.agilent.com). All primers comprised of between 37-45 base pairs with the desired mutation flanked by at least 10 base pairs on either side. Each primer had a melting temperature of >75ºC, a minimum GC content of 40% and ended in a G or a C. Only primers for Q69K required to be generated for the current study as all other primers were already available (24).

The primer sequence for Q69K which corresponds to HXB2 position 976 – 1011, is given below. Nucleotides shown in red represent alterations made to achieve a substitution of a glutamine (Q) with a lysine (K) at position 69 of Gag:

Forward: 5' CAGCTACAACCAGCTCTT**A**AGACAGGAACAGAGGAAC 3'
Reverse: 5' GTTCCTCTGTTCCTGTCT**T**AAGAGCTGGTTGTAGCTG 3'

All primers were HPLC purified.

4.2.1.2 Plasmid vector preparation

The Gag-Protease region of patient SK254 was cloned into a TOPO vector plasmid (pCR2.1-TOPO) for use in SDM. Plasmid vectors were prepared as part of a previous study and donated for use in the present study (23). In total two plasmid vectors were used including: wildtype SK254 in TOPO and SK254 in TOPO containing the A431V Gag mutation.

4.2.1.3 Mutagenic PCR

The SK254 TOPO plasmid was mutated using the QuikChange Multi site-directed mutagenesis kit (Agilent) with relevant primers containing mutations of interest. Briefly, a PCR reaction for both mutant viruses and a positive control was prepared as per Table 4.2. Each PCR reaction was incubated in a thermocycler under the following conditions: 95°C for 2 minutes, 30 cycles of: 95°C for 20 seconds, 55°C for 30 seconds and 65°C for 2 minutes and 54 seconds followed by a 5 minute incubation at 65°C and a 2 minute hold at 37°C. Following completion of the PCR cycle, 1 µl of Dpn1[19] was added to each tube and incubated at 37ºC for 5–10 minutes in order to digest parental DNA.

Table 4-2 Summary of PCR reaction mix components for both sample and control mutagenic reactions, using the QuickChange lightning site-directed mutagenesis kit.

Component	Sample (µl)	Control (µl)	Final concentration
PCR water	16.75	18.5	-
10x Quickchange lightning reaction buffer	2.5	2.5	1x
Quiksolution	0.75	0	-
Fwd Primer	1	1	100 ng
Rvs Primer	1	1	100 ng
Control primer mix	0	1	Proprietary, no information in kit
dNTPs	1	1	Proprietary, no information in kit
Quickchange enzyme blend	1	1	1 U/µl
ds-DNA template (SK254 in TOPO)	1	0	100 ng/ µl
ds-control plasmid	0	1	Proprietary, no information in kit
Total	25	25	

[19] Dpn1 is a restriction endonuclease which functions to digest methylated and hemimethylated DNA. Plasmids used for SDM are generally isolated from *E. coli*. Bacterial cells contain methylated DNA, thus Dpn1 is able to digest the DNA.

4.2.1.4 Transformation

XL10-Gold ultracompetent cells (Agilent) were used for plasmid transformation as per manufacturer's instructions. A total of 45 µl of cells and 2 µl of beta-mercaptoethanol were incubated on ice for 10 minutes. Next, 1.5 µl of the digested mutant plasmid was added and incubated on ice for a further 30 minutes. The transformation reaction was then heat shocked at 42ºC for 30 seconds and promptly transferred onto ice for 2 minutes. Super optimal broth with catabolite repression (SOC)[20] medium (200 µl) was added to each reaction, followed by a 1 hour incubation at 37ºC and 250 rpm in a shaking incubator. Thereafter 100 µl of the transformation reaction was plated on LB-ampicillin agar plates which contained 40 µl of 5-bromo-4-chloro-3-indolyl-β-D-galctopyranoside (X-gal) and 10 µl of 10 mM isopropyl-1-thio-β-D-galctopyranoside (IPTG), which collectively functioned to enable easy visualization of successfully transformed colonies (i.e. white colonies indicated successful transformation whilst blue colonies were not successfully transformed). Plates were incubated at 37ºC for 16 hours.

4.2.1.5 Mutant screening

Following incubation, five single white colonies were picked and touched to a master-plate prior to being boiled at 90ºC in 10 µl of PCR water. The master-plate was incubated at 37ºC for 16 hours and stored thereafter at 4ºC. The DNA from the boiled colony was amplified by Gag-Protese PCR (described in Chapter 2, section 2.2.3) and sequenced (Chapter 2, section 2.2.4), to confirm the presence of the mutation.

Once the presence of the inserted mutation was confirmed the corresponding colony from the stored master-plate was picked and cultured in LB-broth at 37ºC for 16 hours. Purified, concentrated stocks of the mutant plasmid were then prepared using the Genejet plasmid mini-prep kit (Thermo Scientific, USA) as per manufacturer's instructions. DNA was quantified using the nanodrop and aliquots were stored at -80ºC.

4.2.1.6 Generation of mutant virus

Mutant viruses were generated as described in Chapter 3, section 3.2.1. Briefly, the mutated plasmid DNA (i.e. SK254 containing mutations of interest) was amplified by PCR, using 100-mer Gag-Protease primers (described in Chapter 2, section 2.2.3). GXR cells were then co-transfected with the amplified DNA and an NL43Δ*gag-protease* plasmid (as described in

[20] SOC media is a nutrient rich media used in the recovery of E. coli competent cell transformations.

Chapter 3, section 3.2.1.3). The co-transfection product was cultured in R10 media until 25-30% of GXR cells were infected, at which point the viruses were harvested and stored at -80ºC. Validation of all mutant viruses was conducted by extracting RNA from each harvested virus, followed by amplification and sequencing of the Gag-Protease region as described in Chapter 2, section 2.2.2 - 2.2.4.

4.2.1.7 Replication capacity and drug susceptibility assays

Harvested viruses were titred and used in replication capacity (performed in at least triplicate) and phenotypic drug susceptibility assays (performed in at least duplicate), as described in Chapter 3, section 3.2.2 and 3.2.3 respectively.

4.2.2 Western blot cleavage assay

The western blot assay was used to assess the cleavage of Gag by Protease, as previously described (25). The replication capacity assay was scaled up by six times in order to produce enough material for western blotting. Briefly, 6 million GXR cells in 600 µl of R10 was infected with mutant virus at an MOI of 0.3% (viruses were diluted in 2.4 ml of R10) in a 6 well plate, and incubated at 37ºC and 5% CO2 for 24 hours. After 24 hours, 6 ml of R10 was added to each well and incubated for a further 24 hours at 37ºC and 5% Co2. A total of 3 ml of culture was removed from each well from the second day and every day thereafter for the next 5 days, this was replaced with 3 ml of fresh R10. Of the 3 ml removed, 500 µl was used to measure infectivity by flow cytometry whilst 2.5 ml was prepared for use in the western blot assay.

4.2.2.1 Protein extraction and quantification

Protein extraction was performed by centrifuging the 2.5 ml of culture, collected at day 5 and 6 of the replication capacity assay, at 1,500 rpm and 4ºC for 10 minutes. The supernatant was then discarded. Pellets were rinsed in cold PBS and homogenised with a pipette, prior to centrifugation at 1,500 rpm and 4ºC for 10 minutes. Following centrifugation, the supernatant was removed and 150 µl of cold lysis solution containing a 1:200 dilution of cytobuster (Merck Millipore, Germany) and Protease inhibitor cocktail (Sigma) was added to each tube. The contents of each tube was thoroughly homogenised using a pipette and incubated on ice for 30 minutes with intermittent vortexing. Tubes were then centrifuged at 4,600 rpm for 30 minutes to pellet the debris.

The supernatant containing proteins was removed and quantified using the Bradford assay as per manufacturer's instructions (Bio-Rad). Briefly 10 µl of each protein lysate was added to a respective well of a 96 well plate. This was followed by the addition of 200 µl of Bradford reagent. The contents of each well was thoroughly mixed prior to the plate being read on a Biotek ELX 808 absorbance microplate reader (BioTek, Vermont, USA). Standard curves were generated using a 1 mg/ml BSA (Bio-Rad) standard diluted in PCR water. The standard curve was used to extrapolate protein concentrations for each sample.

A total of 10 µg of protein was added to an equal volume of loading buffer. Loading buffer contained 50 µl of beta-mercaptoethanol (Bio-Rad) and 950 µl of Laemmli buffer (Bio-Rad). This was mixed thoroughly by vortexing and incubated at 95ºC in a thermal cycler to breakdown primary, secondary and tertiary structures of proteins.

4.2.2.2 SDS-PAGE

A total of 15 µl of protein lysate (containing 10 µg of protein) solution mixed in loading buffer was added to each well of a Mini-Protean® TGX™ precast gel (Bio-Rad). A Precision Plus Protein™ WesternC™ molecular weight marker (5 µl) was added to each gel to enable protein identification. The gel was run at 150 V for 45 minutes in a Mini-Protean® tetra vertical electrophoresis cell (Bio-Rad) filled with running buffer containing 100 ml of Tris/Glycine/Sodium dodecyl sulphate (SDS) buffer (Bio-Rad) diluted in 900 ml of distilled water.

4.2.2.3 Protein transfer

Protein transfer was conducted using the Trans-Blot® Turbo™ Mini polyvinylide fluoride (PVDF) transfer packs (Bio-Rad) as per manufacturer's instructions. Briefly the SDS-PAGE gel was placed onto the first pre-assembled transfer stack. The PVDF membrane (activated in methanol and rinsed in distilled water) was placed over the gel. A second pre-assembled transfer stack was placed on top of the membrane. Air bubbles were removed and transfer conducted at 25 V for 10 minutes in a Trans-Blot® Turbo™ transfer system.

4.2.2.4 Blocking, probing, viewing and analysis

Following transfer the PVDF membrane was immediately placed in 50 ml of blocking agent containing 2.5 g milk powder (Bio-Rad) dissolved in 50 ml Tris Buffered Saline-(TBS)-Tween solution (prepared by adding 50 ml of 10 x TBS (Bio-Rad) to 450 ml of water and 250 µl Tween [Sigma]). This was incubated at room temperature for 1 hour on a shaking incubator. Following incubation the membrane was rinsed in 20 ml of TBS-Tween solution. The membrane was then probed with 5 µl of HIV p24 mouse monoclonal IgG Ab (Abcam, Massachusetts, USA) diluted in 20 ml of Signalboost primary antibody immunoreaction enhancer solution (Merck Millipore) at 4ºC overnight. Thereafter primary antibody was removed, the membrane was rinsed with TBS-Tween and probed with Rabbit polyclonal secondary Ab to mouse IgG with horse raddish peroxidase (HRP) (Abcam) diluted in Signalboost secondary antibody immunoreaction enhancer solution (Merck Millipore) for 1 hour at room temperature.

This was followed by removal of secondary antibody and rinsing of the membrane with TBS-Tween. Next, 20 ml of LumiGLO chemiluminescent substrate, prepared as per manufacturer's instructions, was added to the membrane and incubated at room temperature for 2 minutes prior to viewing on the Chemidoc^{TM} MP system (Bio-Rad). The Image Lab image acquisition and analysis software (Bio-Rad) was used to view bands and to calculate the density of bands. Band densitometry was used to calculate the p55/p24 ratio for samples representing day 5 and day 6 of the replication capacity assay. The ratio was used to compare processing of p55 between mutant viruses.

4.3 Results

4.3.1 Validation of mutant viruses

Of the 20 site-directed mutant viruses intended for use in this study only five were successfully propagated as recombinant viruses. The remaining 15 viruses were successfully inserted into plasmid vectors however these viruses never grew as recombinant viruses. Of these 15 viruses 14 had PR RAMs which could have had a high associated fitness cost thereby preventing their growth as recombinant viruses especially in the absence of compensatory mutations. The fifteenth mutant that did not grow harbored a combination of Q69K and A431V without PR RAMs. It is possible that Q69K+A431V require other Gag or PR mutations to facilitate its growth. This was supported by the successful growth of Q69K+A431V in the presence of the V82A PR RAM in this study.

Despite the unsuccessful propagation of 15 viruses, the five viruses generated (A431V, V82A, Q69K, A431V+V82A and A431V+V82A+Q69K) (Table 4.3) were sufficient to determine the impact of Q69K on replication capacity, polyprotein cleavage and drug susceptibility.

Sequencing of each mutant virus after harvesting confirmed that all viruses contained the mutations of interest.

Table 4-3 Summary of mutants successfully generated in the current study.

Mutant number	Gag **mu**tat**ions**	Protease **mu**tat**ions**
Mutant 1	Q69K	None
Mutant 7	Q69K + A431V	V82A
Mutant 12	A431V	None
Mutant 13	A431V	V82A
Mutant 17	None	V82A

4.3.2 Replication capacity

Replication capacity was compared amongst NL43-WT, SK254 and the following mutants: A431V, V82A, Q69K, A431+V82A and A431V+V82A+Q69K (Figure 4.1).

Results showed that NL43-WT had a significantly higher replication capacity than SK254 ($p<0.001$) and all mutant viruses ($p<0.0001$). Similarly SK254 had a significantly higher replication capacity than all mutant viruses ($p<0.0001$). Interestingly, Q69K had a significantly higher replication capacity than A431V, V82A, A431V+V82A and A431V+V82A+Q69K. This implies that the V82A and A431V mutation occurs at a fitness cost, and that even combining these mutations does not rescue replication capacity back to that of the wild-type virus (Figure 4.1).

As expected, V82A was associated with the lowest replication capacity possibly due to alterations it caused in the PR substrate binding cleft resulting in reduced cleavage of Gag. The A431V Gag mutation was associated with the second lowest replication capacity. Combining A431V with V82A improved viral replication significantly demonstrating that A431V has a compensatory role in replication capacity, as shown in previous studies (13, 16, 17, 26). Adding Q69K to the A431V+V82A combination further increased replication capacity (Figure 4.1). This increase was significant and suggests that Q69K may function as a compensatory mutation when found together with A431V+V82A. The ranking of in vitro viral fitness for mutants analysed

in this study was V82A < A431V< V82A+A4 31V < A431V+V82A+Q69K < Q69K <SK25 4 < NL43-WT.

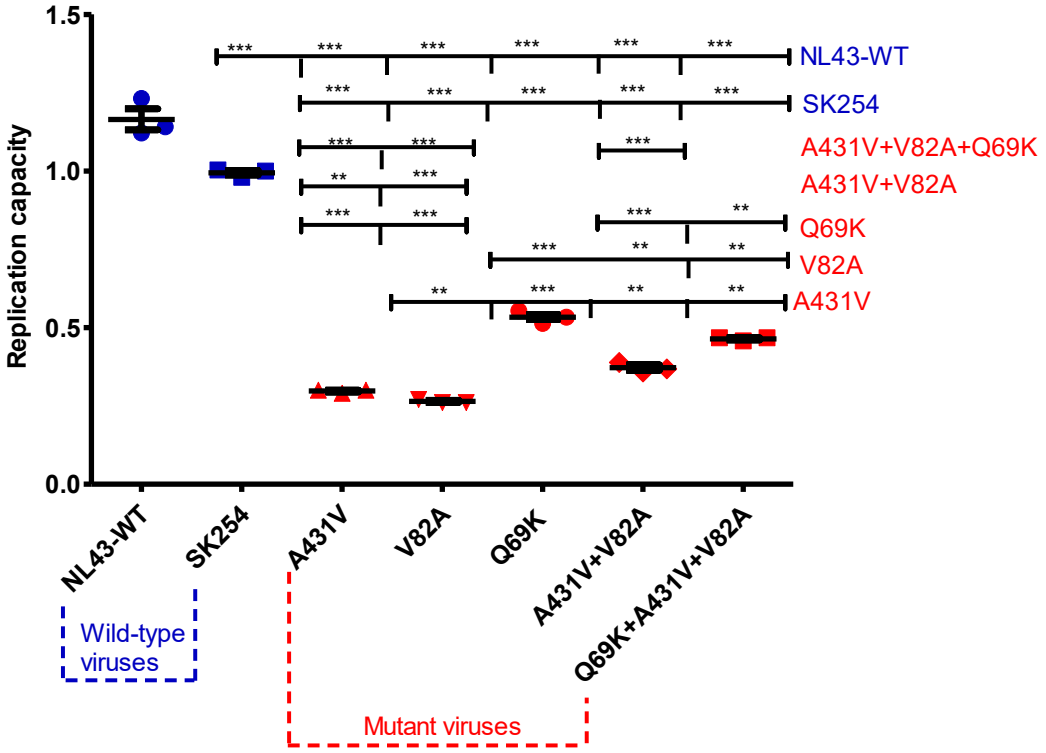


Figure 4-1 Comparison of replication capacity between wild type viruses (i.e. NL43-WT and SK254) and mutant viruses. Mutant viruses contained the following mutations: A431V, V82A, Q69K, A431V+V82A and A431V+V82A+Q69K. The V82A mutation was the only Protease mutation analysed, all other mutations represent Gag mutations. Significant differences are denoted by asterisks were * represents p<0.01, ** represents p<0.001 and *** represents p<0.0001. Blue writing represents wild-type viruses whilst red writing represents mutant viruses. The right hand Y axis is labelled with virus names which correspond to significant differences denoted by stars above each bar. For example, NL43-WT had significantly higher replication than SK254 and all mutant viruses whilst Q69K had significantly higher replication capacity than A431V, V82A, A431V+V82A and A431+V82A+Q69K.

4.3.3 Cleavage assay

The western blot assay was used to assess variations in cleavage amongst mutant viruses. Culture from day five and six of a replication capacity assay was processed and used in the western blot assay. Two blots were run. The first included samples for day five and six of the following viruses: Q69K, Q69K+A431V+V82A, A431V and NL43-WT (Figure 4.2a). The second blot comprised of day five and six samples for the following viruses: V82A, A431V+V82A, SK254 and NL43-WT (Figure 4.2b). Band densitometry analysis was performed using the Image Lab Software, in order to calculate the p55 (Gag polyprotein)/p24 (processed polyprotein) ratio

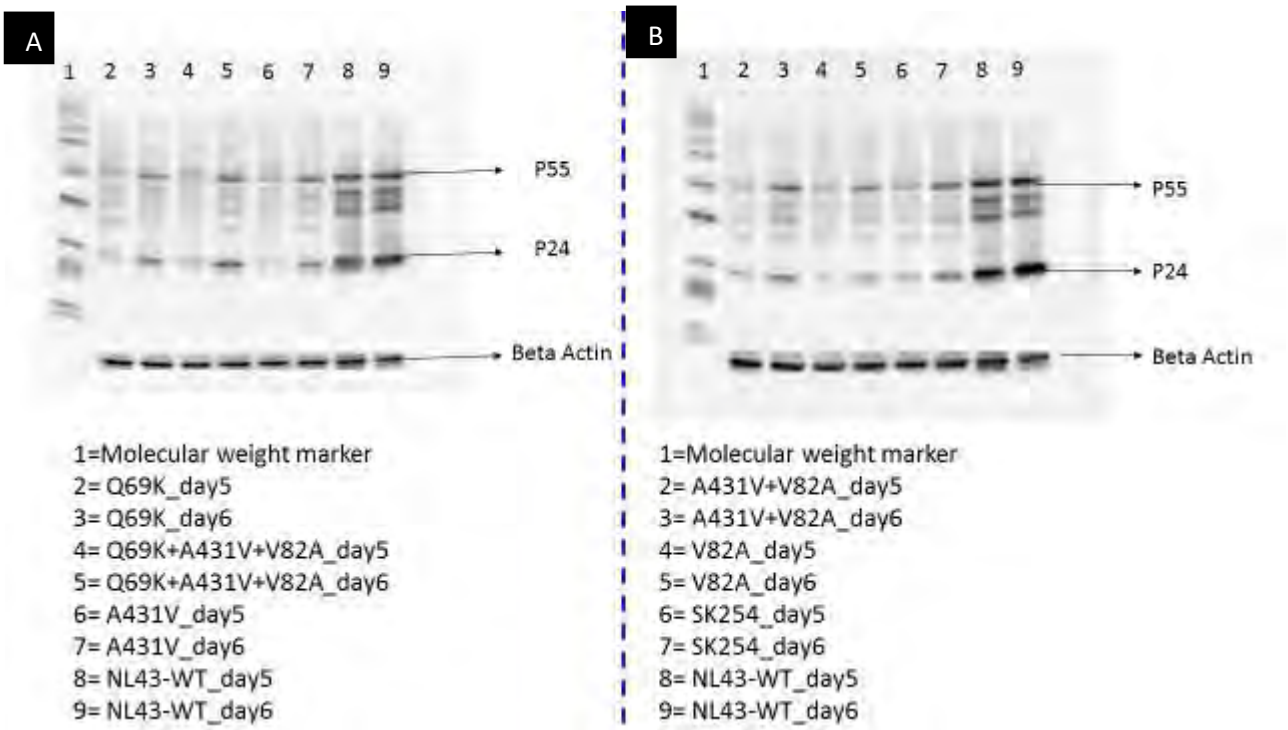
for each virus (Figure 4.2c). In general, a lower ratio represents more efficient p55 processing and vice versa.

Beta actin bands, representing the loading control, were similar for both blots, indicating that approximately equal amounts of protein was loaded for each sample (Figure 4.2a and b).

SK254 (Lane 6,7, blot 2) and NL43-WT (Lane 8 and 9, blot 2) had the most prominent p55 and p24 bands in comparison to all other viruses, indicating that these two wild-type viruses had the highest virus production and cleavage (Figure 4.2b). Band densitometry analysis showed that the p55/p24 ratios for NL43-WT was lower than that for SK254, indicating more advanced polyprotein processing for NL43-WT than SK254. This suggests that NL43-WT had higher replication and cleavage than SK254, which is in line with replication capacity assay results (Figure 4.1).

The mutant virus harbouring Q69K had a lower p55/p24 ratio than all other mutant viruses, (i.e. it had better polyprotein processing than other mutant viruses) (Figure 4.1c). The mutant virus harbouring V82A had the highest p55/p24 ratio in comparison to all mutant viruses. This ratio however was reduced when A431V occurred in conjunction with V82A and was further reduced when Q69K occurred in conjunction with both A431V and V82A (Figure 4.2c). These results support replication capacity data. Whilst A431V is known to act as a compensatory mutation when occurring in conjunction with PR RAMs, this is the first report suggesting that Q69K may have a role in improving polyprotein cleavage in the presence of V82A and A431V.

The ranking of polyprotein clevage for mutants analysed in this study was the same as the ranking for in vitro replicationcapacity V82A < A431V< V82A+A4 31V < A431V+V82A+Q69K < Q69K <SK25 4 < NL43-WT.



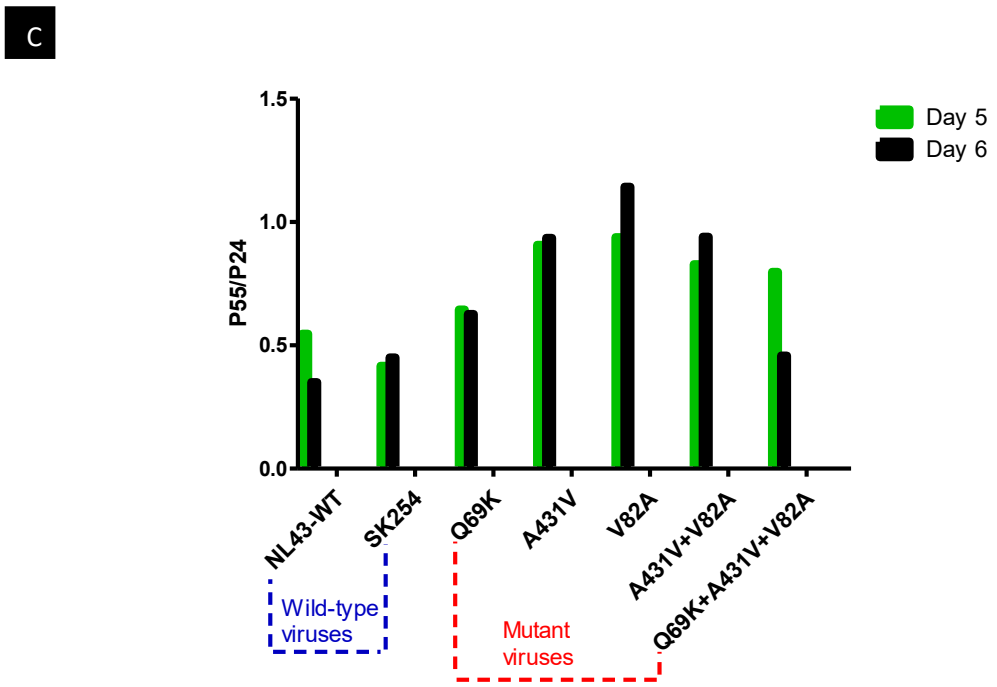


Figure 4-2 Western blot analyses for wild-type viruses (NL43-WT and SK254) and 5 mutant viruses. Mutant viruses harbouring the following mutations were used: A431V, V82A, Q69K, A431V+V82A and A431V+V82A+Q69K. (a) Western blot showing p55, p24 and beta actin bands for day 5 and 6 cell lysate samples representative of the NL43-WT virus and the following mutant viruses: Q69K, V82A+A431V+Q69K and A431V. (b) Western blot showing p55, p24 and beta actin bands for day 5 and 6 cell lysate samples representative of: NL43 and SK254 and the following mutant viruses: A431V+V82A and V82A alone. (c) p55/p24 ratios for all mutant and wild-type virus samples blotted on day 5 and 6 of a replication capacity assay.

4.3.4 Phenotypic drug susceptibility assay

All mutant viruses harbouring the Gag A431V and/or Protease V82A mutation, showed reduced susceptibility to Lopinavir (i.e. FC between 2.43 and 10). This was in line with previous studies which reported that viruses with V82A and A431V occurring either individually or in combination, confers reduced susceptibility to LPV (13, 16, 17, 19). The A431V mutant conferred a higer reduction in susceptibility to LPV (FC: 5.89) than the V82A mutant (FC: 3.79). This supports a previous study that showed that A431V could confer a level of resistance similar to that of a single PR RAM (27).

Interestingly, mutant viruses harbouring the V82A or A431V mutation also showed reduced susceptibility to DRV (FC: 4.43 and 4.63 respectively). This was surprising, since a previous study reported that A431V confers reduced susceptibility to all PI's except DRV (14, 17). Similarly the Stanford HIVdb and the IAS-USA list of mutations do not list V82A as a Protease mutation associated with DRV resistance (6, 28).

In comparison to the mutant harbouring V82A alone, mutants harbouring a combination of A431V+V82A further increased IC50 of LPV by 2.67 fold suggesting that Gag mutations could increase drug susceptibility scores when considered together with PR RAMs. This however did not apply to DRV showing that the effect of mutations on drug susceptibility differs between drugs.

The Q69K Gag mutation was not associated with reduced susceptibility to LPV or DRV when found in isolation of other mutations.

Table 4-4 Summary of drug susceptibility data for each mutant virus (i.e. V82A, A431V, Q69K, V82A+A431V, and A431V+V82A+Q69K).

PID	FC LPV[a]	STDEV[b]	Res Profile[c]	FC DRV[d]	STDEV	Res Profile
V82A	3.79	0.65	RS	4.43	0.09	RS
A431V	5.89	0.004	RS	4.63	0.09	RS
Q69	2.41	0.01	S	1.65	0.18	S
V82A+A431V	6.46	0.11	RS	4.71	0.06	RS
A431V+V82A+Q69K	4.63	0.02	RS	0.05	0.01	S

[a] Represents the fold change in susceptibility of each virus to lopinavir (LPV) using NL43-WT as a reference. [b] Represents the standard deviation between duplicate fold change values for each virus. [c] Represents the resistance classification/profile of each virus described in Chapter 3, Table 3.2. [d] Represents the fold change in susceptibility of each virus to Darunavir (DRV) using NL43-WT as a reference. Abbreviations: PID – patient ID, LPV – lopinavir; DRV – darunavir; FC – fold change, Res profile – resistance profile; R – resistant, RS – reduced susceptibility; S – Susceptible; STDEV – standard deviation

4.4 Discussion

The role of Gag mutations in PI resistance is well documented, particularly for mutations in the NC/p1 and p1/p6 Gag cleavage sites (9, 29-31). There is however limited information on the role of Gag mutations in the matrix and capsid regions in PI resistance. Here we investigated the impact of a novel Gag mutation in the matrix region (i.e. Q69K), occurring individually and in combinations with the Gag A431V and PR V82A mutations, on viral replication, proteolytic cleavage and drug susceptibility. The Q69K mutation was identified to be associated with PI resistance/exposure in Chapter 2 and shown to significantly increase viral fitness when occurring in conjunction with A431V and PR RAMs in patient-derived recombinant viruses in Chapter 3. Results from the current chapter showed that the Q69K mutation is associated with improving replication capacity and viral cleavage when found in conjunction with the V82A PR RAM and the A431V Gag NC/p1 cleavage site mutation.

The V82A PR RAM was the only PR mutation that was successfully propagated both individually and in combination with both the A431V and Q69K Gag mutations. This could be explained by findings of a previous study which showed that the valine at codon 82 is not essential for substrate recognition and thus substrate cleavage could still be permitted (32). Although the mutant PR (containing the V82A mutation) can still recognise the substrate, its ability to cleave the substrate is slightly impaired, since the larger valine is replaced with a smaller alanine at positon 82 causing conformational changes in the structure of PR which have

been shown to reduce contact between viral PR and the NC/p1 region of Gag (13, 26). This conformational change also reduces the binding affinity of the PR active site for PI's thereby conferring resistance to various PI's (26). The NC/p1 Gag cleavage site has been shown to coevolve with V82A, both in the presence and absence of PI's, in order to improve viral cleavage (13, 16, 17). In this instance, the replacement of a smaller alanine at position 431 of Gag with a larger valine creates a protrusion of the NC/p1 peptide into a region within the substrate binding domain of PR which is not usually occupied by the substrate (13, 26) This, allows for improved contact between the cleavage site and the mutant PR, which enables improved viral cleavage and virus production (26). This could explain the significantly higher replication capacity and lower p55/p24 ratios of mutant viruses harbouring V82A in conjunction with A431V in comparison to mutant viruses harbouring the V82A mutation alone (Figure 4.1 and Figure 4.2). It also explains the reduced susceptibility to LPV and DRV for viruses harbouring V82A.

Similarly, the reduced susceptibility seen with A431V alone could be attributed to the continued processing of the Gag polyprotein even in the presence of a PI bound PR, which is possibly facilitated by the protruding NC/p1 region into the substrate binding domain of PR (13, 33, 34).

An interesting result in this study was the identification of the Q69K Gag matrix mutation as a potential compensatory mutation. It was associated with significantly increased replication capacity and improved viral cleavage when found in conjunction with A431V+V82A. This significant increase was only seen at day 6 of the replication capacity assay, possibly due to the slow replicating nature of HIV-1 subtype C compared to HIV-1 subtype B. The Q69K mutation however, was not found to be associated with reduced drug susceptibility. These data combined with data from Chapter 2, showing that Q69K only occurred in PI exposed participants, indicates that Q69K is associated with PI resistance and possibly occurs as a compensatory mutation. Whilst its mechanism of action for replication compensation is unclear, the Q69K mutation could play a role in altering alpha helical structures in MA which would improve accessibility of the MA/CA CS to mutated PR. Additionally such alterations could improve the affinity of PR for the mutated Gag substrate. Collectively, both these actions would improve viral cleavage which manifests as improved viral replication. Such a mechanism has been described for the R76K, Y79F and T81A MA non-CS mutations (35). This mechanism however requires further investigation for Q69K.

4.5 Conclusion

This study confirmed that A431V functions as both a compensatory and primary resistance mutation in PI resistance. In line with other studies we also showed that V82A can occur in isolation of other PR RAMs and that PR can still conduct its enzymatic action on the Gag substrate despite the presence of V82A, however the rate of cleavage is reduced. Lastly this study validated that the Q69K matrix mutation has a compensatory role in rescuing replication capacity and viral cleavage when found in conjunction with A431V+V82A. The mechanism of action of Q69K however requires further investigation.

4.6 References

1. Yang H, Nkeze J, Zhao RY. 2012. Effects of HIV-1 Protease on cellular functions and their potential applications in antiretroviral therapy. Cell Biosci 2:32.
2. Tozser J, Yin FH, Cheng YS, Bagossi P, Weber IT, Harrison RW, Oroszlan S. 1997. Activity of tethered human immunodeficiency virus 1 Protease containing mutations in the flap region of one subunit. Eur J Biochem 244:235-241.
3. Lv Z, Chu Y, Wang Y. 2015. HIV Protease inhibitors: a review of molecular selectivity and toxicity. HIV AIDS (Auckl) 7:95-104.
4. Iyidogan P, Anderson KS. 2014. Current perspectives on HIV-1 antiretroviral drug resistance. Viruses 6:4095-4139.
5. Johnson VA, Calvez V, Gunthard HF, Paredes R, Pillay D, Shafer R, Wensing AM, Richman DD. 2011. 2011 update of the drug resistance mutations in HIV-1. Top Antivir Med 19:156-164.
6. Wensing AM, Calvez V, Gunthard HF, Johnson VA, Paredes R, Pillay D, Shafer RW, Richman DD. 2014. 2014 Update of the drug resistance mutations in HIV-1. Top Antivir Med 22:642-650.
7. Gulnik SV, Suvorov LI, Liu B, Yu B, Anderson B, Mitsuya H, Erickson JW. 1995. Kinetic characterization and cross-resistance patterns of HIV-1 Protease mutants selected under drug pressure. Biochemistry 34:9282-9287.
8. Mahalingam B, Louis JM, Reed CC, Adomat JM, Krouse J, Wang YF, Harrison RW, Weber IT. 1999. Structural and kinetic analysis of drug resistant mutants of HIV-1 Protease. Eur J Biochem 263:238-245.
9. Zhang YM, Imamichi H, Imamichi T, Lane HC, Falloon J, Vasudevachari MB, Salzman NP. 1997. Drug resistance during indinavir therapy is caused by mutations in the Protease gene and in its Gag substrate cleavage sites. J Virol 71:6662-6670.
10. Pettit SC, Henderson GJ, Schiffer CA, Swanstrom R. 2002. Replacement of the P1 amino acid of human immunodeficiency virus type 1 Gag processing sites can inhibit or enhance the rate of cleavage by the viral Protease. J Virol 76:10226-10233.
11. Robinson LH, Myers RE, Snowden BW, Tisdale M, Blair ED. 2000. HIV type 1 Protease cleavage site mutations and viral fitness: implications for drug susceptibility phenotyping assays. AIDS Res Hum Retroviruses 16:1149-1156.
12. Menendez-Arias L. 2013. Molecular basis of human immunodeficiency virus type 1 drug resistance: Overview and recent developments. Antiviral Res 98:93-120.

13. Clavel F, Mammano F. 2010. Role of Gag in HIV Resistance to Protease Inhibitors. Viruses 2:1411-1426.
14. Kolli M, Stawiski E, Chappey C, Schiffer CA. 2009. Human immunodeficiency virus type 1 Protease-correlated cleavage site mutations enhance inhibitor resistance. J Virol 83:11027-11042.
15. Doyon L, Croteau G, Thibeault D, Poulin F, Pilote L, Lamarre D. 1996. Second locus involved in human immunodeficiency virus type 1 resistance to Protease inhibitors. J Virol 70:3763-3769.
16. Dam E, Quercia R, Glass B, Descamps D, Launay O, Duval X, Krausslich HG, Hance AJ, Clavel F, Group AS. 2009. Gag mutations strongly contribute to HIV-1 resistance to Protease inhibitors in highly drug-experienced patients besides compensating for fitness loss. PLoS Pathog 5:e1000345.
17. Nijhuis M, van Maarseveen NM, Lastere S, Schipper P, Coakley E, Glass B, Rovenska M, de Jong D, Chappey C, Goedegebuure IW, Heilek-Snyder G, Dulude D, Cammack N, Brakier-Gingras L, Konvalinka J, Parkin N, Krausslich HG, Brun-Vezinet F, Boucher CA. 2007. A novel substrate-based HIV-1 Protease inhibitor drug resistance mechanism. PLOS Medicine 4:e36.
18. Hsieh SM, Pan SC, Chang SY, Hung CC, Sheng WH, Chen MY, Chang SC. 2013. Differential impact of resistance-associated mutations to Protease inhibitors and nonnucleoside reverse transcriptase inhibitors on HIV-1 replication capacity. AIDS Res Hum Retroviruses 29:1117-1122.
19. Fun A, Wensing AM, Verheyen J, Nijhuis M. 2012. Human Immunodeficiency Virus Gag and Protease: partners in resistance. Retrovirology 9:63.
20. Carter P. 1986. Site-directed mutagenesis. Biochem J 237:1-7.
21. Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 77:51-59.
22. Carrigan PE, Ballar P, Tuzmen S. 2011. Site-directed mutagenesis. Methods Mol Biol 700:107-124.
23. Giandhari J. 2010. The role of the Protease cleavage sites in viral fitness and drug resistance in HIV-1 subtype C. MSC. University of KwaZulu-Natal, Durban.
24. Pillay K. 2015. The impact of the p7/p1 cleavage site mutations on replication capacity and drug resistance in HIV-1 subtype C. Masters of Science. University o Kwa-Zulu Natal, South Africa.
25. Prado JG, Honeyborne I, Brierley I, Puertas MC, Martinez-Picado J, Goulder PJ. 2009. Functional consequences of human immunodeficiency virus escape from an HLA-B*13-restricted CD8+ T-cell epitope in p1 Gag protein. J Virol 83:1018-1025.

26. Prabu-Jeyabalan M, Nalivaika EA, King NM, Schiffer CA. 2004. Structural basis for coevolution of a human immunodeficiency virus type 1 nucleocapsid-p1 cleavage site with a V82A drug-resistant mutation in viral Protease. J Virol 78:12446-12454.
27. Nijhuis M, Wensing AM, Bierman WF, de Jong D, Kagan R, Fun A, Jaspers CA, Schurink KA, van Agtmael MA, Boucher CA. 2009. Failure of treatment with first-line lopinavir boosted with ritonavir can be explained by novel resistance pathways with Protease mutation 76V. J Infect Dis 200:698-709.
28. Stanford University HIV Drug Resistance Database. 9 March 2015 2015. Drug resistance summaries. http://hivdb.stanford.edu/pages/download/resistanceMutations_handout.pdf. Accessed 25 September 2015.
29. Maguire MF, Guinea R, Griffin P, Macmanus S, Elston RC, Wolfram J, Richards N, Hanlon MH, Porter DJ, Wrin T, Parkin N, Tisdale M, Furfine E, Petropoulos C, Snowden BW, Kleim JP. 2002. Changes in human immunodeficiency virus type 1 Gag at positions L449 and P453 are linked to I50V Protease mutants in vivo and cause reduction of sensitivity to amprenavir and improved viral fitness in vitro. J Virol 76:7398-7406.
30. Prado JG, Wrin T, Beauchaine J, Ruiz L, Petropoulos CJ, Frost SD, Clotet B, D'Aquila RT, Martinez-Picado J. 2002. Amprenavir-resistant HIV-1 exhibits lopinavir cross-resistance and reduced replication capacity. AIDS 16:1009-1017.
31. de la Carriere LC, Paulous S, Clavel F, Mammano F. 1999. Effects of human immunodeficiency virus type 1 resistance to Protease inhibitors on reverse transcriptase processing, activity, and drug sensitivity. J Virol 73:3455-3459.
32. Prabu-Jeyabalan M, Nalivaika EA, King NM, Schiffer CA. 2003. Viability of a drug-resistant human immunodeficiency virus type 1 Protease variant: structural insights for better antiviral therapy. J Virol 77:1306-1315.
33. Ali A, Bandaranayake RM, Cai Y, King NM, Kolli M, Mittal S, Murzycki JF, Nalam MN, Nalivaika EA, Ozen A, Prabu-Jeyabalan MM, Thayer K, Schiffer CA. 2010. Molecular Basis for Drug Resistance in HIV-1 Protease. Viruses 2:2509-2535.
34. Parry CM, Kohli A, Boinett CJ, Towers GJ, McCormick AL, Pillay D. 2009. Gag determinants of fitness and drug susceptibility in Protease inhibitor-resistant human immunodeficiency virus type 1. J Virol 83:9094-9101.
35. Parry CM, Kolli M, Myers RE, Cane PA, Schiffer C, Pillay D. 2011. Three residues in HIV-1 matrix contribute to Protease inhibitor susceptibility and replication capacity. Antimicrob Agents Chemother 55:1106-1113.

CHAPTER 5

Identification of low frequency transmitted drug resistance mutations in HIV-1 subtype C acute infection

5 CHAPTER 5: TRANSMITTED DRUG RESISTANCE IN HIV-1 SUBTYPE C ACUTE INFECTION

5.1 Enclosed article: Low frequency drug resistance mutations are common in HIV-1 subtype C acute infection

This chapter comprises of a journal article which has been submitted for review to AIDS. It details the prevalence of TDR with a HIV-1 subtype C acute infection cohort from KwaZulu-Natal, Durban, South Africa. Results for Sanger sequencing and UDPS are presented. Transmitted drug resistance mutations present at both high and low frequency is discussed.

Low frequency drug resistance mutations are common in HIV-1 subtype C acute infection

Urisha SINGH[1], Avashna SINGH[1], Marc NOGUERA-JULIAN[2], Manjeetha JAGGERNATH[1], Amber MOODLEY[1], Tarylee REDDY[3], Krista DONG[4], Bruce D. WALKER[1,4,5], Thumbi NDUNG'U[1,4,6,7] and Michelle L. GORDON[21#]

*HIV Pathogenesis Programme, Doris Duke Medical Research Institute, University of KwaZulu-Natal[1]; Institut de Recerca de la SIDA IrsiCaixa i Uni*tat *VIH, Universi*tat *Autònoma de Barcelona, Universi*tat *de Vic, Catalonia, Spain[2]; Biostatistics Unit, Medical Research Council, Durban, South Africa[3]; Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts, USA[4]; Howard Hughes Medical Institute, Chevy Chase, Maryland[5]; KwaZulu-Natal Research Institute for Tuberculosis and HIV (K-RITH), University of KwaZulu-Natal[6]; Max Planck Institute for Infection Biology, Chariteplatz, D-10117 Berlin, Germany[7]*

This work was supported by the Bill and Melinda Gates Foundation (OPP1066973) and the Collaboration for AIDS Vaccine Discovery, the Mark and Lisa Schwartz Foundation, the International AIDS Vaccine Initiative (IAVI) (UKZNRSA1001), the NIAID (R37AI067073), and the Howard Hughes Medical Institute. Additional funding was received from the South African Research Chairs Initiative, the Victor Daitz Foundation and the Canada-Sub-Saharan Africa (CANSSA) HIV/AIDS Network.

Abstract: 243 words

Text body: 3490 words

Corresponding Author: Michelle L. Gordon. Mailing address: HIV Pathogenesis Programme, University of KwaZulu-Natal, 719 Umbilo Road, Durban, 4013, South Africa. Phone: 27 31 2604998. Fax: 27 31 2604623. Email: Tarinm@ukzn.ac.za.

Abstract

Objective: Widespread roll-out of combination antiretroviral therapy (cART) has improved the quality of life of HIV-1-infected individuals in South Africa but the extent of transmission and persistence of drug resistance mutations (DRMs) is largely unknown. We identified DRMs in individuals with acute HIV-1 subtype C infection in Durban, South Africa, and analysed the persistence of low frequency DRMs and explored their impact on treatment outcomes.

Design and Methods: Sanger sequencing was performed on 45 samples; 32 obtained at a median of 14 days post infection and 13 obtained at a median of one day following onset of plasma viremia (DFOPV). Ultra-deep pyrosequencing (UDPS) was performed on a subset of 14 samples, obtained 1 DFOPV, to identify low frequency DRMs.

Results: Sanger sequencing revealed that one of 45 participants (2%) harbored the K103N non-nucleotide reverse transcriptase inhibitor (NNRTI)-associated DRM. UDPS detected low frequency DRMs in 8 of 14 participants (57%) including: the K65R (1-1.5%) and D67N (3.88%) nucleotide reverse transcriptase inhibitor (NRTI)-associated DRMs, the F53L (17.6%) and M46L (6.3%) Protease inhibitor-associated DRMs, and the T97A (2.90%) integrase strand transfer inhibitor-associated DRM. The K103N DRM persisted for over a year. All low frequency DRMs were transient.

Conclusions: We showed a high prevalence of transmitted or spontaneous emergence of low frequency DRMs in acute infection which do not persist but may affect time to viral suppression for participants on cART. Follow-up of participants on cART is necessary to establish long-term effects of these mutations on treatment.

Keywords
Low frequency drug resistance mutations, HIV-1 subtype C, acute infection

Introduction

Transmitted Drug Resistance (TDR) in HIV-1 presents a risk for the future success and longevity of cART, especially in resource-limited settings in which no genotypic testing is conducted prior to treatment initiation, and options for treatment and salvage therapy are limited. Most TDR studies utilize dideoxynucleotide (Sanger) sequencing for detection of DRMs. However, this method is limited in its ability to detect mutations at frequencies below 15 to 20% [1, 2]. Although the clinical significance of low frequency (i.e. minority variant) mutations remains controversial, studies show that DRMs present at frequencies as low as 1% can significantly impact clinical outcomes of patients on cART [3-11]. Several studies demonstrated that NNRTI-associated DRMs present at <20% of the viral quasispecies were strongly associated with virologic failure of patients initiating an NNRTI-inclusive treatment regimen [1, 3, 6, 7, 12, 13]. This has been attributed to the rapid re-emergence of DRMs under drug selection pressure, and their subsequent dominance in the viral population [14-16].

Whilst low frequency DRMs have been studied in chronically infected treatment naïve individuals [6, 7, 12, 17-19] and recently infected individuals [3, 12, 20] limited information exists for such DRMs in acute infection. The rapid reversion of some DRMs, due to associated high fitness costs [21-30], renders them undetectable during chronic and recent infection. The study of DRMs during acute infection could offer insight into these rapidly reverting low frequency DRMs that can potentially re-emerge from reservoirs under drug selection pressure and impact treatment outcomes [21-32].

One of the few studies using an acute cohort to identify low frequency DRMs, employed a point-mutation assay to detect specific mutations of interest (*viz.* K103N, M184V and L90M) in an HIV-1 subtype B acutely infected cohort [33]. Point-mutation assays are limited in that they can only identify a restricted number of predetermined mutations in a single run [34, 35]. In contrast, deep-sequencing techniques are able to identify all mutations present within a specified genomic region, in a single instrument run [34, 36].

To our knowledge, there are currently no data available on the prevalence of low frequency DRMs in HIV-1 subtype C acute infection. HIV-1 subtype-specific differences in response to cART and development of drug resistance have been reported [37], and therefore information on low frequency DRMs in acute subtype C infection could be beneficial in understanding the landscape of viral diversity in subtype C. Such information could be used to inform decisions on future treatment strategies and clinical management of patients infected with HIV-1 subtype C, the most prevalent subtype globally [38].

Here we performed Sanger sequencing to identify TDR associated mutations in the reverse transcriptase (RT) and Protease (PR) genes from 47 participants acutely infected with HIV-1 subtype C. Thereafter, selected samples were investigated for the presence of low frequency DRMs in RT, PR and integrase (INT) using the Roche 454 UDPS platform. We compared DRMs identified by Sanger sequencing and UDPS and determined which mutations persisted over time. Additionally we performed exploratory assessments of the impact of low frequency DRMs on treatment outcomes.

Methods

Study participants

Samples were available from 47 individuals with acute HIV-1 subtype C infection, from Durban, South Africa, and PCR amplification was successful for 45 samples. Of the 47 individuals, 32 were from the HIV Pathogenesis Programme's (HPP) Acute Infection (AI) cohort [39] and 15 participants were from the Females Rising through Education, Support and Health (FRESH) programme [40, 41]. As previously described, acute infection was defined as: a positive HIV-1 RNA test, a negative or indeterminate rapid immuno-assay with subsequent confirmation of seroconversion by the western blot method [39].

HPP AI study participants were identified by RNA screening of individuals who tested seronegative when presenting for routine HIV counselling and testing at outpatient clinics in the greater Durban area. The date of infection for participants from this cohort was estimated to be 14 days prior to the first HIV-1 positive RNA, as previously described [42]. In the FRESH programme, initiated in November 2012, 18-23 year old HIV uninfected sexually active women were offered a comprehensive empowerment curriculum designed to mitigate HIV infection risk and sampled twice weekly for HIV-1 RNA by finger prick blood draw. Blood samples were collected from participants with evidence of plasma viremia, usually within 24 hours of onset of plasma viremia (OPV) and at regular intervals thereafter [41].

This study was approved by the Biomedical Research Ethics Committee at the University of KwaZulu-Natal. All study participants provided written informed consent before inclusion into the study.

HIV-1 drug resistance testing by Sanger sequencing

Viral RNA was extracted from cryopreserved plasma samples, using the Qiamp Viral RNA Mini-Kit, as per manufacturer's instructions (Qiagen, Valencia, USA). The PR and RT coding regions (HXB2 coordinates 2,293 to 3,509) were amplified and sequenced using the Viroseq HIV-1 genotyping system (Applied Biosystems, Foster City, CA). Sequences were run on the ABI 3130XL genetic analyser (Applied Biosystems), edited in Sequencher version 5.1 and aligned using ClustalX version 2.1. The REGA HIV-1 subtyping tool (http://www.bioafrica.net/rega-genotype/html/subtypinghiv.html) was used to confirm the subtype of all sequences. The International AIDS Society-USA (IAS-USA) list updates of DRMs [43], the World Health Organization (WHO) list of mutations for surveillance of TDR [44] and the Stanford HIV Drug Resistance Database (http://hivdb.stanford.edu/) were used to identify DRMs.

HIV-1 drug resistance testing by UDPS resistance plate assay

For this study, 15 samples from the FRESH cohort were analysed using UDPS to detect low frequency DRMs. Samples selected for UDPS were from recent transmissions (2013 – 2014) in comparison to the HPP AI cohort samples (2007 – 2012). Additionally, samples from the FRESH cohort represented earlier samples post-HIV transmission compared to the HPP AI cohort since participants in the former were tested more frequently.

Ultra-deep pyrosequencing was performed according to instructions provided in the four Plate HIV-1 Drug Resistance Assay Manual – Collaborative Initiative version 3.0 (Roche, Life Sciences). Briefly, viral RNA was extracted using the Qiamp Viral RNA Mini Kit (Qiagen, Valencia) and purified using Agencourt RNAclean XP magnetic beads (Beckman Coulter, Beverly, MA). The cDNA was generated using cDNA synthesis primer plates (Roche, Life Sciences) and a Transcriptor First Strand cDNA Synthesis Kit (Roche, Applied Science). PCR was performed using PCR primer plates pre-spotted with five PCR primer pairs (Roche, Life Sciences), and a FastStart High Fidelity PCR System (Roche, Applied Science). The first four PCR primers overlapped to cover the PR and RT regions of *pol* (HXB2 coordinates: 2,279 to 3,302) and the fifth primer covered the INT region (HXB2 coordinates: 4,352 to 4,739). Primers included a virus-specific sequence for PCR priming attached to a sample specific multiplex identifier sequence (MID) of approximately 10 bp and a 454 sequencing adaptor which allowed for either forward or reverse sequencing (Roche, Life Sciences) [34, 45].

Amplicons were purified using Agencourt Ampure XP magnetic beads (Beckman Coulter, Beverly, MA) and quantified by the Quant-iT PicoGreen dsDNA assay kit (Invitrogen, Carlsbad,

CA). Samples with concentrations below 5 ng/µl were run on the Agilent 2100 bioanlayzer (Agilent Technologies, Waldbronn, Germany). Samples with primer-dimer molar ratios above 3:1 were excluded from further processing.

Equimolar concentrations of all amplicons were pooled and 5 µl of the master-pool was clonally amplified by emulsion-PCR (emPCR) using a GS Junior emPCR Lib-A kit (Roche, Life Sciences). Between 100,000 to 500,000 DNA beads were prepared and loaded onto a picotitre plate (Roche-454, Life Sciences) and run on the GS Junior 454 sequencer, as per manufacturer's instructions (Roche-454 Life Sciences).

The GS Amplicon Variant Analyser software (AVA) (Roche-454 Life Sciences) was used to analyse and compute UDPS results as described elsewhere [1, 34, 46]. Briefly, the software recognizes MIDs and assigns relevant sequence reads to the corresponding amplicon and patient sample. Additionally it aligns each sequence read to a wild-type subtype C reference sequence (GenBank accession number AY772699), trims sequences and quantifies the frequency in variations of nucleotides at each position in the sample sequence relative to the reference sequence [1, 34]. The AVA software also corrects carry forward and incomplete extension errors [46]. Verification of mutations was conducted by manually inspecting flowgrams at positions of interest. Additionally, the frequency of mutations in both the forward and reverse strands were compared, using a two-tailed Fisher's exact test. If variants were detected disproportionately (i.e. $p < 0.001$) in one direction they were not considered as low frequency DRMs [47].

Identification of low frequency DRMs by UDPS

Low frequency DRMs were defined as DRMs detected at frequencies between 1% and 20%, whilst high frequency DRMs were defined as DRMs detected at frequencies above 20%. The cut-off frequency to distinguish low frequency mutations from artifacts caused by technical error was ≥1%. This was consistent with several studies which have shown mutations at a frequency of 1% to be clinically relevant [1, 3, 6, 7, 12, 13, 48]. Validation of UDPS to detect minority variants at frequencies above 1% has been described previously by Vanderbroucke et al., (2011), with concordance and correlation of >0.97 demonstrated for inter and intra run reproducibility [49, 50]. The IAS-USA list updates of DRM's and [43] the WHO list of mutations for surveillance of TDR [44] was used to identify TDR mutations in PR, RT and INT.

Quality control

Negative and positive controls were included in all steps. Sequences generated using Sanger sequencing and UDPS were subjected to analysis for inter-sample and lab strain contamination using BLAST homology searches against each sample and against sequences available in the Los Alamos database (http://blast.ncbi.nlm.nih.gov/Blast.cgi). To visually assess sequences for cross-contamination, neighbor joining trees were drawn in Paup 4.0 and visualized in Figtree v1.1 [51]. A plasmid clone obtained from the Quality Control for Molecular Diagnostics (QCMD) 2013 ENVA HIV Drug Resistance EQA Programme (www.qcmd.org), with known DRMs, was included in one UDPS run.

Exploratory investigation of response to treatment

Seven FRESH participants were initiated on fixed dose cART containing, Tenofovir, Efavirenz and Emtricitabine (2 participants with low frequency DRMs and 5 participants without low frequency DRMs. Six of these participants initiated treatment as per South African 2014 guidelines (i.e.CD4 < 350 cells/mm^3) [52], whilst one participant was placed on treatment a day after diagnosis. Response to treatment (i.e. viral load and CD4$^+$ T-lymphocyte dynamics) was assessed using the Student's T-test. Statistical significance was defined by $P<0.05$.

RESULTS

Participant characteristics

Baseline characteristics of all study participants are summarized in Table 5.1. For the HPP AI cohort, the median viral load was 5.76 $\log_{10}$ copies/ml (IQR, 5.02 – 6.67 $\log_{10}$ copies/ml) and the median CD4$^+$ T-lymphocyte count was 322 cells/mm^3 (IQR, 289 – 519 cells/mm^3). For the FRESH cohort, the median viral load was 4.89 $\log_{10}$ HIV-1 RNA copies/ml (IQR, 4.31 to 5.58 $\log_{10}$ HIV-1 RNA copies per ml) and the median CD4$^+$ T-lymphocyte count was 637 cells/mm^3 (IQR, 390 to 727 cells/mm^3).

Sanger sequencing

A total of 45 out of 47 samples were successfully sequenced by Sanger sequencing with no cross-contamination observed (Figure 5.1a). Of the 45 samples sequenced, only 1 sample (2%) from participant 079 presented with a major TDR-associated DRM (*viz*. the K103N NNRTI-associated DRM). Additional samples from this participant were bulk sequenced at the following

time-points; two weeks after OPV, six months after OPV and one year after OPV. The K103N mutation persisted at all time-points sequenced. These data indicate that there is limited transmission of DRMs in this high prevalence setting.

Detection of low frequency DRMs by UDPS during hyperacute HIV-1 infection

Fourteen of 15 FRESH samples were sequenced with sufficient reads for data analysis. For these 14 samples, UDPS generated a median of 8,281 high quality reads per sample (range, 7,377 to 11,091 reads per sample) with a median read length of 381 bases (range, 377 to 382 bases). This resulted in a mean coverage of 1,783 reads per amplicon (range, 1,475 to 2,218 reads per amplicon).

Two of three UDPS runs were used to assess reproducibility. A concordance of >97% was demonstrated for inter-run reproducibility. All mutant variants were detected proportionately. All known DRMs in the QCMD plasmid clone were detected, including the M41L (85.2%), M184V (99.57%), L210W (99.35%) and T215Y (99.24%) NRTI-associated DRMs. Phylogenetic analysis showed no cross contamination (Figure 5.1b).

Major/ Primary low frequency DRMs were identified in 8 of 14 participants (57%), at the earliest time-point sequenced (Table 5.2). Of the 8 participants, four (268, 272, 312 and 318) presented with the K65R NRTI-associated DRM only at 1.51%, 1.04%, 1.19% and 1.11% respectively. Participant 267 had the T97A (2.78%) integrase strand transfer inhibitor (InSTI)-associated DRM and the D67N (3.92%) NRTI-associated DRM whilst participant 271 had the F53L (17.6%) Protease inhibitor (PI)-associated DRM. Participant 079 had K103N (100%), consistent with Sanger sequencing, as well as the V90I (7.77%) NNRTI-associated DRM and the K65R (1.08%) NRTI-associated DRM. Participant 036 presented with the M46I/L (7%) PI-associated DRM (Table 2). These data demonstrate a high prevalence of low frequency DRMs in this cohort.

Analysis of low frequency DRMs at additional time-points using UDPS

A subset of participant samples with low frequency DRMs (267, 268, 271, 079 and 036), were sequenced by UDPS at additional time-points where available, including: peak viremia, presumed viral set-point, mid-time-point and/or the last available time-point (Figure 5.2).

For participant 267, the T97A InSTI-associated DRM and D67N NRTI-associated DRM were detected at frequencies of 2.90% and 3.88% respectively, 2 days following OPV. Neither of these DRMs was detectable 19 days later (i.e. at peak viremia).

Participant 268 presented with the K65R NRTI-associated DRM (1.51%) one day following OPV. This DRM was undetectable by UDPS at the last time-point sequenced 164 days later.

The F53L PI-associated DRM was detected one day following OPV in participant 271. This mutation was undetectable seven days later, which coincided with peak viremia. No additional DRMs were detected at day 15 (i.e. presumed viral set-point) or day 249 (i.e. last time-point).

Participant 079 harbored the K103N (99.9%) and V90I (7.86%) NNRTI-associated DRMs as well as the K65R (1.08%) NRTI-associated DRM detected three days following OPV. The K103N mutation persisted at a high frequency (>99%). The participant also developed the K219R (8.17%) NRTI-associated DRM at the last sequenced time-point, 329 days after OPV.

Participant 036 initially presented with the M46L (6.32%) PI-associated DRM, three days following OPV. This mutation was undetectable seven days later (peak viremia), at which point the K65R (1.16%) NRTI-associated DRM was detected. Again, the K65R mutation reverted seven days later (presumed viral set-point), with no additional DRMs detected by day 22. Interestingly, the Y143C (1.19%) InSTI-associated DRM and the K219Q (1.64%) NRTI-associated DRM were detected at the last tested time-point, approximately 332 days after OPV.

Overall, these data show a high prevalence of transmission or spontaneous emergence of low frequency DRMs during acute HIV-1 subtype C infection. However, the majority of these DRMs do not persist within the host and are mostly transient.

Response to treatment

We next performed an exploratory assessment of whether the presence of low frequency DRMs compromised cART effectiveness in patients initiating therapy. Two participants that harbored low frequency DRMs (267 and 036) achieved undetectable viral loads by day 156 and 120 respectively (median: 138 days) (Figure 5.2a and 5.2e). Three of five participants with no low frequency DRMs (093, 208 and 309) had undetectable viral load by day 174, 70 and 98 respectively (median: 98 days), whilst one participant (312) showed complete viral suppression by day 27 (Figure 5.3). This participant however received treatment one day after diagnosis and was therefore not compared to participants that only received treatment once their $CD4^+$ count dropped below 350 cells/mm^3. The fifth participant in this group never attained complete viral suppression (186), possibly due to non-adherence, and developed the K103N DRM 311 days after treatment initiation, indicative of poor adherence (Figure 5.3d).

Participants without low frequency DRMs had a significantly higher viral load at treatment initiation than participants with low frequency DRMs (p=0.0452), however they reached viral suppression 40 days sooner than participants with low frequency DRMs (i.e. median of 98 versus 138 days). This however did not reach statistical significance in this small cohort. There were no significant differences in: number of days to treatment initiation, $CD4^+$ T-lymphocyte counts at treatment initiation and $CD4^+$ T-lymphocyte counts at viral suppression between the two groups. This data indicates that low frequency DRMs possibly impacts treatment outcomes.

DISCUSSION

The extent of TDR in resource-limited, high prevalence settings is largely unknown and yet has important public health implications. Here, we show a low prevalence of TDR mutations detected by Sanger sequencing (2%) in persons diagnosed with acute infection in KwaZulu-Natal Province, South Africa, one of the highest incidence regions in the world, where treatment has been available in the public sector since 2004. However, UDPS revealed a high prevalence of low frequency TDR mutations (57%). Although these low frequency DRMs were transient, participants harboring them had lower viral loads than participants without DRMs, at treatment initiation, yet took longer to achieve complete viral suppression, suggesting that low frequency DRMs impacts treatment outcomes. While the small sample size is recognized as a limitation, these findings support previous studies [3-11].

Variations in viral load and mutation persistence is likely attributable to fitness costs (measured as variation in replicative ability) associated with DRMs [21-27]. Generally mutant viruses have lower replicative fitness than wild-type virus [28]. In the absence of drug selection pressure, less fit viruses revert their mutations to improve replicative fitness [29, 30]. Over time, the more fit strain predominates and the mutant virus decreases in frequency [3, 28]. Here we show that K103N persisted at one year after OPV. This is in line with several studies which showed that K103N strains replicate at similar rates to wild-type viruses and persist at high frequencies for up to five years [53-56].

Consistent with another study, we found that PI-associated DRMs revert rapidly (seven days) (Figure 1) [25]. This is likely due to high fitness costs associated with mutations in PR, which alter the structure of the PR binding pocket and subsequently impairs proteolytic cleavage of the natural substrate, resulting in reduced viral replication [57, 58]. To our knowledge, this is the first study in South Africa (SA) to identify transmitted PI-associated DRMs. This is possibly driven by an increased number of patients receiving PI-inclusive treatment regimens. However PI-associated TDR could be under-reported in SA due to the combined use of Sanger

sequencing and recently infected cohorts (generally defined as being infected for 3 to 6 months) for TDR surveillance, both of which would hinder the identification of low frequency rapidly reverting PI-associated DRMs. [25, 59].

The K65R NRTI-associated DRM was the most commonly occurring low frequency DRM (42.8%). These were all found at <2% of the viral population. Studies show that K65R is most prevalent in subtype C [60-63], due to sequence variations and RNA structure which favors its sporadic appearance [64]. Even though K65R can occur “sporadically”, this does not exclude the possibility of TDR [65]. Recent studies showed that transmitted low frequency DRMs are more likely to have clinical implications in comparison to “sporadically” appearing low frequency DRMs, largely attributable to faster disappearance of “sporadically” occurring DRMs [65]. Whilst unclear if K65R occurred sporadically or as a transmitted DRM, we show that it reverted seven days after first identification in participant 036 who achieved complete viral suppression 120 days after treatment initiation with a Tenofovir inclusive regimen (Figure 2e). Further follow-up of participant 036 is suggested to confirm if low frequency DRMs impact treatment outcomes at a later stage.

Although the use of InSTIs is limited in SA, our study detected two participants with InSTI-associated DRMs: T97A in participant 267 and Y143C in participant 036, both of which are associated with Raltegravir resistance (Figure 2a and 2e). Whilst T97A could be a transmitted mutation, studies have shown that it occurs as a natural polymorphism in varying frequencies across subtypes [66, 67]. The Y143C DRM that occurred one year after OPV was possibly a result of error prone de novo replication and not TDR, as the absence of a second peak in viral load eliminates the possibility of dual infection or superinfection as its cause. The presence of InSTI-associated DRMs in this cohort could indicate that these mutations occur sporadically in subtype C. This could have implications on the efficacy of InSTI’s for treatment of HIV-1 subtype C.

With studies showing that viral reservoirs are established earlier than 10 days after onset of clinical symptoms in primary HIV infection [31, 32], the probability of low frequency mutations identified in this study to exist in reservoirs are high. We recommend investigating viral diversity in reservoirs during acute infection to elucidate archived viral populations. Additionally, further work to determine the impact of low frequency PI and InSTI-associated mutations on clinical outcomes is suggested.

CONCLUSION

Low frequency DRMs are common in HIV-1 subtype C acute infection. While most revert rapidly, they may still be present in viral reservoirs and could impact future therapy options. Low frequency DRMs may contribute to increased time to viral suppression. However two participants with low frequency DRMs showed viral suppression within 6 months. The long term outcomes of treatment have yet to be determined.

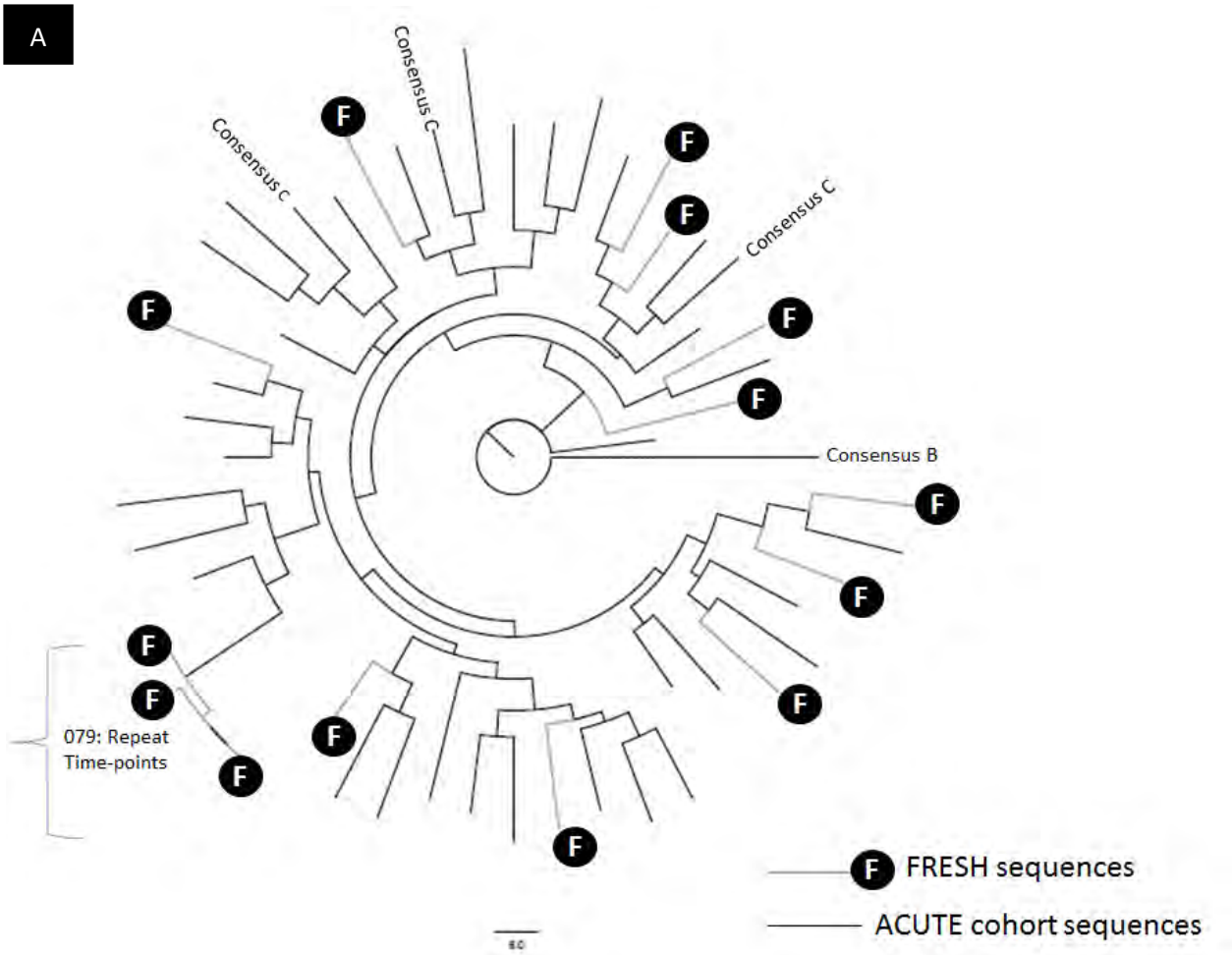
ACKNOWLEDGEMENTS:

We would like to thank the FRESH and HPP AI participants, the FRESH and HIV Pathogenesis Programme (HPP) staff, Dr Luke Swenson and Dr Zaza Ndhlovu.

Author contributions: U.S.: manuscript writing, ultra-deep pyrosequencing, Sanger sequencing, data collation, phylogenetic analysis, sequence analysis, AVA bioinformatics and statistical analysis. A.S.: Sanger sequencing and ultra-deep pyrosequencing. N.N-J.: Bioinformatics. M.J.: Site clinician. A.M.: Clinical site manager and data validation. T.R.: Statistical analysis. K.D.: Cohort design. B.D.W.: Study design, cohort design, manuscript editing. T.N.: Cohort design, study design, manuscript writing. M.L.G.: manuscript writing and study design.

Figure 5-1 Neighbor joining phylogenetic trees of sequences from participants belonging to the FRESH cohort and acute infection cohort. (a) Sanger sequences from FRESH cohort participants are represented by grey branches with a black circle containing an F at the end of each branch, whilst Sanger sequences from acute cohort sequences are represented by black branches with no label. Sequences at various time-points from participant 079 clustered together. This highlights the similarity of samples. No clustering was observed between all other sequences sequenced by Sanger sequencing. (b) Sanger sequences (grey branches) and ultra-deep sequences (black branches) for all FRESH participants is depicted. Sequences ending with a U represents those generated by UDPS whilst those ending with an SS represents sequences generated by Sanger sequencing. The first three digits of the sample ID represents the patient ID. For selected participants, four additional digits follow the patient ID, this denotes the time-point sequenced. No unusual clustering between different PIDs were noticed indicative of no cross contamination. As expected samples at different time-points clustered with each other. Furthermore, Sanger sequences and sequences generated by UDPS, from the same participant, clustered together.

A


Consensus C
Consensus C
Consensus C
Consensus B
079: Repeat
Time-points
F FRESH sequences
ACUTE cohort sequences

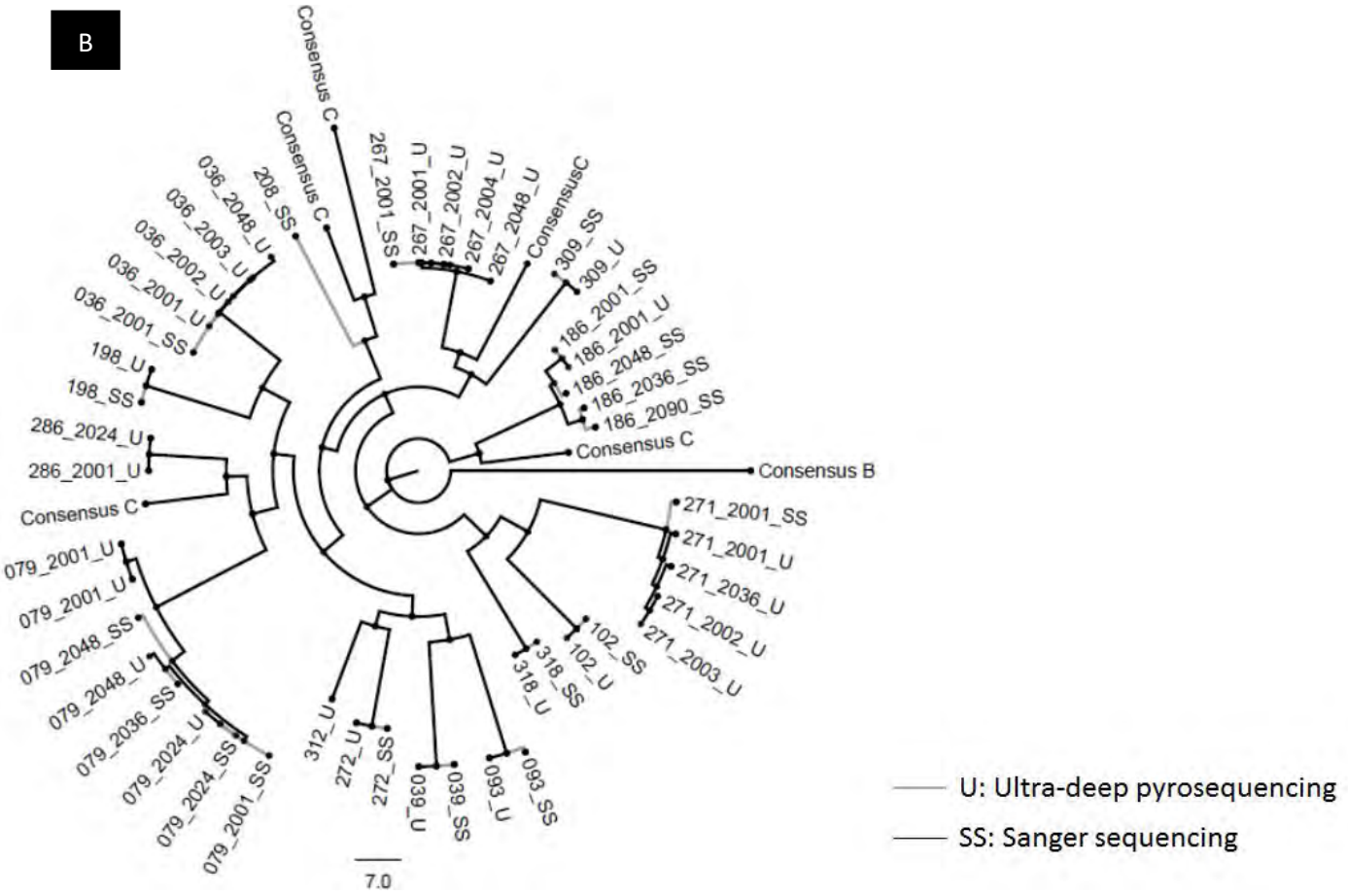
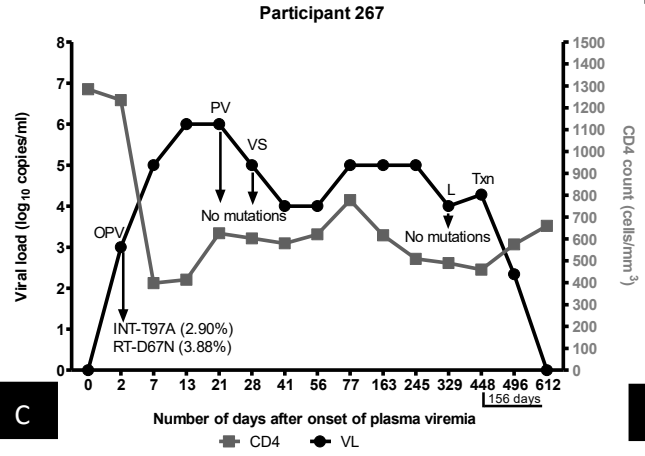
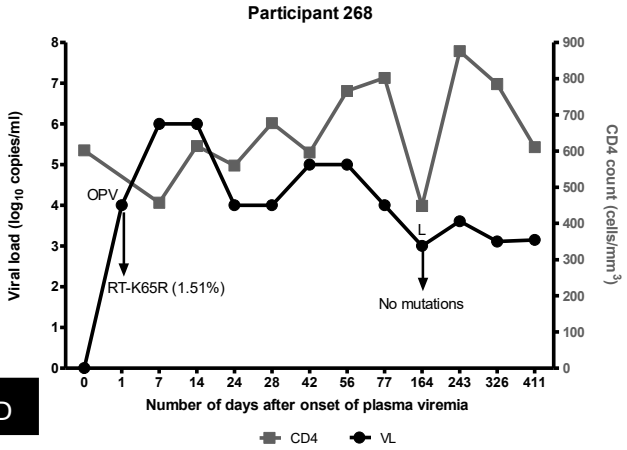

B
Consensus C
Consensus C
267_2001_SS
267_2001_U
267_2002_U
267_2004_U
267_2048_U
ConsensusC
309_SS
309_U
186_2001_SS
186_2001_U
186_2048_SS
186_2036_SS
186_2090_SS
Consensus C
Consensus B
271_2001_SS
271_2001_U
271_2036_U
271_2002_U
271_2003_U
102_SS
102_U
318_SS
318_U
093_SS
093_U
039_SS
039_U
272_SS
272_U
312_U
079_2001_SS
079_2024_SS
079_2024_U
079_2036_SS
079_2048_U
079_2048_SS
079_2001_U
079_2001_U
Consensus C
286_2001_U
286_2024_U
198_SS
198_U
036_2001_SS
036_2001_U
036_2002_U
036_2003_U
036_2048_U
208_SS
Consensus C
7.0
U: Ultra-deep pyrosequencing
SS: Sanger sequencing

Figure 5-2 Graphical representat**ion of CD4$^+$ T-cell count (grey line), viral load (black line) and DRM's found at selected time-points for participant 267, 268, 271, 079 and 036.**Time-points tested included; OPV: onset of plasma viremia (this represents the first sample obtained following initial detection of plasma viremia), PV: plasma viremia, VS: viral set-point, M: mid-point, L: last tested time-point. Participant 267, 268 and 271 did not have any persistence of mutations following OPV. For Participant 079 the RT-K103N mutation persisted at the mid and last time-point sequenced. Additionally, participant 079 developed the RT-K219R mutation at the last sequenced time-point. Participant 036 developed the RT-K65R mutation at VS and the InSTI-Y143C DRM and RT-K219Q DRM at the last sequenced time-point. Participant 267 and 036 were placed on cART. Both participants showed complete viral suppression by day 156 and 120 post treatment initiation respectively.

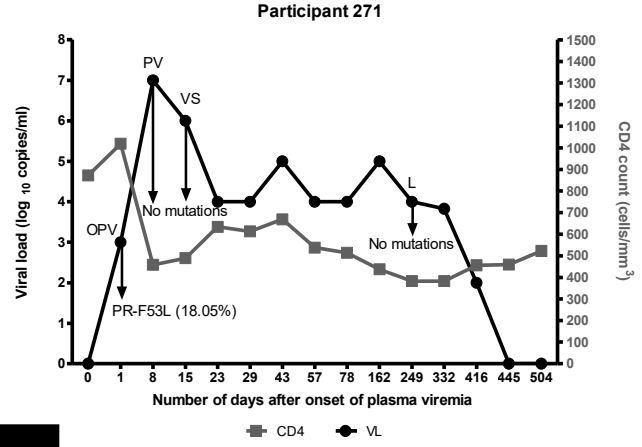
A


Participant 267
Viral load (log10 copies/ml)
CD4 count (cells/mm3)
PV
VS
OPV
No mutations
L
Txn
No mutations
INT-T97A (2.90%)
RT-D67N (3.88%)
0 2 7 13 21 28 41 56 77 163 245 329 448 496 612
156 days
Number of days after onset of plasma viremia
CD4
VL

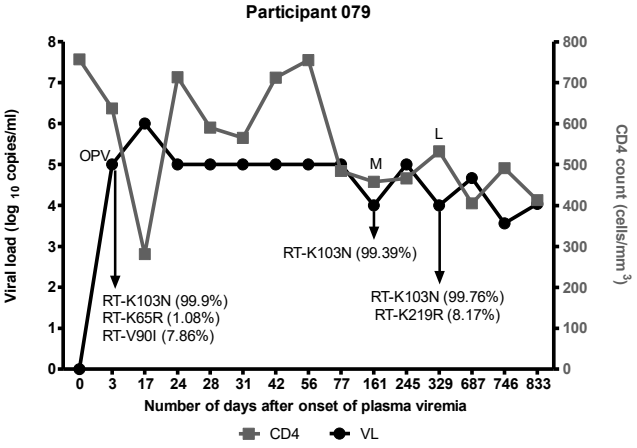
B


Participant 268
Viral load (log10 copies/ml)
CD4 count (cells/mm3)
OPV
RT-K65R (1.51%)
L
No mutations
0 1 7 14 24 28 42 56 77 164 243 326 411
Number of days after onset of plasma viremia
CD4
VL

C


Participant 271
Viral load (log 10 copies/ml)
CD4 count (cells/mm3)
PV
VS
No mutations
OPV
L
No mutations
PR-F53L (18.05%)
0 1 8 15 23 29 43 57 78 162 249 332 416 445 504
Number of days after onset of plasma viremia
CD4
VL

D


Participant 079
Viral load (log 10 copies/ml)
CD4 count (cells/mm3)
OPV
M
L
RT-K103N (99.39%)
RT-K103N (99.9%)
RT-K65R (1.08%)
RT-V90I (7.86%)
RT-K103N (99.76%)
RT-K219R (8.17%)
0 3 17 24 28 31 42 56 77 161 245 329 687 746 833
Number of days after onset of plasma viremia
CD4
VL

E

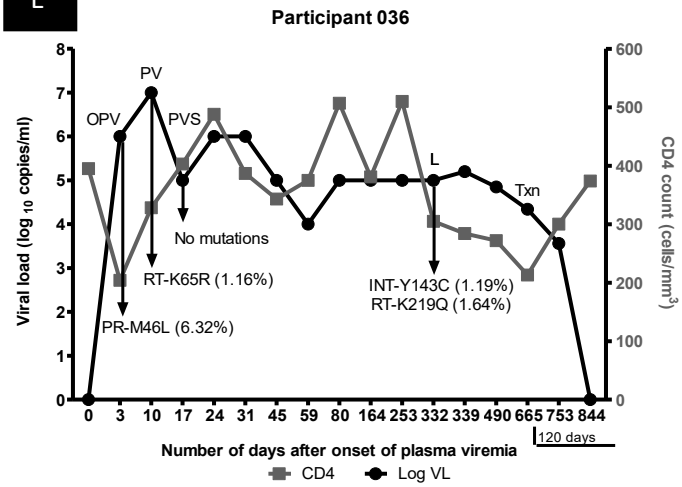
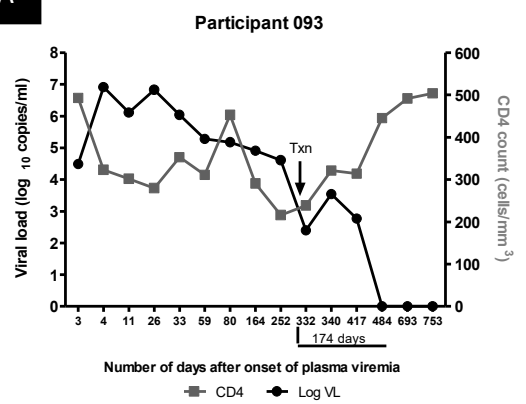
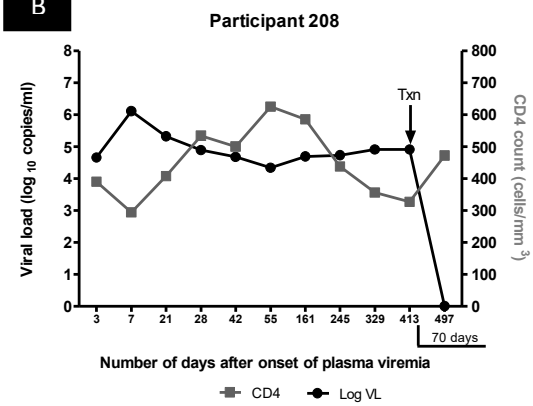

Participant 036
Viral load (log 10 copies/ml)
CD4 count (cells/mm3)
PV
OPV
PVS
L
Txn
No mutations
RT-K65R (1.16%)
INT-Y143C (1.19%)
RT-K219Q (1.64%)
PR-M46L (6.32%)
0 3 10 17 24 31 45 59 80 164 253 332 339 490 665 753 844
120 days
Number of days after onset of plasma viremia
CD4
Log VL

Figure 5-3 Comparison of CD4+ T-cell counts (grey) and log viral load (black) of five participants with no low frequency DRMs that were placed on cART Four participants were placed on treatment as per government guidelines (093, 186, 208 and 309) whilst one participant (312) was placed on early treatment a day after diagnosis. The day of treatment initiation (Txn) is indicated by a black arrow. Black boxes below the x-axis indicate the number of days taken to reach complete viral suppression post treatment initiation. Participant 093 reached complete viral suppression 174 days post treatment initiation, whilst participant 208 and 309 had undetectable viral loads by day 70 and 98 post treatment initiation respectively. All 3 of these participants showed an increase in $CD4^+$ T-cell counts which correlated with decreased viral load. Participant 186 received treatment for 429 days, during which time the viral load remained detectable and the $CD4^+$ T-cell count continued to decline. This was likely due to poor adherence to treatment, as evidenced by the development of the RT-K103N DRM which was detected by Sanger sequencing 311 days after treatment initiation. Participant 312 displayed an undetectable viral load 27 days post treatment initiation with an accompanied gradual increase in $CD4^+$ T-lymphocyte count.

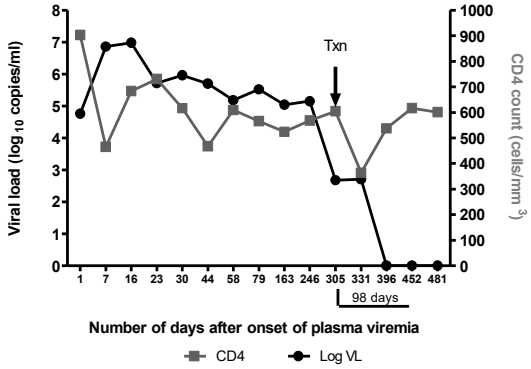
A


Participant 093
Viral load (log 10 copies/ml)
CD4 count (cells/mm3)
Txn
3 4 11 26 33 59 80 164 252 332 340 417 484 693 753
174 days
Number of days after onset of plasma viremia
CD4
Log VL

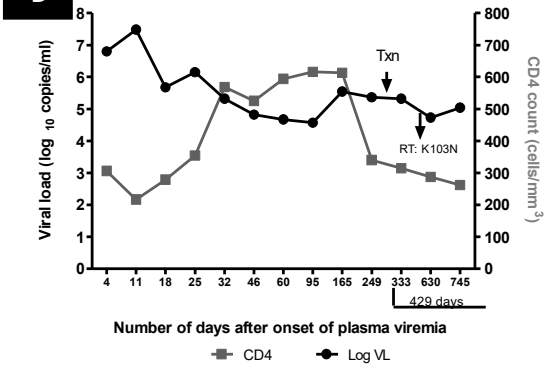
B


Participant 208
Viral load (log 10 copies/ml)
CD4 count (cells/mm3)
Txn
3 7 21 28 42 55 161 245 329 413 497
70 days
Number of days after onset of plasma viremia
CD4
Log VL

C


Participant 309
Viral load (log 10 copies/ml)
CD4 count (cells/mm3)
Txn
1 7 16 23 30 44 58 79 163 246 305 331 396 452 481
98 days
Number of days after onset of plasma viremia
CD4
Log VL

D


Participant 186
Viral load (log 10 copies/ml)
CD4 count (cells/mm3)
Txn
RT: K103N
4 11 18 25 32 46 60 95 165 249 333 630 745
429 days
Number of days after onset of plasma viremia
CD4
Log VL

E

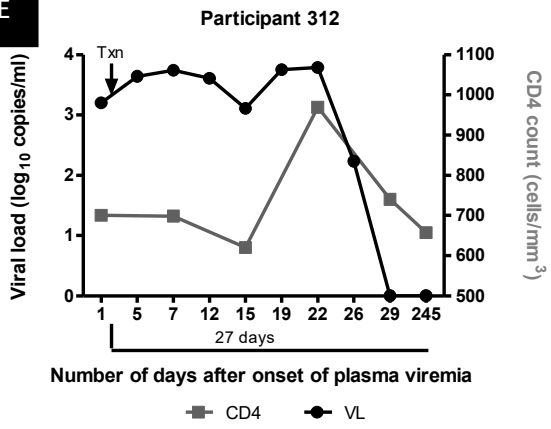

Participant 312
Viral load (log10 copies/ml)
CD4 count (cells/mm3)
Txn
1 5 7 12 15 19 22 26 29 245
27 days
Number of days after onset of plasma viremia
CD4
VL

Table 5-1 Overview of participant characteristics for the AI cohort and the FRESH cohort.

Characteristic	Acute Infection Cohort (n=32) Value (IQR) for the parameter	FRESH Cohort (n=15) Value (IQR) for the parameter
Gender (%) **Male** **Female**	 50 50	 0 100
Median age at sampling (years)	29 (25 – 37.5)	21 (20 – 22)
Fiebig Stage at time of testing[a] (%) **Stage I** **Stage II** **Stage V**	 72 3 5	 87 13
Subtype% **Subtype C**	 100	 100
Median Number of days post infection	14 (14 – 32)[b]	1 (1 – 3)[c]
Median viral load (log_{10} copies/ml)	5.8 (5.21 – 6.74)	4.76 (4.29 – 5.58)
Median $CD4^+$ T-Lymphocytes (cells/mm^3)	322 (289 – 519)	668 (411 – 859)
Median viral load set-point[d] (log_{10} copies/ml)	4.8 (4.1 – 5.2)	4.72 (4.07 – 4.99)
Median rate of $CD4^+$ T-lymphocyte decline (cells/mm^3 per month)	-4.4 (-9.8 to -0.3)	-4.1 (-7 to 5.4)

[a]Fiebig staging for 11 participants was as reported by Ndhlovu et al., 2015 [41], the remaining participants were staged as per Fiebig et al., 2003 [68]. [b]Represents the number of days post infection, calculated by adding 14 days to the date of the first positive test as described previously [42]. [c]Represents the number of days following onset of plasma viremia as described previously [41]. [d]The viral load set-point was defined as the average viral load from 3 to 12 months post infection.

Table 5-2 Comparison between DRMs detected by Sanger sequencing and UDPS at the earliest available time-point after onset of plasma viremia (OPV).

PID	# Days after OPV	Frequency of DRMs detected by UDPS and Sanger sequencing (% by UDPS)		Frequency of DRMs detected by UDPS only (% by UDPS)			
		NNRTI	PR[a]	NRTI	NNRTI	PR[a]	InSTI
36	8	None	None	**K65R (1.16)**	None	**M46L (6.32)**	None
39	6	None	T74S (99.39)	None	None	None	None
79	8	**K103N (99.93)**	A71V (0.3)	**K65R (1.08)**	**V90I (7.86)**	None	None
93	8	None	None	None	None	None	None
102	8	None	None	None	None	None	None
186	11	None	None	None	None	None	None
198	22	None	None	None	None	None	None
267	7	None	None	**D67N (3.88)**	None	None	**T97A (2.90)**
268	7	None	None	**K65R (1.51)**	None	None	None
271	5	None	None	None	None	F53L (17.6)[b]	None
272	6	None	None	**K65R (1.04)**	None	None	None
309	5	None	None	None	None	None	None
312	9	None	None	**K65R (1.19)**	None	None	None
318	6	None	None	**K65R (1.11)**	None	None	None

[a]Major PI DRMs are highlighted in bold.. [b]The F53L PI-associated DRM is listed as a minor PR DRM in the IAS-USA list of DRMs and in the Stanford HIVdb, however it is listed as a PI TDR associated mutation in the World Health Organization list of mutations for surveillance of transmitted drug resistant HIV strains. Abbreviations: OPV, Onset of plasma viremia; DRM, drug resistance mutation; UDPS, ultra-deep pyrosequencing; PR, Protease, NNRTI, non-nucleo(s)tide reverse transcriptase inhibitors; NRTI, nucleo(s)tide reverse transcriptase inhibitor; InSTI, integrase strand transfer inhibitor

References

1. Xiaobai Z, Xi C, Tian H, Williams AB, Wang H, He J*, et al.* Prevalence of WHO transmitted drug resistance mutations by deep sequencing in antiretroviral-naive subjects in Hunan Province, China. *PLoS One* 2014,9:e98740.

2. Palmer S, Kearney M, Maldarelli F, Halvas EK, Bixby CJ, Bazmi H*, et al.* Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. *J Clin Microbiol* 2005,43:406-413.

3. Johnson JA, Li JF, Wei X, Lipscomb J, Irlbeck D, Craig C*, et al.* Minority HIV-1 drug resistance mutations are present in antiretroviral treatment-naive populations and associate with reduced treatment efficacy. *PLoS Med* 2008,5:e158.

4. Svarovskaia ES, Margot NA, Bae AS, Waters JM, Goodman D, Zhong L*, et al.* Low-level K65R mutation in HIV-1 reverse transcriptase of treatment-experienced patients exposed to abacavir or didanosine. *J Acquir Immune Defic Syndr* 2007,46:174-180.

5. Palmer S, Boltz V, Maldarelli F, Kearney M, Halvas EK, Rock D*, et al.* Selection and persistence of non-nucleoside reverse transcriptase inhibitor-resistant HIV-1 in patients starting and stopping non-nucleoside therapy. *AIDS* 2006,20:701-710.

6. Paredes R, Lalama CM, Ribaudo HJ, Schackman BR, Shikuma C, Giguel F*, et al.* Pre-existing minority drug-resistant HIV-1 variants, adherence, and risk of antiretroviral treatment failure. *J Infect Dis* 2010,201:662-671.

7. Li JZ, Paredes R, Ribaudo HJ, Svarovskaia ES, Metzner KJ, Kozal MJ*, et al.* Low-frequency HIV-1 drug resistance mutations and risk of NNRTI-based antiretroviral treatment failure: a systematic review and pooled analysis. *JAMA* 2011,305:1327-1335.

8. Palmer S, Boltz V, Martinson N, Maldarelli F, Gray G, McIntyre J*, et al.* Persistence of nevirapine-resistant HIV-1 in women after single-dose nevirapine therapy for prevention of maternal-to-fetal HIV-1 transmission. *Proc Natl Acad Sci U S A* 2006,103:7094-7099.

9. Jourdain G, Ngo-Giang-Huong N, Le Coeur S, Bowonwatanuwong C, Kantipong P, Leechanachai P*, et al.* Intrapartum exposure to nevirapine and subsequent maternal responses to nevirapine-based antiretroviral therapy. *N Engl J Med* 2004,351:229-240.

10. Kapoor A, Jones M, Shafer RW, Rhee SY, Kazanjian P, Delwart EL. Sequencing-based detection of low-frequency human immunodeficiency virus type 1 drug-resistant mutants by an RNA/DNA heteroduplex generator-tracking assay. *J Virol* 2004,78:7112-7123.

11. Lecossier D, Shulman NS, Morand-Joubert L, Shafer RW, Joly V, Zolopa AR*, et al.* Detection of minority populations of HIV-1 expressing the K103N resistance mutation in patients failing nevirapine. *J Acquir Immune Defic Syndr* 2005,38:37-42.

12. Simen BB, Simons JF, Hullsiek KH, Novak RM, Macarthur RD, Baxter JD*, et al.* Low-abundance drug-resistant viral variants in chronically HIV-infected, antiretroviral treatment-naive patients significantly impact treatment outcomes. *J Infect Dis* 2009,199:693-701.
13. Kuritzkes DR, Lalama CM, Ribaudo HJ, Marcial M, Meyer WA, 3rd, Shikuma C*, et al.* Preexisting resistance to nonnucleoside reverse-transcriptase inhibitors predicts virologic failure of an efavirenz-based regimen in treatment-naive HIV-1-infected subjects. *J Infect Dis* 2008,197:867-870.
14. Charpentier C, Dwyer DE, Mammano F, Lecossier D, Clavel F, Hance AJ. Role of minority populations of human immunodeficiency virus type 1 in the evolution of viral resistance to Protease inhibitors. *J Virol* 2004,78:4234-4247.
15. Metzner KJ, Bonhoeffer S, Fischer M, Karanicolas R, Allers K, Joos B*, et al.* Emergence of minor populations of human immunodeficiency virus type 1 carrying the M184V and L90M mutations in subjects undergoing structured treatment interruptions. *J Infect Dis* 2003,188:1433-1443.
16. Metzner KJ, Giulieri SG, Knoepfel SA, Rauch P, Burgisser P, Yerly S*, et al.* Minority quasispecies of drug-resistant HIV-1 that lead to early therapy failure in treatment-naive and -adherent patients. *Clin Infect Dis* 2009,48:239-247.
17. Pou C, Noguera-Julian M, Perez-Alvarez S, Garcia F, Delgado R, Dalmau D*, et al.* Improved prediction of salvage antiretroviral therapy outcomes using ultrasensitive HIV-1 drug resistance testing. *Clin Infect Dis* 2014,59:578-588.
18. Cozzi-Lepri A, Noguera-Julian M, Di Giallonardo F, Schuurman R, Daumer M, Aitken S*, et al.* Low-frequency drug-resistant HIV-1 and risk of virological failure to first-line NNRTI-based ART: a multicohort European case-control study using centralized ultrasensitive 454 pyrosequencing. *J Antimicrob Chemother* 2015,70:930-940.
19. Casadella M, Manzardo C, Noguera-Julian M, Ferrer E, Domingo P, Perez-Alvarez S*, et al.* Clinical value of ultradeep HIV-1 genotyping and tropism testing in late presenters with advanced disease. *AIDS* 2015,29:1493-1504.
20. Little SJ, Holte S, Routy JP, Daar ES, Markowitz M, Collier AC*, et al.* Antiretroviral drug resistance amongst patients recently infected with HIV. *N Engl J Med* 2002,347:385-394.
21. Gadhamsetty S, Dixit NM. Estimating frequencies of minority nevirapine-resistant strains in chronically HIV-1-infected individuals naive to nevirapine by using stochastic simulations and a mathematical model. *J Virol* 2010,84:10230-10240.
22. Little SJ, Frost SD, Wong JK, Smith DM, Pond SL, Ignacio CC*, et al.* Persistence of transmitted drug resistance among subjects with primary human immunodeficiency virus infection. *J Virol* 2008,82:5510-5518.

23. Garcia-Lerma JG, Nidtha S, Blumoff K, Weinstock H, Heneine W. Increased ability for selection of zidovudine resistance in a distinct class of wild-type HIV-1 from drug-naive persons. *Proc Natl Acad Sci U S A* 2001,98:13907-13912.
24. Brenner B, Routy JP, Quan Y, Moisi D, Oliveira M, Turner D*, et al.* Persistence of multidrug-resistant HIV-1 in primary infection leading to superinfection. *AIDS* 2004,18:1653-1660.
25. Yang WL, Kouyos RD, Boni J, Yerly S, Klimkait T, Aubert V*, et al.* Persistence of transmitted HIV-1 drug resistance mutations associated with fitness costs and viral genetic backgrounds. *PLoS Pathog* 2015,11:e1004722.
26. Martinez-Picado J, Martinez MA. HIV-1 reverse transcriptase inhibitor resistance mutations and fitness: a view from the clinic and ex vivo. *Virus Res* 2008,134:104-123.
27. Cong ME, Heneine W, Garcia-Lerma JG. The fitness cost of mutations associated with human immunodeficiency virus type 1 drug resistance is modulated by mutational interactions. *J Virol* 2007,81:3037-3041.
28. Li JZ, Kuritzkes DR. Clinical implications of HIV-1 minority variants. *Clin Infect Dis* 2013,56:1667-1674.
29. Gandhi RT, Wurcel A, Rosenberg ES, Johnston MN, Hellmann N, Bates M*, et al.* Progressive reversion of human immunodeficiency virus type 1 resistance mutations in vivo after transmission of a multiply drug-resistant virus. *Clin Infect Dis* 2003,37:1693-1698.
30. Hirsch MS, Gunthard HF, Schapiro JM, Brun-Vezinet F, Clotet B, Hammer SM*, et al.* Antiretroviral drug resistance testing in adult HIV-1 infection: 2008 recommendations of an International AIDS Society-USA panel. *Clin Infect Dis* 2008,47:266-285.
31. Chun TW, Engel D, Berrey MM, Shea T, Corey L, Fauci AS. Early establishment of a pool of latently infected, resting CD4(+) T cells during primary HIV-1 infection. *Proc Natl Acad Sci U S A* 1998,95:8869-8873.
32. Chun TW, Fauci AS. HIV reservoirs: pathogenesis and obstacles to viral eradication and cure. *AIDS* 2012,26:1261-1268.
33. Metzner KJ, Rauch P, Walter H, Boesecke C, Zollner B, Jessen H*, et al.* Detection of minor populations of drug-resistant HIV-1 in acute seroconverters. *AIDS* 2005,19:1819-1825.
34. Avidor B, Girshengorn S, Matus N, Talio H, Achsanov S, Zeldis I*, et al.* Evaluation of a benchtop HIV ultradeep pyrosequencing drug resistance assay in the clinical laboratory. *J Clin Microbiol* 2013,51:880-886.
35. Halvas EK, Aldrovandi GM, Balfe P, Beck IA, Boltz VF, Coffin JM*, et al.* Blinded, multicenter comparison of methods to detect a drug-resistant mutant of human immunodeficiency virus type 1 at low frequency. *J Clin Microbiol* 2006,44:2612-2614.

36. Barzon L, Lavezzo E, Militello V, Toppo S, Palu G. Applications of next-generation sequencing technologies to diagnostic virology. *Int J Mol Sci* 2011,12:7861-7884.
37. Evans D. Ten years on ART where to now. *S Afr Med J* 2013,103:229-230.
38. Alcalde R, Guimaraes ML, Duarte AJ, Casseb J. Clinical, epidemiological and molecular features of the HIV-1 subtype C and recombinant forms that are circulating in the city of Sao Paulo, Brazil. *Virol J* 2012,9:156.
39. Radebe M, Nair K, Chonco F, Bishop K, Wright JK, van der Stok M*, et al.* Limited immunogenicity of HIV CD8+ T-cell epitopes in acute Clade C virus infection. *J Infect Dis* 2011,204:768-776.
40. Anahtar MN, Byrne EH, Doherty KE, Bowman BA, Yamamoto HS, Soumillon M*, et al.* Cervicovaginal bacteria are a major modulator of host inflammatory responses in the female genital tract. *Immunity* 2015,42:965-976.
41. Ndhlovu ZM, Kamya P, Mewalal N, Kloverpris HN, Nkosi T, Pretorius K*, et al.* Magnitude and Kinetics of CD8(+) T Cell Activation during Hyperacute HIV Infection Impact Viral Set Point. *Immunity* 2015,43:591-604.
42. van Loggerenberg F, Mlisana K, Williamson C, Auld SC, Morris L, Gray CM*, et al.* Establishing a cohort at high risk of HIV infection in South Africa: challenges and experiences of the CAPRISA 002 acute infection study. *PLoS One* 2008,3:e1954.
43. Wensing AM, Calvez V, Gunthard HF, Johnson VA, Paredes R, Pillay D*, et al.* 2014 Update of the drug resistance mutations in HIV-1. *Top Antivir Med* 2014,22:642-650.
44. Bennett DE, Camacho RJ, Otelea D, Kuritzkes DR, Fleury H, Kiuchi M*, et al.* Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. *PLoS One* 2009,4:e4724.
45. Borras E, Jurado I, Hernan I, Gamundi MJ, Dias M, Marti I*, et al.* Clinical pharmacogenomic testing of KRAS, BRAF and EGFR mutations by high resolution melting analysis and ultra-deep pyrosequencing. *BMC Cancer* 2011,11:406.
46. Messiaen P, Verhofstede C, Vandenbroucke I, Dinakis S, Van Eygen V, Thys K*, et al.* Ultra-deep sequencing of HIV-1 reverse transcriptase before start of an NNRTI-based regimen in treatment-naive patients. *Virology* 2012,426:7-11.
47. Wang C, Mitsuya Y, Gharizadeh B, Ronaghi M, Shafer RW. Characterization of mutation spectra with ultra-deep pyrosequencing: application to HIV-1 drug resistance. *Genome Res* 2007,17:1195-1201.
48. Garcia-Diaz A, Guerrero-Ramos A, McCormick AL, Macartney M, Conibear T, Johnson MA*, et al.* Evaluation of the Roche prototype 454 HIV-1 ultradeep sequencing drug resistance assay in a routine diagnostic laboratory. *J Clin Virol* 2013,58:468-473.
49. Vandenbroucke I, Van Marck H, Verhasselt P, Thys K, Mostmans W, Dumont S*, et al.* Minor variant detection in amplicons using 454 massive parallel pyrosequencing:

experiences and considerations for successful applications. *Biotechniques* 2011,51:167-177.

50. Knapp DJ, McGovern RA, Poon AF, Zhong X, Chan D, Swenson LC*, et al.* "Deep" sequencing accuracy and reproducibility using Roche/454 technology for inferring co-receptor usage in HIV-1. *PLoS One* 2014,9:e99508.
51. Rambaut A. Figtree. In; 2015.
52. National Department of Health. National consolidated guidelines for the prevention of mother to child transmission (PMTCT) of HIV and the management of HIV in children, adolescents and adults. In: National Department of Health 2014.
53. Deeks SG, Wrin T, Liegler T, Hoh R, Hayden M, Barbour JD*, et al.* Virologic and immunologic consequences of discontinuing combination antiretroviral-drug therapy in HIV-infected patients with detectable viremia. *N Engl J Med* 2001,344:472-480.
54. Capetti AF, Gabris AI, Drago L, Vigevani GM. Can a K103N HIV strain stably overcome the wild type in the absence of non-nucleoside reverse transcriptase inhibitor selective pressure? *AIDS* 2005,19:633-634.
55. Nicastri E, Sarmati L, d'Ettorre G, Palmisano L, Parisi SG, Uccella I*, et al.* Replication capacity, biological phenotype, and drug resistance of HIV strains isolated from patients failing antiretroviral therapy. *J Med Virol* 2003,69:1-6.
56. Margot NA, Lu B, Cheng A, Miller MD. Resistance development over 144 weeks in treatment-naive patients receiving tenofovir disoproxil fumarate or stavudine with lamivudine and efavirenz in Study 903. *HIV Med* 2006,7:442-450.
57. Dauber DS, Ziermann R, Parkin N, Maly DJ, Mahrus S, Harris JL*, et al.* Altered substrate specificity of drug-resistant human immunodeficiency virus type 1 Protease. *J Virol* 2002,76:1359-1368.
58. Malim MH, Emerman M. HIV-1 sequence variation: drift, shift, and attenuation. *Cell* 2001,104:469-472.
59. Yanik EL, Napravnik S, Hurt CB, Dennis A, Quinlivan EB, Sebastian J*, et al.* Prevalence of transmitted antiretroviral drug resistance differs between acutely and chronically HIV-infected patients. *J Acquir Immune Defic Syndr* 2012,61:258-262.
60. Brenner BG, Oliveira M, Doualla-Bell F, Moisi DD, Ntemgwa M, Frankel F*, et al.* HIV-1 subtype C viruses rapidly develop K65R resistance to tenofovir in cell culture. *AIDS* 2006,20:F9-13.
61. Doualla-Bell F, Avalos A, Brenner B, Gaolathe T, Mine M, Gaseitsiwe S*, et al.* High prevalence of the K65R mutation in human immunodeficiency virus type 1 subtype C isolates from infected patients in Botswana treated with didanosine-based regimens. *Antimicrob Agents Chemother* 2006,50:4182-4185.

62. Hosseinipour MC, van Oosterhout JJ, Weigel R, Phiri S, Kamwendo D, Parkin N*, et al.* The public health approach to identify antiretroviral therapy failure: high-level nucleoside reverse transcriptase inhibitor resistance among Malawians failing first-line antiretroviral therapy. *AIDS* 2009,23:1127-1134.
63. Orrell C, Walensky RP, Losina E, Pitt J, Freedberg KA, Wood R. HIV type-1 clade C resistance genotypes in treatment-naive patients and after first virological failure in a large community antiretroviral therapy programme. *Antivir Ther* 2009,14:523-531.
64. Coutsinos D, Invernizzi CF, Xu H, Brenner BG, Wainberg MA. Factors affecting template usage in the development of K65R resistance in subtype C variants of HIV type-1. *Antivir Chem Chemother* 2010,20:117-131.
65. Metzner KJ, Scherrer AU, Preiswerk B, Joos B, von Wyl V, Leemann C*, et al.* Origin of minority drug-resistant HIV-1 variants in primary HIV-1 infection. *J Infect Dis* 2013,208:1102-1112.
66. Ceccherini-Silberstein F, Van Baelen K, Armenia D, Trignetti M, Rondelez E, Fabeni L*, et al.* Secondary integrase resistance mutations found in HIV-1 minority quasispecies in integrase therapy-naive patients have little or no effect on susceptibility to integrase inhibitors. *Antimicrob Agents Chemother* 2010,54:3938-3948.
67. Parczewski M, Bander D, Urbanska A, Boron-Kaczmarska A. HIV-1 integrase resistance among antiretroviral treatment naive and experienced patients from Northwestern Poland. *BMC Infect Dis* 2012,12:368.
68. Fiebig EW, Wright DJ, Rawal BD, Garrett PE, Schumacher RT, Peddada L*, et al.* Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. *AIDS* 2003,17:1871-1879.

CHAPTER 6

General discussion

6 CHAPTER 6: GENERAL DISCUSSION AND CONCLUSIONS

6.1 Discussion

South Africa has one of the world's highest HIV infection rates with approximately 6.8 million people (out of a population of 50.59 million) living with HIV (1). In addition to this, SA has one of the world's largest ARV programs serving over 1.7 million people, and reaching almost 60% of HIV infected individuals that are eligible for ARV treatment in the country (1, 2). The country aims to have 80% of all HIV positive candidates who are eligible for ARV treatment on ARV's by 2016 as part of the National Strategic Plan on HIV/AIDS, Tuberculosis and Sexually Transmitted Infections. The major concern however, in having a large ARV program in a low-income country, is the development of drug resistance (i.e. acquired drug resistance) and its subsequent transmission. Acquired drug resistance occurs as a result of the development of resistance mutations within the HIV genome due to drug-selection pressure in individuals receiving ARV treatment. Transmitted drug resistance occurs when individuals who were not previously infected with HIV are infected with a drug-resistant variant of the virus (3). Adherence to treatment is the key to preventing/controlling acquired drug resistance which in turn would reduce the transmission of resistant variants of HIV. However several barriers to adherence exists which hamper the success of treatment. These include: poor service delivery, drug stock-outs, stigma and discrimination associated with ARV use, drug toxicity, socio-economic issues and lack of knowledge on the importance of adherence (4).

Within South Africa, reports of increased resistance to first-line regimens has been followed by concomitant increases in TDR to NRTI's and NNRTI's (3, 5). A recent report showed that NRTI and NNRTI TDR had increased from a low threshold level in 2007 to a moderate threshold level by 2008 (5-15%) (3). This increase would directly impact on the success of the country's first-line regimen, resulting in an increase in first-line failures and is likely to be accompanied by an increase in the number of patients receiving second line treatment. Currently, it is estimated that 10% of all people receiving cART in SA are on a PI inclusive treatment regimen. Interestingly, a recent study reported that 40% of all patients on second line cART in SA experience virologic failure, in the absence of mutations in PR (6-10). Whilst this could be due to poor treatment adherence, studies have shown that resistance to PI's can occur in the absence of mutations in PR (11-15). These studies recognise mutations in Gag as contributors to PI resistance (11, 13, 14, 16). The majority of studies investigating the role of Gag in PI resistance have used HIV-1 subtype B cohorts, despite HIV-1 subtype C being the most prevalent subtype globally (17). To date, only one study investigating the role of Gag in PI resistance exists for HIV-1 subtype C.

This study however had a small sample size (n=20) and was based on a paediatric cohort receiving RTV or LPV/r (18). Furthermore this study only investigated drug susceptibility with no analysis conducted on replication capacity, an integral component in PI resistance, since several studies have reported that mutations in Gag contribute to PI resistance via an increase in Gag polyprotein processing (19-25). Given that Gag is highly polymorphic amongst subtypes (26), a study investigating the role of Gag in PI resistance which addresses both replication capacity and drug susceptibility in HIV-1 subtype C is warranted.

With the high rate of virologic failure to both first and second line treatment, as discussed above, the need for more studies on TDR is recognised. To date, the majority of data used to generate TDR reports has primarily employed Sanger sequencing, which means that only mutations occurring at frequencies >15-20% are accounted for (27, 28). With various studies reporting that RAMs present at frequencies as low as 1% in the viral quasispecies could significantly impact treatment outcomes, the use of more sensitive techniques to detect low frequency mutations would be beneficial (29-37). Furthermore, most TDR studies have used recently infected cohorts of patients to study TDR. With the high fitness cost associated with most RAMs, they revert rapidly and thus would be undetectable in recent infection (i.e. 3-6 months post infection). Collectively this data highlights the need for studies which investigate low frequency TDR mutations in acute infection.

The current study used a two-pronged approach to address drug resistance surveillance within the context of acquired and transmitted drug resistance in SA. As part of the first approach; a cohort of 80 HIV-1 subtype C participants failing a PI inclusive treatment regimen (i.e. PCS cohort) were investigated to identify mutations in Gag and PR associated with PI resistance/exposure. Thereafter replication capacity of each virus was measured and selected viruses were used in drug susceptibility analysis in order to identify the impact of mutations on replication capacity and drug susceptibility. Lastly, the impact of selected mutations on replication capacity and drug susceptibility was validated using site-directed mutagenesis. As part of the second arm, prevalence of TDR mutations in a cohort of 45 HIV-1 subtype C acutely infected individuals was investigated using Sanger sequencing. Thereafter a subset of 14 participant samples, selected based on sample availability, was sequenced by UDPS to identify low frequency mutations. Lastly, we explored the impact of low frequency TDR mutations on treatment outcomes by comparing treatment outcomes of two participants with low frequency mutations to that of five participants without low frequency mutations.

Of the participants from the PCS cohort used in the current study, PR RAMs were detected in 34 out of 80 participants (i.e. 42.5%) (38). This implied that the remaining 57.5% (46/80) of

participants failed their PI inclusive treatment regimen as a result of either poor drug adherence or an alternate pathway of PI resistance. Of these 46 participants, only seven did not have detectable LPV levels suggesting that PI failure for the remaining participants was likely due to an alternate pathway of resistance. With the role of Gag in PI resistance being well documented (11, 13, 15), we investigated the Gag gene of each of these participants. Gag mutations were divided into three groups (a) PI exposure associated Gag mutations; (b) resistance associated Gag mutations (i.e. rGag) and (c) novel Gag mutations (i.e. nGag). Both rGag mutations and PI exposure associated Gag mutations used in this study were from existing literature (Chapter 1, Table 1.3) whilst nGag mutations represented mutations that occurred at a significantly higher frequency in the PCS cohort in comparison to the control cohorts (i.e. HIV-1 subtype C treatment naïve and acute infection cohorts).

All participants in our cohort presented with PI exposure associated Gag mutations (i.e. E12K, V35I, G62R, V370A/M, S373P/Q/T, A374P, T375N, I376V, G381S, I389T, I401T and H219Q). The E12K, V370A, T375N and G381S mutations all occurred as natural polymorphisms in HIV-1 subtype C (i.e. they were found at frequencies >50% in the treatment naïve cohorts). The role of these mutations in PI resistance is largely unknown, with one study suggesting that E12K may be involved in altering viral replication in the presence of PI's (39) and a second study showing that V370A confers resistance to Bevirimat (40), whilst no information on T375N or G381S exists within this context. Although the role of these mutations in PI resistance is unclear, their high prevalence in subtype C viruses in comparison to subtype B viruses highlights the polymorphic nature of Gag between subtypes and shows that research on subtype B viruses may not always translate to subtype C viruses (41).

We identified eight rGag mutations in this study (i.e. R76K, Y79F, V128I, A431V, K436R, L449F/P, R452K and P453L). The R76K and Y79F mutations appeared as natural polymorphisms in HIV-1 subtype C, further highlighting variations between HIV-1 subtype C and B Gag. Interestingly a previous study reported that a combination of R76K and Y79F was associated with increasing replication capacity in the presence of PR RAMs (16). This could suggest that HIV-1 subtype C viruses may have a replicative advantage in the presence of PR RAMs. The NC/p1 CS mutations: A431V and K436R, as well as the p1/p6 CS mutations: L449F and R452K, have all been associated with enhancing PI resistance in the presence of PR RAMs (13, 14, 20, 25). These CS mutations have been shown to alter CS structure thereby increasing accessibility of mutant PR to the substrate (19). This allows for continued substrate cleavage in the presence of PI's and thus these rGag CS mutations are said to enhance PI resistance through improving polyprotein processing (13, 14, 20, 25) . Bearing in mind that previous studies showed all rGag mutations identified in this study to impact on polyprotein processing in the

presence of PR RAMs, it was interesting to find that all 34 participants with PR RAMs in our cohort harboured ≥1 rGag mutation. Having both rGag mutations and PR RAMs would enable sustained substrate recognition and cleavage by the mutated PR. The combined presence of rGag and PR RAMs in our cohort could suggest that Gag and Protease co-evolve under PI drug selection pressure. This is supported by a recent study which showed that L449F co-evolves with D30N and N88S in PR to enable continued cleavage in the presence of PI's (23, 42).

Whether Gag or PR mutations appear first in PI resistance is currently a leading area of interest within the field. Our study showed that 72% (i.e. 33/46) of participants without PR RAMs presented with ≥1 rGag mutation described above. This suggests that development of mutations in Gag may precede the development of mutations in PR and could indicate that Gag mutations may be required for PR RAMs to develop. This is supported by several studies which also showed Gag mutations to occur before PR RAMs (39, 43, 44). These findings indicate that genotypic testing of PR is not sufficient to identify PI resistance, and advocates for the inclusion of Gag mutations in resistance algorithms, since Gag mutations may be an indicator/predictor for the development of PR RAMs and PI resistance.

The current study also identified four novel Gag mutations (i.e. Q69K, S111C/I, T239A/S and I256V), in the amino terminal of Gag, which were associated with PI resistance/exposure. Similar to rGag mutations, all 34 participants harbouring a PR RAM (and an rGag mutation) also harboured ≥1 nGag mutation, whilst 80% (i.e. 37/46) of participants without PR RAMs harboured ≥1 nGag mutation. This shows similar development of nGag and rGag mutations and could indicate that nGag mutations also precede PR RAM development further supporting the theory that Gag mutations precede the development of PR RAMs. In exploring the role of nGag mutations in PI resistance, we found Q69K to be associated with significantly increased replication capacity when found together with A431V+PR RAMs. Investigations using SDM validated that Q69K is a compensatory mutation which increases replication capacity when found in conjunction with A431V+V82A. Interestingly though, Q69K had no impact on reducing drug susceptibility either individually or in combination (i.e. A431V+PR RAMs+Q69K). The mechanism of action of the Q69K mutation could thus be similar to other rGag mutations such as L449F, R452K and P453L which have all been shown to improve polyprotein processing in the presence of PR RAMs, but do not confer reduced PI drug susceptibility when found alone (20, 22, 25, 45).

In exploring replication capacity of viruses with combinations of PR RAMs, rGag and nGag mutations, we showed that viruses with PR RAMs had a significantly lower replication capacity than viruses without PR RAMs. This was consistent with previous studies (46-48) and was

expected since PR RAMs alter the substrate binding cleft thereby reducing substrate cleavage (49-51). Viruses with PR RAMs+rGag mutations had a significantly lower replication capacity than viruses with rGag mutations only, suggesting that PR RAMs drive the reduction of viral fitness in PI resistance. A recent study also showed a similar result (46). Interestingly viruses with rGag mutations showed no variation in replication capacity when compared to: viruses with nGag mutations and viruses without Gag mutations. This was surprising since most rGag mutations are found in Gag CS's and have been associated with altering CS structure to facilitate better recognition by mutant PR (11, 52). One would thus expect to observe a reduced replication capacity in viruses with rGag mutations and wild-type PR. This however was not the case and highlighted that wild-type and mutant PR can recognise and cleave Gag CS's despite peptide variations.

Analysis of drug susceptibility data for 18 selected patients demonstrated, as expected, that PR RAMs are associated with significantly reduced susceptibility to LPV (>15 FC in IC_{50}) and DRV (>6 FC in IC_{50}) (20, 23, 53). This is largely due to alterations in the PR active site caused by PR RAMs which reduces the affinity for PI's thereby resulting in a higher IC_{50} (11). Interestingly, we showed that combinations of rGag mutations (in the absence of PR mutations) can also impact drug susceptibility, as demonstrated previously (16, 39, 54, 55). In our cohort, a virus with R76K+Y79F+K436R+L449P+I256V exhibited a 5.2 fold increase in IC_{50} for DRV. Similarly, a virus harbouring R76K+R453L had a 23.88 fold increase in IC_{50} for LPV and a 6.73 fold increase in IC_{50} for DRV. Lastly, a virus with R76K+K436R+Q69K+S111C displayed a 7.40 fold increase in IC_{50} for LPV. These data highlights that combinations of rGag mutations are associated with reduced PI susceptibility and further supports the recommendation to include Gag mutations associated with PI resistance in resistance algorithms.

Whilst PI virologic failure can be explained for most of the participants in the PCS cohort by the presence of PR RAMs, rGag mutations and poor adherence, there were a few participants that did not harbour PR RAMs or rGag mutations for whom PI virologic failure is difficult to explain. Each of these patients had detectable LPV levels indicative of treatment adherence and suggestive of an alternate pathway to PI resistance. Interestingly, a recent study showed that sequence variation at the cytoplasmic tail of the Env protein conferred resistance to PI's in the absence of mutations in PR (56). This region of HIV-1 is currently ignored in studies investigating PI resistance and thus we recommend future studies on PI resistance to include Env together with Gag and PR.

The current study was the first to investigate the prevalence of low frequency TDR mutations in PR, RT and INT within an HIV-1 subtype C acute infection cohort. Low frequency TDR

associated mutations were identified in 57% of participants by UDPS. We showed the presence of TDR associated mutations in PR for the first time in a South African study (i.e. M46I and F53L). Interestingly both of these mutations reverted seven days after onset of plasma viremia, highlighting the high fitness costs associated with PR RAMs. Similarly, the T97A, D67N and K65R mutations also reverted within one to two weeks after onset of plasma viremia. The quick reversion of these TDR associated mutations, shows that their presence would go undetected in recently infected individuals. As such the prevalence of TDR mutations could be under estimated in SA, since most studies investigating TDR mutations have utilised recently infected cohorts (3, 57). In our study, UDPS detected TDR mutations in 57% of participants (i.e. 8/14) whilst Sanger sequencing detected TDR in 2.2% of participants (1/45). This highlights that Sanger sequencing does not provide a true reflection of TDR mutations and shows the importance of using deep sequencing methods to obtain a true view of TDR. Collectively this data shows that the best view of TDR mutations would be obtained by using acute infection cohorts and deep sequencing techniques. It is however important to note that the high prevalence of low frequency mutations as identified by UDPS (i.e. 57%), could be related to either transmission events or "sporadic" emergence of mutants as a result of de novo viral replication. If the latter is true, the data may not reflect transmission events and implications for resistance to regimens may be different.

In exploring the impact of low frequency mutations on treatment outcomes, we found that on average, participants with low frequency TDR mutations who were initiated on FDC ART took 40 days longer to reach viral suppression than participants without low frequency TDR associated mutations. Interestingly, these participants also had significantly lower viral load than participants without low frequency TDR mutations, however they still took a longer time to achieve viral suppression. These data concur with other studies that have reported that low frequency TDR mutations impacts treatment outcomes and highlights the need for deep sequencing based genotypic testing. The need for cost-effective next generation sequencing technologies in low income countries is thus recognised.

6.2 Conclusions

Variability in replication capacity and drug susceptibility is driven primarily by mutations in PR, with mutations in Gag playing mostly a compensatory role to enhance polyprotein processing. Mutations in Gag can however also confer reduced susceptibility to PI's in the absence of PR RAMs either when found alone, as was the case for A431V, or when found in combinations. Furthermore Gag mutations may occur before PR RAMs and could be a precursor of PI resistance. Collectively this data advocates for the inclusion of Gag in PI resistance algorithms.

Deep sequencing technologies identify a significantly higher number of TDR associated mutations than Sanger sequencing as a result of the limit of detection associated with Sanger sequencing (i.e. it can only detect mutations present at a frequency >15-20%). Low frequency mutations impact treatment outcomes thereby highlighting the importance of genotypic screening prior to treatment initiation. The use of deep sequencing for genotypic screening would be more beneficial than using Sanger sequencing. There is thus a need for the development of cost-effective deep-sequencing technologies for use in resource-limited settings, like SA.

Overall this study showed that both the carboxy and amino terminal of Gag play an important role in PI resistance. Furthermore, TDR mutations to all four drug classes (i.e. NRTI, NNRTI, PI and INI) exists in SA. Genotypic resistance testing using deep sequencing technologies would be beneficial in maintaining the success and longevity of cART in SA.

6.3 Study limitations

- Although we measured LPV levels for each patient as an indicator of treatment adherence, we do recognise that that "white-coat" adherence (i.e. cases were patients take treatment prior to clinic visits) is a problem and that measuring adherence is difficult and not always truly reflective of patient behaviour.
- The high cost associated with drug susceptibility testing limited the number of patient samples we tested. Nonetheless, the study was designed to answer specific questions which could be addressed with the susceptibility analysis conducted.
- The inability to calculate an upper limit of detection for phenotypic tests using lab-derived viruses limited our drug resistance classifications to susceptible or reduced susceptibility only, without any intermediate categorisations of drug susceptibility. Whilst this is recognised as a limitation, it was not detrimental to our study since we wanted to determine if mutations affected drug susceptibility, a question which could be answered by describing only the FC in IC_{50} values for LPV or DRV.
- The small sample size used in the study to detect low frequency TDR associated mutations is acknowledged as a limitation. It must however be recognised that the FRESH acute cohort is the first of its kind globally since participants were followed prior to developing HIV with bi-weekly tests conducted to identify HIV-1 infection at its earliest detectable point. Most patients in this cohort were thus identified as early as Fiebig stage 1 and are truly representative of acute infection.

- The Env region of HIV-1 was not included in the present study. As such mechanisms of PI failure in several patients could not be determined and it remains unclear if some of these patients failed treatment due to poor adherence or because of the Env based alternate pathway of PI resistance.
- An additional limitation of this study was the insertion of subtype C patient derived gag-protease into a subtype B backbone (pNL43) to create recombinant viruses which were used to assess viral fitness and drug susceptibility of patient derived gag-protease viruses. Ideally this study should have employed a subtype C backbone (e.g. pMJ4), however generation of recombinant viruses in pMJ4 has been a challenge in our laboratory. Whilst it is possible that some gag-protease mutations associated with replication capacity or drug susceptibility could have been influenced by other components of the subtype B backbone, previous studies have shown that the subtype of a backbone has no effect on replication capacity (58) or drug susceptibility results (59).

6.4 Recommendations for future studies

- Genotyping the Env region of HIV-1 for the PCS cohort is recommended in order to identify mutations in gp41 that are associated with PI resistance.
- Including Env in Gag-Protease recombinant viruses is suggested. Employing these viruses in viral replication and drug susceptibility assays would be beneficial to determine how virus functionality is affected when these three regions are combined under drug selection pressure.
- Identification of the quasispecies harboured in viral reservoirs is recommended in order to elucidate if drug resistant minority variant mutants are harboured in viral reservoirs.

6.5 References

1. UNAIDS. 2012. Regional Fact Sheet - Sub Saharan Africa. http://www.unaids.org/en/media/unaids/contentassets/documents/epidemiology/2012/gr2012/2012_FS_regional_ssa_en.pdf. Accessed 15 March
2. SA AIDS. 2013. Southern African HIV and AIDS information dissemination service. South Africa" National Strategic Plan 2012-2016 – Living with HIV. http://www.safaids.net/content/south-africa-national-strategic-plan-2012-2016-living-hiv Accessed 17 March.
3. World Health Organisation. 2012. Drug Resistance Report www.who.int/hiv/pub/drugresistance/report2012/en/index.html Accessed 13 March
4. Reda AA, Biadgilign S. 2012. Determinants of Adherence to Antiretroviral Therapy among HIV-Infected Patients in Africa. AIDS Res Treat 2012:574656.
5. Marconi VC, Sunpath H, Lu Z, Gordon M, Koranteng-Apeagyei K, Hampton J, Carpenter S, Giddy J, Ross D, Holst H, Losina E, Walker BD, Kuritzkes DR, South Africa Resistance Cohort Study T. 2008. Prevalence of HIV-1 drug resistance after failure of a first highly active antiretroviral therapy regimen in KwaZulu Natal, South Africa. Clin Infect Dis 46:1589-1597.
6. Levison JH, Orrell C, Gallien S, Kuritzkes DR, Fu N, Losina E, Freedberg KA, Wood R. 2012. Virologic failure of Protease inhibitor-based second-line antiretroviral therapy without resistance in a large HIV treatment program in South Africa. PLoS One 7:e32144.
7. Richard A. Murphy HS, Zhigang Lu, Neville Chelin, Elena Losina, Michelle Gordon, Douglas Ross, Aba D. Ewusi, Lynn T. Matthews, Daniel R. Kuritzkes, Vincent C. Marconi. 2010. Outcomes After Virologic Failure of First-line ART in South Africa. AIDS 24:1007-1012.
8. Hosseinipour MC, Kumwenda JJ, Weigel R, Brown LB, Mzinganjira D, Mhango B, Eron JJ, Phiri S, van Oosterhout JJ. 2010. Second-line treatment in the Malawi antiretroviral programme: high early mortality, but good outcomes in survivors, despite extensive drug resistance at baseline. HIV Med 11:510-518.
9. Fox MP, Ive P, Long L, Maskew M, Sanne I. 2010. High rates of survival, immune reconstitution, and virologic suppression on second-line antiretroviral therapy in South Africa. J Acquir Immune Defic Syndr 53:500-506.
10. Levison JH, Orrell C, Losina E, Lu Z, Freedberg KA, Wood R. 2011. Early outcomes and the virological effect of delayed treatment switching to second-line therapy in an antiretroviral roll-out programme in South Africa. Antivir Ther 16:853-861.

11. Fun A, Wensing AM, Verheyen J, Nijhuis M. 2012. Human Immunodeficiency Virus Gag and Protease: partners in resistance. Retrovirology 9:63.
12. Parry CM, Kohli A, Boinett CJ, Towers GJ, McCormick AL, Pillay D. 2009. Gag determinants of fitness and drug susceptibility in Protease inhibitor-resistant human immunodeficiency virus type 1. J Virol 83:9094-9101.
13. Nijhuis M, van Maarseveen NM, Lastere S, Schipper P, Coakley E, Glass B, Rovenska M, de Jong D, Chappey C, Goedegebuure IW, Heilek-Snyder G, Dulude D, Cammack N, Brakier-Gingras L, Konvalinka J, Parkin N, Krausslich HG, Brun-Vezinet F, Boucher CA. 2007. A novel substrate-based HIV-1 Protease inhibitor drug resistance mechanism. PLOS Medicine 4:e36.
14. Dam E, Quercia R, Glass B, Descamps D, Launay O, Duval X, Krausslich HG, Hance AJ, Clavel F, Group AS. 2009. Gag mutations strongly contribute to HIV-1 resistance to Protease inhibitors in highly drug-experienced patients besides compensating for fitness loss. PLoS Pathog 5:e1000345.
15. Clavel F, Mammano F. 2010. Role of Gag in HIV Resistance to Protease Inhibitors. Viruses 2:1411-1426.
16. Parry CM, Kolli M, Myers RE, Cane PA, Schiffer C, Pillay D. 2011. Three residues in HIV-1 matrix contribute to Protease inhibitor susceptibility and replication capacity. Antimicrob Agents Chemother 55:1106-1113.
17. Alcalde R, Guimaraes ML, Duarte AJ, Casseb J. 2012. Clinical, epidemiological and molecular features of the HIV-1 subtype C and recombinant forms that are circulating in the city of Sao Paulo, Brazil. Virol J 9:156.
18. Giandhari J, Basson AE, Coovadia A, Kuhn L, Abrams EJ, Strehlau R, Morris L, Hunt GM. 2015. Genetic Changes in HIV-1 Gag-Protease Associated with Protease Inhibitor-Based Therapy Failure in Pediatric Patients. AIDS Res Hum Retroviruses 31:776-782.
19. Prabu-Jeyabalan M, Nalivaika EA, King NM, Schiffer CA. 2004. Structural basis for coevolution of a human immunodeficiency virus type 1 nucleocapsid-p1 cleavage site with a V82A drug-resistant mutation in viral Protease. J Virol 78:12446-12454.
20. Kolli M, Stawiski E, Chappey C, Schiffer CA. 2009. Human immunodeficiency virus type 1 Protease-correlated cleavage site mutations enhance inhibitor resistance. J Virol 83:11027-11042.
21. Shibata J, Sugiura W, Ode H, Iwatani Y, Sato H, Tsang H, Matsuda M, Hasegawa N, Ren F, Tanaka H. 2011. Within-host co-evolution of Gag P453L and Protease D30N/N88D demonstrates virological advantage in a highly Protease inhibitor-exposed HIV-1 case. Antiviral Res 90:33-41.
22. Maguire MF, Guinea R, Griffin P, Macmanus S, Elston RC, Wolfram J, Richards N, Hanlon MH, Porter DJ, Wrin T, Parkin N, Tisdale M, Furfine E, Petropoulos C, Snowden

BW, Kleim JP. 2002. Changes in human immunodeficiency virus type 1 Gag at positions L449 and P453 are linked to I50V Protease mutants in vivo and cause reduction of sensitivity to amprenavir and improved viral fitness in vitro. J Virol 76:7398-7406.

23. Kolli M, Lastere S, Schiffer CA. 2006. Co-evolution of nelfinavir-resistant HIV-1 Protease and the p1-p6 substrate. Virology 347:405-409.
24. Feher A, Weber IT, Bagossi P, Boross P, Mahalingam B, Louis JM, Copeland TD, Torshin IY, Harrison RW, Tozser J. 2002. Effect of sequence polymorphism and drug resistance on two HIV-1 Gag processing sites. Eur J Biochem 269:4114-4120.
25. Verheyen J, Litau E, Sing T, Daumer M, Balduin M, Oette M, Fatkenheuer G, Rockstroh JK, Schuldenzucker U, Hoffmann D, Pfister H, Kaiser R. 2006. Compensatory mutations at the HIV cleavage sites p7/p1 and p1/p6-Gag in therapy-naive and therapy-experienced patients. Antivir Ther 11:879-887.
26. Sutherland KA, Mbisa JL, Cane PA, Pillay D, Parry CM. 2014. Contribution of Gag and Protease to variation in susceptibility to Protease inhibitors between different strains of subtype B human immunodeficiency virus type 1. J Gen Virol 95:190-200.
27. Xiaobai Z, Xi C, Tian H, Williams AB, Wang H, He J, Zhen J, Chiarella J, Blake LA, Turenchalk G, Kozal MJ. 2014. Prevalence of WHO transmitted drug resistance mutations by deep sequencing in antiretroviral-naive subjects in Hunan Province, China. PLoS One 9:e98740.
28. Palmer S, Kearney M, Maldarelli F, Halvas EK, Bixby CJ, Bazmi H, Rock D, Falloon J, Davey RT, Jr., Dewar RL, Metcalf JA, Hammer S, Mellors JW, Coffin JM. 2005. Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. J Clin Microbiol 43:406-413.
29. Johnson JA, Li JF, Wei X, Lipscomb J, Irlbeck D, Craig C, Smith A, Bennett DE, Monsour M, Sandstrom P, Lanier ER, Heneine W. 2008. Minority HIV-1 drug resistance mutations are present in antiretroviral treatment-naive populations and associate with reduced treatment efficacy. PLoS Med 5:e158.
30. Svarovskaia ES, Margot NA, Bae AS, Waters JM, Goodman D, Zhong L, Borroto-Esoda K, Miller MD. 2007. Low-level K65R mutation in HIV-1 reverse transcriptase of treatment-experienced patients exposed to abacavir or didanosine. J Acquir Immune Defic Syndr 46:174-180.
31. Palmer S, Boltz V, Maldarelli F, Kearney M, Halvas EK, Rock D, Falloon J, Davey RT, Jr., Dewar RL, Metcalf JA, Mellors JW, Coffin JM. 2006. Selection and persistence of non-nucleoside reverse transcriptase inhibitor-resistant HIV-1 in patients starting and stopping non-nucleoside therapy. AIDS 20:701-710.

32. Paredes R, Lalama CM, Ribaudo HJ, Schackman BR, Shikuma C, Giguel F, Meyer WA, 3rd, Johnson VA, Fiscus SA, D'Aquila RT, Gulick RM, Kuritzkes DR. 2010. Pre-existing minority drug-resistant HIV-1 variants, adherence, and risk of antiretroviral treatment failure. J Infect Dis 201:662-671.
33. Li JZ, Paredes R, Ribaudo HJ, Svarovskaia ES, Metzner KJ, Kozal MJ, Hullsiek KH, Balduin M, Jakobsen MR, Geretti AM, Thiebaut R, Ostergaard L, Masquelier B, Johnson JA, Miller MD, Kuritzkes DR. 2011. Low-frequency HIV-1 drug resistance mutations and risk of NNRTI-based antiretroviral treatment failure: a systematic review and pooled analysis. JAMA 305:1327-1335.
34. Palmer S, Boltz V, Martinson N, Maldarelli F, Gray G, McIntyre J, Mellors J, Morris L, Coffin J. 2006. Persistence of nevirapine-resistant HIV-1 in women after single-dose nevirapine therapy for prevention of maternal-to-fetal HIV-1 transmission. Proc Natl Acad Sci U S A 103:7094-7099.
35. Jourdain G, Ngo-Giang-Huong N, Le Coeur S, Bowonwatanuwong C, Kantipong P, Leechanachai P, Ariyadej S, Leenasirimakul P, Hammer S, Lallemant M. 2004. Intrapartum exposure to nevirapine and subsequent maternal responses to nevirapine-based antiretroviral therapy. N Engl J Med 351:229-240.
36. Kapoor A, Jones M, Shafer RW, Rhee SY, Kazanjian P, Delwart EL. 2004. Sequencing-based detection of low-frequency human immunodeficiency virus type 1 drug-resistant mutants by an RNA/DNA heteroduplex generator-tracking assay. J Virol 78:7112-7123.
37. Lecossier D, Shulman NS, Morand-Joubert L, Shafer RW, Joly V, Zolopa AR, Clavel F, Hance AJ. 2005. Detection of minority populations of HIV-1 expressing the K103N resistance mutation in patients failing nevirapine. J Acquir Immune Defic Syndr 38:37-42.
38. Pillay K. 2015. The impact of the p7/p1 cleavage site mutations on replication capacity and drug resistance in HIV-1 subtype C. Masters of Science. University o Kwa-Zulu Natal, South Africa.
39. Gatanaga H, Suzuki Y, Tsang H, Yoshimura K, Kavlick MF, Nagashima K, Gorelick RJ, Mardy S, Tang C, Summers MF, Mitsuya H. 2002. Amino acid substitutions in Gag protein at non-cleavage sites are indispensable for the development of a high multitude of HIV-1 resistance against Protease inhibitors. J Biol Chem 277:5952-5961.
40. Fun A, van Maarseveen NM, Pokorna J, Maas RE, Schipper PJ, Konvalinka J, Nijhuis M. 2011. HIV-1 Protease inhibitor mutations affect the development of HIV-1 resistance to the maturation inhibitor bevirimat. Retrovirology 8:70.
41. Sutherland KA, Parry CM, McCormick A, Kapaata A, Lyagoba F, Kaleebu P, Gilks CF, Goodall R, Spyer M, Kityo C, Pillay D, Gupta RK, Group DV. 2015. Evidence for

Reduced Drug Susceptibility without Emergence of Major Protease Mutations following Protease Inhibitor Monotherapy Failure in the SARA Trial. PLoS One 10:e0137834.

42. Kolli M, Ozen A, Kurt-Yilmaz N, Schiffer CA. 2014. HIV-1 Protease-substrate coevolution in nelfinavir resistance. J Virol 88:7145-7154.
43. Ghosn J, Delaugerre C, Flandre P, Galimand J, Cohen-Codar I, Raffi F, Delfraissy JF, Rouzioux C, Chaix ML. 2011. Polymorphism in Gag gene cleavage sites of HIV-1 non-B subtype and virological outcome of a first-line lopinavir/ritonavir single drug regimen. PLoS One 6:e24798.
44. Bally F, Martinez R, Peters S, Sudre P, Telenti A. 2000. Polymorphism of HIV type 1 Gag p7/p1 and p1/p6 cleavage sites: clinical significance and implications for resistance to Protease inhibitors. AIDS Res Hum Retroviruses 16:1209-1213.
45. Prado JG, Wrin T, Beauchaine J, Ruiz L, Petropoulos CJ, Frost SD, Clotet B, D'Aquila RT, Martinez-Picado J. 2002. Amprenavir-resistant HIV-1 exhibits lopinavir cross-resistance and reduced replication capacity. AIDS 16:1009-1017.
46. Hsieh SM, Pan SC, Chang SY, Hung CC, Sheng WH, Chen MY, Chang SC. 2013. Differential impact of resistance-associated mutations to Protease inhibitors and nonnucleoside reverse transcriptase inhibitors on HIV-1 replication capacity. AIDS Res Hum Retroviruses 29:1117-1122.
47. Nicastri E, Sarmati L, d'Ettorre G, Palmisano L, Parisi SG, Uccella I, Rianda A, Concia E, Vullo V, Vella S, Andreoni M. 2003. Replication capacity, biological phenotype, and drug resistance of HIV strains isolated from patients failing antiretroviral therapy. J Med Virol 69:1-6.
48. Iyidogan P, Anderson KS. 2014. Current perspectives on HIV-1 antiretroviral drug resistance. Viruses 6:4095-4139.
49. Dauber DS, Ziermann R, Parkin N, Maly DJ, Mahrus S, Harris JL, Ellman JA, Petropoulos C, Craik CS. 2002. Altered substrate specificity of drug-resistant human immunodeficiency virus type 1 Protease. J Virol 76:1359-1368.
50. Malim MH, Emerman M. 2001. HIV-1 sequence variation: drift, shift, and attenuation. Cell 104:469-472.
51. Lee MC, Deng J, Briggs JM, Duan Y. 2005. Large-scale conformational dynamics of the HIV-1 integrase core domain and its catalytic loop mutants. Biophys J 88:3133-3146.
52. Prabu-Jeyabalan M, Nalivaika E, Schiffer CA. 2002. Substrate shape determines specificity of recognition for HIV-1 Protease: analysis of crystal structures of six substrate complexes. Structure 10:369-381.
53. Nijhuis M, Wensing AM, Bierman WF, de Jong D, Kagan R, Fun A, Jaspers CA, Schurink KA, van Agtmael MA, Boucher CA. 2009. Failure of treatment with first-line

lopinavir boosted with ritonavir can be explained by novel resistance pathways with Protease mutation 76V. J Infect Dis 200:698-709.

54. Aoki M, Venzon DJ, Koh Y, Aoki-Ogata H, Miyakawa T, Yoshimura K, Maeda K, Mitsuya H. 2009. Non-cleavage site Gag mutations in amprenavir-resistant human immunodeficiency virus type 1 (HIV-1) predispose HIV-1 to rapid acquisition of amprenavir resistance but delay development of resistance to other Protease inhibitors. J Virol 83:3059-3068.

55. Myint L, Matsuda M, Matsuda Z, Yokomaku Y, Chiba T, Okano A, Yamada K, Sugiura W. 2004. Gag non-cleavage site mutations contribute to full recovery of viral fitness in Protease inhibitor-resistant human immunodeficiency virus type 1. Antimicrob Agents Chemother 48:444-452.

56. Rabi SA, Laird GM, Durand CM, Laskey S, Shan L, Bailey JR, Chioma S, Moore RD, Siliciano RF. 2013. Multi-step inhibition explains HIV-1 Protease inhibitor pharmacodynamics and resistance. J Clin Invest 123:3848-3860.

57. Parboosing R, Naidoo A, Gordon M, Taylor M, Vella V. 2011. Resistance to antiretroviral drugs in newly diagnosed, young treatment-naive HIV-positive pregnant women in the province of KwaZulu-Natal, South Africa. J Med Virol 83:1508-1513.

58. Wright JK, Brumme ZL, Carlson JM, Heckerman D, Kadie CM, Brumme CJ, Wang B, Losina E, Miura T, Chonco F, van der Stok M, Mncube Z, Bishop K, Goulder PJ, Walker BD, Brockman MA, Ndung'u T. 2010. Gag-Protease-mediated replication capacity in HIV-1 subtype C chronic infection: associations with HLA type and clinical parameters. J Virol 84:10820-10831.

59. Nauwelaers D, Van Houtte M, Winters B, Steegen K, Van Baelan K, Chi E, Zhou M, Steiner D, Bonesteel R, Aston C, Stuyver L. 2011. A Synthetic HIV-1 Subtype C Backbone Generates Comparable PR and RT Resistance Profiles to a Subtype B Backbone in a Recombinant Virus Assay. Plos One 6(5): e19643.

7 APPENDIX

7.1 HIV-1 subtype C treatment naïve sequence accession numbers

HIV-1 subtype C treatment naïve sequences: Los Alamos accession numbers							
AY463217	DQ164107	DQ445634	HM593221	HM593341	HM593461	JF704305	JF704425
AY463218	DQ164108	DQ445635	HM593222	HM593342	HM593462	JF704306	JF704426
AY463219	DQ164109	DQ445637	HM593223	HM593343	HM593463	JF704307	JF704427
AY463220	DQ164110	EU347706	HM593224	HM593344	HM593464	JF704308	JF704428
AY463221	DQ164111	EU347707	HM593225	HM593345	HM593465	JF704309	JF704429
AY463222	DQ164112	EU347708	HM593226	HM593346	HM593466	JF704310	JF704430
AY463223	DQ164113	GQ999989	HM593227	HM593347	HM593467	JF704311	JF704431
AY463224	DQ164114	HM593106	HM593228	HM593348	HM593468	JF704312	JF704432
AY463225	DQ164115	HM593107	HM593229	HM593349	HM593469	JF704313	JF704433
AY463226	DQ164116	HM593108	HM593230	HM593350	HM593470	JF704314	JF704434
AY463227	DQ164117	HM593109	HM593231	HM593351	HM593471	JF704315	JF704435
AY463228	DQ164118	HM593110	HM593232	HM593352	HM593472	JF704316	JF704436
AY463229	DQ164119	HM593111	HM593233	HM593353	HM593473	JF704317	JF704437
AY463230	DQ164121	HM593112	HM593234	HM593354	HM593474	JF704318	JF704438
AY463231	DQ164122	HM593113	HM593235	HM593355	HM593475	JF704319	JF704439
AY463232	DQ164126	HM593114	HM593236	HM593356	HM593476	JF704320	JF704440
AY463233	DQ164127	HM593115	HM593237	HM593357	HM593477	JF704321	JF704441
AY463234	DQ164129	HM593116	HM593238	HM593358	HM593478	JF704322	JF704442
AY463236	DQ275642	HM593117	HM593239	HM593359	HM593479	JF704323	JF704443
AY463237	DQ275643	HM593118	HM593240	HM593360	HM593480	JF704324	JF704444
AY703908	DQ275644	HM593119	HM593241	HM593361	HM593481	JF704325	JF704445
AY703909	DQ275645	HM593120	HM593242	HM593362	HM593482	JF704326	JF704446
AY703910	DQ275646	HM593121	HM593243	HM593363	HM593483	JF704327	JF704447
AY703911	DQ275647	HM593122	HM593244	HM593364	HM593484	JF704328	JF704448
AY772690	DQ275648	HM593123	HM593245	HM593365	HM593485	JF704329	JF704449
AY772691	DQ275649	HM593124	HM593246	HM593366	HM593486	JF704330	JF704450
AY772692	DQ275650	HM593125	HM593247	HM593367	HM593487	JF704331	JF704451
AY772693	DQ275651	HM593126	HM593248	HM593368	HM593488	JF704332	JF704452

HIV-1 subtype C treatment naïve sequences: Los Alamos accession numbers							
AY772694	DQ275652	HM593127	HM593249	HM593369	HM593489	JF704333	JF704453
AY772695	DQ275653	HM593128	HM593250	HM593370	HM593490	JF704334	JF704454
AY772696	DQ275654	HM593129	HM593251	HM593371	HM593491	JF704335	JF704455
AY772698	DQ275655	HM593130	HM593252	HM593372	HM593492	JF704336	JF704456
AY772699	DQ275656	HM593131	HM593253	HM593373	HM593493	JF704337	JF704457
AY772700	DQ275657	HM593132	HM593254	HM593374	HM593494	JF704338	JF704458
AY878054	DQ275658	HM593133	HM593255	HM593375	HM593495	JF704339	JF704459
AY878055	DQ275659	HM593134	HM593256	HM593376	HM593496	JF704340	JF704460
AY878056	DQ275660	HM593135	HM593257	HM593377	HM593497	JF704341	JF704461
AY878057	DQ275661	HM593136	HM593258	HM593378	HM593498	JF704342	JF704462
AY878058	DQ275664	HM593137	HM593259	HM593379	HM593499	JF704343	JF704464
AY878059	DQ351216	HM593138	HM593260	HM593380	HM593500	JF704344	JF704465
AY878060	DQ351217	HM593139	HM593261	HM593381	HM593501	JF704345	JF704466
AY878061	DQ351218	HM593140	HM593262	HM593382	HM593502	JF704346	JF704467
AY878062	DQ351219	HM593141	HM593263	HM593383	HM593503	JF704347	JF704468
AY878063	DQ351220	HM593142	HM593264	HM593384	HM593504	JF704348	JF704469
AY878064	DQ351221	HM593143	HM593265	HM593385	HM593505	JF704349	JF704470
AY878065	DQ351222	HM593144	HM593266	HM593386	HM593506	JF704350	JF704471
AY878068	DQ351223	HM593145	HM593267	HM593387	HM593507	JF704351	JF704472
AY878070	DQ351224	HM593146	HM593268	HM593388	HM593508	JF704352	JF704473
AY878071	DQ351225	HM593147	HM593269	HM593389	HM593509	JF704353	JF704474
AY878072	DQ351226	HM593148	HM593270	HM593390	HM593510	JF704354	JF704475
AY901965	DQ351227	HM593149	HM593271	HM593391	JF704235	JF704355	JF704476
AY901966	DQ351228	HM593150	HM593272	HM593392	JF704236	JF704356	JF704477
AY901967	DQ351229	HM593152	HM593273	HM593393	JF704237	JF704357	JF704478
AY901968	DQ351230	HM593153	HM593274	HM593394	JF704238	JF704358	JF704479
AY901969	DQ351231	HM593154	HM593275	HM593395	JF704239	JF704359	JF704480
AY901970	DQ351232	HM593155	HM593276	HM593396	JF704240	JF704360	JF704481
AY901971	DQ351233	HM593156	HM593277	HM593397	JF704241	JF704361	JF704482
AY901972	DQ351234	HM593157	HM593278	HM593398	JF704242	JF704362	JF704483
AY901973	DQ351235	HM593158	HM593279	HM593399	JF704243	JF704363	JF704484
AY901974	DQ351237	HM593159	HM593280	HM593400	JF704244	JF704364	JF704485

HIV-1 subtype C treatment naïve sequences: Los Alamos accession numbers							
AY901975	DQ369976	HM593160	HM593281	HM593401	JF704245	JF704365	JF704486
AY901976	DQ369977	HM593161	HM593282	HM593402	JF704246	JF704366	JF704487
AY901977	DQ369978	HM593162	HM593283	HM593403	JF704247	JF704367	JF704488
AY901978	DQ369979	HM593163	HM593284	HM593404	JF704248	JF704368	JF704489
AY901979	DQ369980	HM593164	HM593285	HM593405	JF704249	JF704369	JF704490
AY901980	DQ369981	HM593165	HM593286	HM593406	JF704250	JF704370	JF704492
AY901981	DQ369982	HM593166	HM593287	HM593407	JF704251	JF704371	JF704494
DQ011165	DQ369983	HM593167	HM593288	HM593408	JF704252	JF704372	JF704496
DQ011166	DQ369984	HM593168	HM593289	HM593409	JF704253	JF704373	JF704497
DQ011167	DQ369985	HM593169	HM593290	HM593410	JF704254	JF704374	JF704498
DQ011169	DQ369986	HM593170	HM593291	HM593411	JF704255	JF704375	JF704499
DQ011170	DQ369987	HM593171	HM593292	HM593412	JF704256	JF704376	JF704500
DQ011171	DQ369988	HM593172	HM593293	HM593413	JF704257	JF704377	JF704501
DQ011172	DQ369989	HM593173	HM593294	HM593414	JF704258	JF704378	JF704502
DQ011173	DQ369990	HM593174	HM593295	HM593415	JF704259	JF704379	JF704503
DQ011174	DQ369991	HM593175	HM593296	HM593416	JF704260	JF704380	JF704504
DQ011175	DQ369992	HM593176	HM593297	HM593417	JF704261	JF704381	JF704505
DQ011176	DQ369993	HM593177	HM593298	HM593418	JF704262	JF704382	JF704506
DQ011177	DQ369994	HM593178	HM593299	HM593419	JF704263	JF704383	JF704507
DQ011178	DQ369995	HM593179	HM593300	HM593420	JF704264	JF704384	JF704508
DQ011179	DQ369996	HM593180	HM593301	HM593421	JF704265	JF704385	JF704509
DQ011180	DQ369997	HM593181	HM593302	HM593422	JF704266	JF704386	JF704510
DQ056404	DQ396364	HM593182	HM593303	HM593423	JF704267	JF704387	JF704511
DQ056405	DQ396365	HM593183	HM593304	HM593424	JF704268	JF704388	JF704512
DQ056406	DQ396366	HM593184	HM593305	HM593425	JF704269	JF704389	JF704513
DQ056408	DQ396367	HM593185	HM593306	HM593426	JF704270	JF704390	JF704514
DQ056409	DQ396368	HM593186	HM593307	HM593427	JF704271	JF704391	JF704515
DQ056410	DQ396369	HM593187	HM593308	HM593428	JF704272	JF704392	JF704516
DQ056411	DQ396370	HM593188	HM593309	HM593429	JF704273	JF704393	JF704517
DQ056412	DQ396371	HM593189	HM593310	HM593430	JF704274	JF704394	JF704518
DQ056413	DQ396372	HM593190	HM593311	HM593431	JF704275	JF704395	JF704519
DQ056414	DQ396373	HM593191	HM593312	HM593432	JF704276	JF704396	JF704520

HIV-1 subtype C treatment naïve sequences: Los Alamos accession numbers							
DQ056415	DQ396374	HM593192	HM593313	HM593433	JF704277	JF704397	JF704521
DQ056416	DQ396375	HM593193	HM593314	HM593434	JF704278	JF704398	JF704522
DQ056417	DQ396376	HM593194	HM593315	HM593435	JF704279	JF704399	JF704523
DQ056418	DQ396377	HM593195	HM593316	HM593436	JF704280	JF704400	JF704524
DQ093585	DQ396378	HM593196	HM593317	HM593437	JF704281	JF704401	JF704525
DQ093586	DQ396379	HM593197	HM593318	HM593438	JF704282	JF704402	JF704526
DQ093587	DQ396380	HM593199	HM593319	HM593439	JF704283	JF704403	JF704527
DQ093588	DQ396381	HM593200	HM593320	HM593440	JF704284	JF704404	JF704528
DQ093589	DQ396382	HM593201	HM593321	HM593441	JF704285	JF704405	JF704529
DQ093590	DQ396383	HM593202	HM593322	HM593442	JF704286	JF704406	JF704530
DQ093591	DQ396384	HM593203	HM593323	HM593443	JF704287	JF704407	JF704531
DQ093592	DQ396385	HM593204	HM593324	HM593444	JF704288	JF704408	JF704532
DQ093593	DQ396386	HM593205	HM593325	HM593445	JF704289	JF704409	JF704533
DQ093594	DQ396387	HM593206	HM593326	HM593446	JF704290	JF704410	JF704534
DQ093595	DQ396388	HM593207	HM593327	HM593447	JF704291	JF704411	JF704535
DQ093596	DQ396389	HM593208	HM593328	HM593448	JF704292	JF704412	JF704536
DQ093597	DQ396390	HM593209	HM593329	HM593449	JF704293	JF704413	JF704537
DQ093598	DQ396391	HM593210	HM593330	HM593450	JF704294	JF704414	JF704538
DQ093599	DQ396392	HM593211	HM593331	HM593451	JF704295	JF704415	JF704539
DQ093600	DQ396393	HM593212	HM593332	HM593452	JF704296	JF704416	JF704540
DQ093601	DQ396394	HM593213	HM593333	HM593453	JF704297	JF704417	JF704541
DQ093602	DQ396395	HM593214	HM593334	HM593454	JF704298	JF704418	JF704542
DQ093604	DQ396396	HM593215	HM593335	HM593455	JF704299	JF704419	
DQ093605	DQ396397	HM593216	HM593336	HM593456	JF704300	JF704420	
DQ093607	DQ396399	HM593217	HM593337	HM593457	JF704301	JF704421	
DQ164104	DQ445631	HM593218	HM593338	HM593458	JF704302	JF704422	
DQ164105	DQ445632	HM593219	HM593339	HM593459	JF704303	JF704423	
DQ164106	DQ445633	HM593220	HM593340	HM593460	JF704304	JF704424	

7.2 HIV-1 subtype B treatment naïve sequence accession numbers

HIV-1 subtype B treatment naïve sequence Los Alamos accession numbers							
AB428551	AB873242	AB873291	AB873340	AB873390	AB873440	AB873490	AB873539
AB428552	AB873243	AB873292	AB873341	AB873391	AB873441	AB873491	AB873540
AB428553	AB873244	AB873293	AB873342	AB873392	AB873442	AB873492	AB873541
AB428554	AB873245	AB873294	AB873343	AB873393	AB873443	AB873493	AB873542
AB428555	AB873246	AB873295	AB873344	AB873394	AB873444	AB873494	AB873543
AB428556	AB873247	AB873296	AB873345	AB873395	AB873445	AB873495	AB873544
AB428557	AB873248	AB873297	AB873346	AB873396	AB873446	AB873496	AB873545
AB428558	AB873249	AB873298	AB873347	AB873397	AB873447	AB873497	AB873546
AB428559	AB873250	AB873299	AB873348	AB873398	AB873448	AB873498	AB873547
AB428560	AB873251	AB873300	AB873349	AB873399	AB873449	AB873499	AB873548
AB428561	AB873252	AB873301	AB873350	AB873400	AB873451	AB873500	AB873549
AB428562	AB873253	AB873302	AB873351	AB873401	AB873452	AB873501	AB873550
AB873205	AB873254	AB873303	AB873352	AB873402	AB873453	AB873502	AB873551
AB873206	AB873255	AB873304	AB873353	AB873403	AB873454	AB873503	AB873552
AB873207	AB873256	AB873305	AB873354	AB873404	AB873455	AB873504	AB873553
AB873208	AB873257	AB873306	AB873355	AB873405	AB873456	AB873505	AB873554
AB873209	AB873258	AB873307	AB873356	AB873406	AB873457	AB873506	AB873555
AB873210	AB873259	AB873308	AB873357	AB873407	AB873458	AB873507	AB873556
AB873211	AB873260	AB873309	AB873358	AB873408	AB873459	AB873508	AB873557
AB873212	AB873261	AB873310	AB873359	AB873409	AB873460	AB873509	AB873558
AB873213	AB873262	AB873311	AB873360	AB873410	AB873461	AB873510	AB873559
AB873214	AB873263	AB873312	AB873361	AB873411	AB873462	AB873511	AB873560
AB873215	AB873264	AB873313	AB873362	AB873412	AB873463	AB873512	AB873561
AB873216	AB873265	AB873314	AB873363	AB873413	AB873464	AB873513	AB873563
AB873217	AB873266	AB873315	AB873364	AB873414	AB873465	AB873514	AB873564
AB873218	AB873267	AB873316	AB873365	AB873415	AB873466	AB873515	AB873565
AB873219	AB873268	AB873317	AB873366	AB873416	AB873467	AB873516	AB873566
AB873220	AB873269	AB873318	AB873367	AB873417	AB873468	AB873517	AB873567
AB873221	AB873270	AB873319	AB873368	AB873419	AB873469	AB873518	AB873568

HIV-1 subtype B treatment naïve sequence Los Alamos accession numbers							
AB873222	AB873271	AB873320	AB873369	AB873420	AB873470	AB873519	AB873569
AB873223	AB873272	AB873321	AB873370	AB873421	AB873471	AB873520	AB873570
AB873224	AB873273	AB873322	AB873371	AB873422	AB873472	AB873521	AB873571
AB873225	AB873274	AB873323	AB873372	AB873423	AB873473	AB873522	AB873572
AB873226	AB873275	AB873324	AB873373	AB873424	AB873474	AB873523	AB873573
AB873227	AB873276	AB873325	AB873374	AB873425	AB873475	AB873524	AB873574
AB873228	AB873277	AB873326	AB873375	AB873426	AB873476	AB873525	AB873575
AB873229	AB873278	AB873327	AB873376	AB873427	AB873477	AB873526	AB873576
AB873230	AB873279	AB873328	AB873377	AB873428	AB873478	AB873527	AB873577
AB873231	AB873280	AB873329	AB873378	AB873429	AB873479	AB873528	AB873578
AB873232	AB873281	AB873330	AB873379	AB873430	AB873480	AB873529	AB873579
AB873233	AB873282	AB873331	AB873381	AB873431	AB873481	AB873530	AB873580
AB873234	AB873283	AB873332	AB873382	AB873432	AB873482	AB873531	AB873581
AB873235	AB873284	AB873333	AB873383	AB873433	AB873483	AB873532	AB873582
AB873236	AB873285	AB873334	AB873384	AB873434	AB873484	AB873533	AB873583
AB873237	AB873286	AB873335	AB873385	AB873435	AB873485	AB873534	AB873584
AB873238	AB873287	AB873336	AB873386	AB873436	AB873486	AB873535	AB873585
AB873239	AB873288	AB873337	AB873387	AB873437	AB873487	AB873536	AB873586
AB873240	AB873289	AB873338	AB873388	AB873438	AB873488	AB873537	AB873587
AB873241	AB873290	AB873339	AB873389	AB873439	AB873489	AB873538	AB873588
AB873589	AY786818	AY786888	AY786937	EF121387	EU241975	EU242027	EU242078
AB873590	AY786819	AY786889	AY786938	EF121388	EU241976	EU242028	EU242079
AB873591	AY786830	AY786890	AY786939	EF121389	EU241977	EU242029	EU242080
AB873592	AY786831	AY786891	AY786952	EF121390	EU241978	EU242030	EU242081
AB873593	AY786832	AY786892	AY786962	EF121391	EU241979	EU242031	EU242082
AB873594	AY786833	AY786893	DQ127534	EF121392	EU241980	EU242032	EU242083
AB873595	AY786834	AY786894	DQ396398	EF121393	EU241981	EU242033	EU242085
AB873596	AY786835	AY786895	DQ487188	EF121394	EU241982	EU242034	EU242086
AB873597	AY786836	AY786896	DQ487190	EF175209	EU241983	EU242035	EU242087
AB873598	AY786837	AY786897	DQ487191	EF175210	EU241984	EU242036	EU242088
AB873599	AY786838	AY786898	DQ853436	EF175211	EU241985	EU242037	EU242089
AB873600	AY786839	AY786899	DQ853437	EF175212	EU241986	EU242038	EU242090

HIV-1 subtype B treatment naïve sequence Los Alamos accession numbers							
AB873601	AY786840	AY786900	DQ853438	EU241938	EU241987	EU242039	EU242091
AY314044	AY786841	AY786901	DQ853439	EU241939	EU241988	EU242041	EU242092
AY314045	AY786842	AY786902	DQ853440	EU241940	EU241989	EU242042	EU242093
AY314046	AY786843	AY786903	DQ853441	EU241941	EU241990	EU242043	EU242094
AY314047	AY786844	AY786904	DQ853442	EU241942	EU241991	EU242044	EU242095
AY314048	AY786845	AY786905	DQ853443	EU241943	EU241992	EU242045	EU242096
AY314049	AY786846	AY786906	DQ853444	EU241944	EU241993	EU242046	EU242097
AY314050	AY786848	AY786907	DQ853445	EU241945	EU241994	EU242047	EU242098
AY314052	AY786849	AY786908	DQ853446	EU241946	EU241995	EU242048	EU242099
AY786790	AY786850	AY786909	DQ853447	EU241947	EU241996	EU242049	EU242100
AY786791	AY786851	AY786910	DQ853448	EU241948	EU241997	EU242050	EU242101
AY786792	AY786852	AY786911	DQ853449	EU241949	EU241998	EU242051	EU242102
AY786793	AY786853	AY786912	DQ853450	EU241950	EU242000	EU242052	EU242103
AY786794	AY786854	AY786913	DQ853451	EU241951	EU242001	EU242053	EU242104
AY786795	AY786855	AY786914	DQ853452	EU241952	EU242002	EU242054	EU242105
AY786796	AY786856	AY786915	DQ853453	EU241953	EU242003	EU242055	EU242106
AY786797	AY786857	AY786916	DQ853454	EU241954	EU242004	EU242056	EU242107
AY786798	AY786858	AY786917	DQ853466	EU241955	EU242005	EU242057	EU242108
AY786799	AY786859	AY786918	DQ853467	EU241956	EU242006	EU242058	EU242109
AY786800	AY786870	AY786919	DQ853468	EU241957	EU242007	EU242059	EU242110
AY786801	AY786871	AY786920	DQ853469	EU241958	EU242008	EU242060	EU242111
AY786802	AY786872	AY786921	DQ853470	EU241959	EU242009	EU242061	EU242112
AY786803	AY786873	AY786922	DQ853471	EU241960	EU242010	EU242062	EU242113
AY786804	AY786874	AY786923	DQ853472	EU241961	EU242011	EU242063	EU242114
AY786805	AY786875	AY786924	DQ853473	EU241962	EU242012	EU242064	EU242115
AY786806	AY786876	AY786925	DQ853474	EU241963	EU242013	EU242065	EU242116
AY786807	AY786877	AY786926	DQ853475	EU241964	EU242014	EU242066	EU242117
AY786808	AY786878	AY786927	DQ996244	EU241965	EU242015	EU242067	EU242118
AY786809	AY786879	AY786928	DQ996245	EU241966	EU242017	EU242068	EU242120
AY786810	AY786880	AY786929	DQ996248	EU241967	EU242018	EU242069	EU242121
AY786811	AY786881	AY786930	DQ996249	EU241968	EU242019	EU242070	EU242122
AY786812	AY786882	AY786931	DQ996250	EU241969	EU242020	EU242071	EU242123

HIV-1 subtype B treatment naïve sequence Los Alamos accession numbers							
AY786813	AY786883	AY786932	DQ996252	EU241970	EU242021	EU242073	EU242124
AY786814	AY786884	AY786933	DQ996253	EU241971	EU242022	EU242074	EU242125
AY786815	AY786885	AY786934	DQ996254	EU241972	EU242023	EU242075	EU242126
AY786816	AY786886	AY786935	EF121385	EU241973	EU242025	EU242076	EU242127
AY786817	AY786887	AY786936	EF121386	EU241974	EU242026	EU242077	EU242128
EU242129	EU242179	EU242229	EU242278	EU242329	EU242379	EU242432	EU242483
EU242130	EU242180	EU242230	EU242279	EU242330	EU242380	EU242433	EU242484
EU242131	EU242181	EU242231	EU242280	EU242331	EU242382	EU242434	EU242485
EU242132	EU242182	EU242232	EU242281	EU242332	EU242383	EU242435	EU242486
EU242133	EU242183	EU242233	EU242282	EU242333	EU242384	EU242436	EU242487
EU242134	EU242184	EU242234	EU242283	EU242334	EU242385	EU242437	EU242488
EU242135	EU242185	EU242235	EU242284	EU242335	EU242386	EU242438	EU242489
EU242136	EU242186	EU242236	EU242285	EU242336	EU242387	EU242439	EU242491
EU242137	EU242187	EU242237	EU242286	EU242337	EU242388	EU242440	EU242492
EU242138	EU242188	EU242238	EU242287	EU242338	EU242389	EU242441	EU242493
EU242139	EU242189	EU242239	EU242288	EU242339	EU242390	EU242442	EU242494
EU242140	EU242190	EU242240	EU242289	EU242340	EU242391	EU242443	EU242495
EU242141	EU242191	EU242241	EU242290	EU242341	EU242392	EU242444	EU242496
EU242142	EU242192	EU242242	EU242291	EU242342	EU242393	EU242445	EU242497
EU242143	EU242193	EU242243	EU242292	EU242343	EU242394	EU242446	EU242498
EU242145	EU242194	EU242244	EU242293	EU242344	EU242395	EU242447	EU242499
EU242146	EU242195	EU242245	EU242295	EU242345	EU242396	EU242448	EU242500
EU242147	EU242196	EU242246	EU242296	EU242346	EU242397	EU242449	EU242501
EU242148	EU242197	EU242247	EU242297	EU242347	EU242398	EU242450	EU242502
EU242149	EU242198	EU242248	EU242298	EU242348	EU242399	EU242451	EU242503
EU242150	EU242199	EU242249	EU242299	EU242349	EU242400	EU242452	EU242504
EU242151	EU242200	EU242250	EU242300	EU242350	EU242401	EU242453	EU517762
EU242152	EU242201	EU242251	EU242301	EU242351	EU242402	EU242454	EU517763
EU242153	EU242203	EU242252	EU242302	EU242352	EU242403	EU242455	EU517764
EU242154	EU242204	EU242253	EU242303	EU242353	EU242404	EU242456	EU517765
EU242155	EU242205	EU242254	EU242304	EU242354	EU242405	EU242457	EU517766
EU242156	EU242206	EU242255	EU242305	EU242355	EU242406	EU242458	EU517767

HIV-1 subtype B treatment naïve sequence Los Alamos accession numbers							
EU242157	EU242207	EU242256	EU242306	EU242356	EU242407	EU242459	EU517768
EU242158	EU242208	EU242257	EU242307	EU242357	EU242408	EU242460	EU517769
EU242159	EU242209	EU242258	EU242308	EU242358	EU242410	EU242461	EU517770
EU242160	EU242210	EU242259	EU242309	EU242359	EU242412	EU242462	EU517771
EU242161	EU242211	EU242260	EU242310	EU242360	EU242413	EU242463	EU517772
EU242162	EU242212	EU242261	EU242311	EU242361	EU242414	EU242464	EU517773
EU242163	EU242213	EU242262	EU242312	EU242362	EU242416	EU242465	EU517774
EU242164	EU242214	EU242263	EU242313	EU242363	EU242417	EU242466	EU517775
EU242165	EU242215	EU242264	EU242314	EU242364	EU242418	EU242467	EU517776
EU242166	EU242216	EU242265	EU242315	EU242365	EU242419	EU242468	EU517777
EU242167	EU242217	EU242266	EU242317	EU242366	EU242420	EU242469	EU517778
EU242168	EU242218	EU242267	EU242318	EU242367	EU242421	EU242472	EU517779
EU242169	EU242219	EU242268	EU242319	EU242368	EU242422	EU242473	EU517780
EU242170	EU242220	EU242269	EU242320	EU242370	EU242423	EU242474	EU517781
EU242171	EU242221	EU242270	EU242321	EU242371	EU242424	EU242475	EU517782
EU242172	EU242222	EU242271	EU242322	EU242372	EU242425	EU242476	EU517783
EU242173	EU242223	EU242272	EU242323	EU242373	EU242426	EU242477	EU517784
EU242174	EU242224	EU242273	EU242324	EU242374	EU242427	EU242478	EU517785
EU242175	EU242225	EU242274	EU242325	EU242375	EU242428	EU242479	EU517786
EU242176	EU242226	EU242275	EU242326	EU242376	EU242429	EU242480	EU517787
EU242177	EU242227	EU242276	EU242327	EU242377	EU242430	EU242481	EU517788
EU242178	EU242228	EU242277	EU242328	EU242378	EU242431	EU242482	EU517790
EU242129	EU242179	EU242229	EU242278	EU242329	EU242379	EU242432	EU242483
EU242130	EU242180	EU242230	EU242279	EU242330	EU242380	EU242433	EU242484
EU242131	EU242181	EU242231	EU242280	EU242331	EU242382	EU242434	EU242485
EU242132	EU242182	EU242232	EU242281	EU242332	EU242383	EU242435	EU242486
EU242133	EU242183	EU242233	EU242282	EU242333	EU242384	EU242436	EU242487
EU242134	EU242184	EU242234	EU242283	EU242334	EU242385	EU242437	EU242488
EU242135	EU242185	EU242235	EU242284	EU242335	EU242386	EU242438	EU242489
EU242136	EU242186	EU242236	EU242285	EU242336	EU242387	EU242439	EU242491
EU242137	EU242187	EU242237	EU242286	EU242337	EU242388	EU242440	EU242492
EU242138	EU242188	EU242238	EU242287	EU242338	EU242389	EU242441	EU242493

HIV-1 subtype B treatment naïve sequence Los Alamos accession numbers							
EU242139	EU242189	EU242239	EU242288	EU242339	EU242390	EU242442	EU242494
EU242140	EU242190	EU242240	EU242289	EU242340	EU242391	EU242443	EU242495
EU242141	EU242191	EU242241	EU242290	EU242341	EU242392	EU242444	EU242496
EU242142	EU242192	EU242242	EU242291	EU242342	EU242393	EU242445	EU242497
EU242143	EU242193	EU242243	EU242292	EU242343	EU242394	EU242446	EU242498
EU242145	EU242194	EU242244	EU242293	EU242344	EU242395	EU242447	EU242499
EU242146	EU242195	EU242245	EU242295	EU242345	EU242396	EU242448	EU242500
EU242147	EU242196	EU242246	EU242296	EU242346	EU242397	EU242449	EU242501
EU242148	EU242197	EU242247	EU242297	EU242347	EU242398	EU242450	EU242502
EU242149	EU242198	EU242248	EU242298	EU242348	EU242399	EU242451	EU242503
EU242150	EU242199	EU242249	EU242299	EU242349	EU242400	EU242452	EU242504
EU242151	EU242200	EU242250	EU242300	EU242350	EU242401	EU242453	EU517762
EU242152	EU242201	EU242251	EU242301	EU242351	EU242402	EU242454	EU517763
EU242153	EU242203	EU242252	EU242302	EU242352	EU242403	EU242455	EU517764
EU242154	EU242204	EU242253	EU242303	EU242353	EU242404	EU242456	EU517765
EU242155	EU242205	EU242254	EU242304	EU242354	EU242405	EU242457	EU517766
EU242156	EU242206	EU242255	EU242305	EU242355	EU242406	EU242458	EU517767
EU242157	EU242207	EU242256	EU242306	EU242356	EU242407	EU242459	EU517768
EU242158	EU242208	EU242257	EU242307	EU242357	EU242408	EU242460	EU517769
EU242159	EU242209	EU242258	EU242308	EU242358	EU242410	EU242461	EU517770
EU242160	EU242210	EU242259	EU242309	EU242359	EU242412	EU242462	EU517771
EU242161	EU242211	EU242260	EU242310	EU242360	EU242413	EU242463	EU517772
EU242162	EU242212	EU242261	EU242311	EU242361	EU242414	EU242464	EU517773
EU242163	EU242213	EU242262	EU242312	EU242362	EU242416	EU242465	EU517774
EU242164	EU242214	EU242263	EU242313	EU242363	EU242417	EU242466	EU517775
EU242165	EU242215	EU242264	EU242314	EU242364	EU242418	EU242467	EU517776
EU242166	EU242216	EU242265	EU242315	EU242365	EU242419	EU242468	EU517777
EU242167	EU242217	EU242266	EU242317	EU242366	EU242420	EU242469	EU517778
EU242168	EU242218	EU242267	EU242318	EU242367	EU242421	EU242472	EU517779
EU242169	EU242219	EU242268	EU242319	EU242368	EU242422	EU242473	EU517780
EU242170	EU242220	EU242269	EU242320	EU242370	EU242423	EU242474	EU517781
EU242171	EU242221	EU242270	EU242321	EU242371	EU242424	EU242475	EU517782

HIV-1 subtype B treatment naïve sequence Los Alamos accession numbers							
EU242172	EU242222	EU242271	EU242322	EU242372	EU242425	EU242476	EU517783
EU242173	EU242223	EU242272	EU242323	EU242373	EU242426	EU242477	EU517784
EU242174	EU242224	EU242273	EU242324	EU242374	EU242427	EU242478	EU517785
EU242175	EU242225	EU242274	EU242325	EU242375	EU242428	EU242479	EU517786
EU242176	EU242226	EU242275	EU242326	EU242376	EU242429	EU242480	EU517787
EU242177	EU242227	EU242276	EU242327	EU242377	EU242430	EU242481	EU517788
EU242178	EU242228	EU242277	EU242328	EU242378	EU242431	EU242482	EU517790
EU517791	FJ155192	GQ371264	GQ371337	GQ371410	GQ371475	GQ371538	GQ371607
EU517792	FJ155195	GQ371265	GQ371338	GQ371411	GQ371477	GQ371539	GQ371608
EU517793	FJ155199	GQ371266	GQ371339	GQ371412	GQ371478	GQ371540	GQ371609
EU517794	FJ155200	GQ371267	GQ371340	GQ371413	GQ371480	GQ371543	GQ371610
EU517795	FJ155202	GQ371268	GQ371341	GQ371414	GQ371481	GQ371544	GQ371612
EU517796	FJ155205	GQ371271	GQ371344	GQ371415	GQ371482	GQ371545	GQ371615
EU517797	FJ155208	GQ371273	GQ371348	GQ371416	GQ371483	GQ371546	GQ371617
EU517798	FJ155209	GQ371274	GQ371349	GQ371417	GQ371484	GQ371547	GQ371619
EU517799	FJ155211	GQ371275	GQ371351	GQ371418	GQ371485	GQ371548	GQ371624
EU517800	FJ670525	GQ371277	GQ371352	GQ371420	GQ371486	GQ371552	GQ371626
EU517801	FJ670531	GQ371278	GQ371354	GQ371421	GQ371488	GQ371553	GQ371627
EU517802	FJ853620	GQ371279	GQ371355	GQ371422	GQ371489	GQ371554	GQ371628
EU517803	FJ853622	GQ371280	GQ371356	GQ371423	GQ371490	GQ371555	GQ371629
EU517804	GQ371216	GQ371281	GQ371357	GQ371425	GQ371491	GQ371556	GQ371630
EU517805	GQ371217	GQ371282	GQ371359	GQ371427	GQ371492	GQ371558	GQ371632
EU517806	GQ371218	GQ371284	GQ371361	GQ371428	GQ371493	GQ371559	GQ371633
EU517807	GQ371219	GQ371286	GQ371363	GQ371429	GQ371495	GQ371561	GQ371634
EU517808	GQ371221	GQ371287	GQ371364	GQ371430	GQ371497	GQ371562	GQ371635
EU517809	GQ371222	GQ371289	GQ371367	GQ371431	GQ371498	GQ371563	GQ371636
EU517810	GQ371223	GQ371290	GQ371368	GQ371432	GQ371499	GQ371565	GQ371637
EU517811	GQ371224	GQ371291	GQ371369	GQ371433	GQ371500	GQ371566	GQ371640
EU517812	GQ371225	GQ371293	GQ371370	GQ371434	GQ371501	GQ371567	GQ371641
EU517813	GQ371226	GQ371294	GQ371373	GQ371435	GQ371502	GQ371569	GQ371645
EU517814	GQ371228	GQ371296	GQ371374	GQ371436	GQ371503	GQ371570	GQ371646
EU786679	GQ371229	GQ371297	GQ371376	GQ371437	GQ371504	GQ371571	GQ371647

HIV-1 subtype B treatment naïve sequence Los Alamos accession numbers							
EU786680	GQ371234	GQ371298	GQ371378	GQ371438	GQ371505	GQ371574	GQ371648
FJ155079	GQ371235	GQ371302	GQ371379	GQ371439	GQ371506	GQ371577	GQ371649
FJ155087	GQ371236	GQ371303	GQ371380	GQ371440	GQ371507	GQ371578	GQ371651
FJ155102	GQ371237	GQ371306	GQ371381	GQ371441	GQ371508	GQ371579	GQ371653
FJ155114	GQ371239	GQ371307	GQ371382	GQ371442	GQ371510	GQ371580	GQ371655
FJ155118	GQ371241	GQ371308	GQ371383	GQ371443	GQ371511	GQ371581	GQ371656
FJ155120	GQ371242	GQ371309	GQ371384	GQ371444	GQ371512	GQ371583	GQ371658
FJ155126	GQ371244	GQ371312	GQ371385	GQ371449	GQ371513	GQ371584	GQ371660
FJ155131	GQ371245	GQ371313	GQ371386	GQ371450	GQ371514	GQ371585	GQ371662
FJ155132	GQ371246	GQ371314	GQ371388	GQ371451	GQ371515	GQ371587	GQ371663
FJ155136	GQ371248	GQ371315	GQ371389	GQ371452	GQ371516	GQ371588	GQ371664
FJ155137	GQ371249	GQ371316	GQ371390	GQ371453	GQ371517	GQ371589	GQ371665
FJ155142	GQ371250	GQ371318	GQ371391	GQ371454	GQ371520	GQ371591	GQ371666
FJ155144	GQ371251	GQ371321	GQ371393	GQ371455	GQ371521	GQ371592	GQ371667
FJ155147	GQ371253	GQ371323	GQ371396	GQ371457	GQ371522	GQ371593	GQ371668
FJ155149	GQ371254	GQ371324	GQ371397	GQ371458	GQ371523	GQ371594	GQ371669
FJ155150	GQ371256	GQ371325	GQ371399	GQ371459	GQ371524	GQ371596	GQ371670
FJ155152	GQ371257	GQ371327	GQ371400	GQ371460	GQ371525	GQ371598	GQ371673
FJ155169	GQ371258	GQ371328	GQ371401	GQ371462	GQ371526	GQ371599	GQ371674
FJ155173	GQ371259	GQ371331	GQ371402	GQ371463	GQ371529	GQ371600	GQ371675
FJ155175	GQ371260	GQ371332	GQ371403	GQ371466	GQ371532	GQ371601	GQ371676
FJ155185	GQ371261	GQ371333	GQ371406	GQ371467	GQ371533	GQ371602	GQ371679
FJ155186	GQ371262	GQ371334	GQ371407	GQ371468	GQ371535	GQ371605	GQ371681
FJ155190	GQ371263	GQ371335	GQ371409	GQ371472	GQ371537	GQ371606	GQ371682
GQ371683	GQ371754	JN024106	JN024163	JN024213	JN024268	JN024318	JN024368
GQ371685	GQ371755	JN024108	JN024164	JN024214	JN024269	JN024319	JN024369
GQ371686	GQ371756	JN024109	JN024165	JN024215	JN024270	JN024320	JN024370
GQ371688	GQ371758	JN024110	JN024166	JN024216	JN024271	JN024321	JN024372
GQ371689	GQ371759	JN024111	JN024167	JN024217	JN024272	JN024322	JN024373
GQ371690	GQ371760	JN024112	JN024168	JN024218	JN024273	JN024323	JN024374
GQ371693	GQ371761	JN024113	JN024169	JN024219	JN024274	JN024324	JN024375
GQ371694	GQ371762	JN024114	JN024170	JN024220	JN024275	JN024325	JN024376

HIV-1 subtype B treatment naïve sequence Los Alamos accession numbers							
GQ371695	GQ371763	JN024115	JN024171	JN024222	JN024276	JN024326	JN024377
GQ371696	GQ372988	JN024116	JN024172	JN024224	JN024277	JN024327	JN024378
GQ371697	GQ372990	JN024117	JN024173	JN024226	JN024278	JN024328	JN024379
GQ371698	GU331318	JN024118	JN024174	JN024227	JN024279	JN024329	JN024380
GQ371699	GU331321	JN024119	JN024175	JN024228	JN024280	JN024330	JN024381
GQ371700	GU331326	JN024121	JN024176	JN024229	JN024281	JN024331	JN024383
GQ371703	GU331347	JN024122	JN024177	JN024231	JN024282	JN024332	JN024384
GQ371704	GU331353	JN024123	JN024179	JN024232	JN024284	JN024333	JN024385
GQ371706	GU331360	JN024124	JN024180	JN024234	JN024285	JN024334	JN024386
GQ371708	GU331367	JN024125	JN024181	JN024235	JN024286	JN024335	JN024387
GQ371709	GU331374	JN024126	JN024182	JN024236	JN024287	JN024336	JN024388
GQ371710	GU331375	JN024127	JN024183	JN024237	JN024288	JN024337	JN024389
GQ371711	GU331379	JN024128	JN024184	JN024238	JN024289	JN024338	JN024390
GQ371712	GU331402	JN024130	JN024185	JN024239	JN024290	JN024339	JN024391
GQ371713	GU331411	JN024131	JN024186	JN024241	JN024291	JN024340	JN024392
GQ371715	GU362881	JN024132	JN024187	JN024242	JN024292	JN024341	JN024393
GQ371716	GU362886	JN024133	JN024188	JN024243	JN024293	JN024342	JN024395
GQ371718	HM586187	JN024134	JN024189	JN024244	JN024294	JN024343	JN024396
GQ371719	HM586191	JN024135	JN024190	JN024245	JN024295	JN024344	JN024397
GQ371721	HM586194	JN024136	JN024191	JN024246	JN024296	JN024345	JN024398
GQ371722	HM586196	JN024137	JN024192	JN024247	JN024297	JN024346	JN024399
GQ371723	HM586198	JN024138	JN024193	JN024248	JN024298	JN024347	JN024400
GQ371724	HM586200	JN024139	JN024194	JN024249	JN024299	JN024348	JN024401
GQ371725	HM586202	JN024140	JN024195	JN024250	JN024300	JN024349	JN024402
GQ371726	HM586204	JN024141	JN024196	JN024251	JN024301	JN024350	JN024403
GQ371729	HM586206	JN024142	JN024197	JN024252	JN024302	JN024351	JN024404
GQ371730	HM586207	JN024143	JN024198	JN024253	JN024303	JN024352	JN024405
GQ371731	HM586208	JN024144	JN024199	JN024254	JN024304	JN024353	JN024406
GQ371732	HM586209	JN024145	JN024200	JN024255	JN024305	JN024354	JN024407
GQ371733	HM586210	JN024146	JN024201	JN024256	JN024306	JN024355	JN024408
GQ371734	HM586211	JN024147	JN024202	JN024257	JN024307	JN024356	JN024409
GQ371736	HM586212	JN024148	JN024203	JN024258	JN024308	JN024357	JN024410

HIV-1 subtype B treatment naïve sequence Los Alamos accession numbers							
GQ371737	HQ215554	JN024150	JN024204	JN024259	JN024309	JN024358	JN024411
GQ371740	HQ215556	JN024152	JN024205	JN024260	JN024310	JN024359	JN024412
GQ371742	HQ215568	JN024154	JN024206	JN024261	JN024311	JN024360	JN024413
GQ371744	HQ215577	JN024155	JN024207	JN024262	JN024312	JN024361	JN024414
GQ371745	JN024100	JN024157	JN024208	JN024263	JN024313	JN024362	JN024415
GQ371747	JN024101	JN024158	JN024209	JN024264	JN024314	JN024363	JN024416
GQ371751	JN024103	JN024159	JN024210	JN024265	JN024315	JN024364	JN024418
GQ371752	JN024104	JN024160	JN024211	JN024266	JN024316	JN024366	JN024419
GQ371753	JN024105	JN024162	JN024212	JN024267	JN024317	JN024367	JN024420
JN024421	JN024471	JN024524	JN400483	JX264249	JX264303	JX264354	JX264406
JN024422	JN024472	JN024525	JN400490	JX264251	JX264304	JX264355	JX264407
JN024423	JN024473	JN024526	JN400491	JX264252	JX264305	JX264356	JX264408
JN024424	JN024474	JN024527	JN408075	JX264253	JX264306	JX264357	JX264409
JN024425	JN024475	JN024528	JN408076	JX264254	JX264307	JX264358	JX264410
JN024426	JN024477	JN024529	JN408077	JX264255	JX264309	JX264359	JX446795
JN024427	JN024478	JN024530	JN704057	JX264256	JX264310	JX264360	JX446796
JN024428	JN024479	JN024531	JN860769	JX264257	JX264311	JX264361	JX446797
JN024429	JN024481	JN024532	JQ302321	JX264258	JX264312	JX264362	JX446798
JN024430	JN024483	JN024533	JQ302402	JX264259	JX264313	JX264363	JX446799
JN024431	JN024484	JN024534	JQ302403	JX264261	JX264314	JX264364	JX446800
JN024432	JN024485	JN024535	JQ302440	JX264262	JX264315	JX264365	JX446801
JN024434	JN024486	JN024536	JQ900848	JX264263	JX264316	JX264366	JX446802
JN024435	JN024487	JN024537	JQ900850	JX264264	JX264317	JX264367	JX446803
JN024436	JN024488	JN024538	JQ900857	JX264265	JX264318	JX264368	JX446804
JN024437	JN024489	JN024539	JQ900861	JX264266	JX264319	JX264369	JX446805
U69588							

7.3 Ethics clearance certificate

UNIVERSITY OF
KWAZULU-NATAL
INYUVESI
YAKWAZULU-NATALI

RESEARCH OFFICE
BIOMEDICAL RESEARCH ETHICS ADMINISTRATION
Westville Campus
Govan Mbeki Building
Private Bag X 54001
Durban
4000
KwaZulu-Natal, SOUTH AFRICA
Tel: 27 31 2604769 - Fax: 27 31 260-4609
Email: [redacted]
Website: http://research.ukzn.ac.za/Research/Ethics/BiomedicalResearchEthics.aspx

28 August 2014

Ms Urisha Singh
719 Umbilo Road
Doris Duke Medical Research Institute
Level 1, Nelson R Mandela School of Medicine
Congella
Durban
4013
Urisha.singh@gmail.com

Dear Ms Singh

PROTOCOL: Acquired and Transmitted Drug Resistance in HIV-1 Subtype C: Implications of Mutations on Viral Replication, Cleavage and Drug Susceptibility REF: BE347/13

RECERTIFICATION APPLICATION APPROVAL NOTICE

Approved:	07 October 2014
Expiration of Ethical Approval:	06 October 2015

I wish to advise you that your application for Recertification received on 01 August 2014 for the above protocol has been noted and approved by a sub-committee of the Biomedical Research Ethics Committee (BREC) for another approval period. The start and end dates of this period are indicated above.

If any modifications or adverse events occur in the project before your next scheduled review, you must submit them to BREC for review. Except in emergency situations, no change to the protocol may be implemented until you have received written BREC approval for the change.

The approval will be ratified by a full Committee at a meeting to be held on **09 September 2014.**

Yours sincerely

Mrs A Marimuthu
Senior Administrator: Biomedical Research Ethics