Investigation of the HIV diversity in the Cape Winelands, Overberg and West Coast districts of the Western Cape Province of South Africa.

Mikasi Sello Given

Thesis submitted in partial fulfilment of the requirements of the degree of Masters of Sciences in Medical Sciences (Medical Virology) at the Faculty of Medicine and Health Sciences, Stellenbosch University

Promoter

Dr. Graeme Brendon Jacobs

Co-Promoter

Associate Professor Susan Engelbrecht

March 2017

Table of Contents

Declarations

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

March 2017

Summary

The Western Cape Province of South Africa has a well-established program that monitors active combination antiretroviral therapy (cART) against HIV-1. The HIV-1 prevalence rate in the Province has increased from 5.0% in 2011 to 18.0% in 2015. South Africa has the highest rate of infections worldwide (19.2%). In this study, we analyzed the Protease (PR), Reverse Transcriptase (RT) and Intergrase (IN) regions of HIV-1 for diversity and resistance-associated mutations (RAMs) from samples obtained from the Cape Winelands, West Coast and Overberg districts of the Province, where no such study has ever been conducted. Samples were received from our diagnostic laboratory for HIV-1 viral load testing, through the National Health Laboratory Services (NHLS). Two hundred and five (205) patient samples with a viral load of 2000 copies/ml and above were included, based on Gall et al., (2012) who showed that a sensitivity of at least 2000 copies/ml is a limit of amplification for the SuperScsript ® III one-step RT with Platinum Taq DNA Polymerase kit, used in this study. We screened for HIV-1 diversity and RAMs using the pol PR, RT and IN regions with a laboratory-based PCR and sequencing protocol. Sequence-specific subtype analyses were executed with the REGA HIV subtyping tool 3.0, Recombinant Identification Program (RIP) 3.0 and subtype classification using evolutionary algorithms (SCUEAL) software. Sequences were screened for RAMs using the Stanford University HIV Drug Resistance Database (HIVdb) 8.1. We successfully PCR amplified 170 (82.9%) PR and 166 (80.9%) RT fragments. For the IN region, only 176 samples had sufficient plasma and RNA left after genotyping of the PR and RT regions. For IN we successfully amplified 143 (81.3%) of the patient samples. A total of 197 (96.1%) samples could be amplified for at least one of the pol regions. Of these, 62 (53.4%) PR, 103 (62.0%) RT and 93 (86.1%) IN sequences were obtained, respectively. We could successfully sequence 173 (84.4%) of the samples included. HIV-1 subtype C was predominant (n = 144; 93.7%), with 5.3% of other subtypes detected. This includes A1 (n = 2; 1.3%), B (n = 4; 2.6%), D (n = 1; 0.7%) and H (n =1; 0.7%). No major RAMs were detected against PI and IN inhibitors. Minor RAMs were detected in 4 PR (3.7%) and 15 IN (16.1%) sequences analysed. RAMs against RT inhibitors were detected in 63 (61.7%) of the sequences analyzed. This includes 39 NRTI mutations (36.1%) and 71 NNRTI mutations (63.5%) identified. As the national cART program continues to expand, HIV-1 diversity, viral load monitoring and drug resistance screening remains critical for the success of cART outcomes and reducing transmission rates. Our results reflect that subtype C is still the driving force of the epidemic in South Africa. However, we cannot ignore the potential impact of non-C subtypes. Sequence analyses confirm that the majority of patients

receiving viral load testing have major RAMs against RT inhibitors used in first line therapy. Better surveillance systems for HIV diversity and drug resistance testing are required to ensure success of cART.

Opsomming

Die Wes-Kaap Provinsie van Suid-Afrika het 'n goed gevestigde program wat aktief kombinasie antiretrovirale terapie (cART) teen MIV-1 monitor. Die MIV-1 voorkomssyfer in die Provinsie het toegeneem van 5,0% in 2011 tot 18,0% in 2015. Suid-Afrika het die hoogste koers van infeksies wêreldwyd (19,2%). In hierdie studie het ons die Protease (PR), Trutranskriptase (RT) en Intergrase (IN) streke van MIV-1 vir diversiteit en weerstand geassosieerde mutasies (RAMS) ontleed. Die monsters is verky vanuit die Kaapse Wynland, Weskus en Overberg distrikte van die Provinsie, waar geen sodanige studie ooit gedoen is nie. Monsters wat getoets was vir MIV-1 viruslading is van ons diagnostiese laboratorium seksie deur die NHLS verkry. 205 pasiënt monsters met 'n virale lading van 2000 kopieë / ml of meer is in ons studie ingesluit. Ons kriteria is gebaseer op Gall *et al*., (2012) wat getoon het dat 'n sensitiwiteit van ten minste 2000 kopieë / ml 'n beperking is vir amplifisering met die SuperScsript ® III een-stap RT met Platinum Taq DNA-polimerase kit, wat in hierdie studie gebruik is. Ons ondersoeke is gedoen met ’n laboratorium-gebaseerde PCR en DNS volgorde bepalings protokol. MIV-1 subtipe analises is uitgevoer met die volgende aanlyn sagteware: REGA 3.0; RIP 3.0 en SCUEAL. Die Stanford Universiteit se MIV weerstand aanlyn databasis (HIVdb) 8.1 is gebruik om die voorkomssyfer van RAMs te bepaal. Uit 'n versameling van 205 monsters het ons suksesvol 170 (82.9%) PR en 166 (80.9%) RT fragmente geamplifiseer. Vir die IN gebied was slegs 176 monsters oor met genoeg plasma en RNA vir verdere toetse. Vir die IN gebied kon ons 143 (81.3%) monsters suksesvol amplifiseer. Ons kon vir 197 (96.1%) van die monsters ten minste een van die *pol* streke amplifiseer. Hiervan is 62 (53.4%) PR, 103 (62.0%) RT en 93 IN (86.1%) DNS volgordes onderskeidelik verkry. Ons kon suksesvol die MIV-1 DNS volgorde bepaal van ten minste een pol gebied van 173 (84.4%) van ons totale monsters. MIV-1 subtipe C was oorheersend (n = 144; 93,7%), met 5.3% wat as ander subtipes geklasifiseer is. Dit sluit in A1 (n = 2; 1.3%), B (n = 4; 2.6%), D (n = 1; 0.7%) en H (N = 1; 0.7%). Daar was geen groot RAMs teen PR en IN middels nie. Kleiner RAMs is identifiseer in 4 (3.7%) PR en 15 (16.1%) volgordes wat ontleed is. RAMs teen RT-middels is opgespoor in 63 (61,7%) van die volgordes ontleed. Dit sluit in 39 NRTI mutasies (36,1%) en 71 NNRTI mutasies (63,5%) geïdentifiseer. Soos die nasionale cART program uitgebrei word, word MIV-1 diversiteit, viruslading en weerstandstoetse al hoe belangriker, veral om die sukses van cART te bepaal. Ons resultate toon dat subtipe C steeds die dryfkrag van die epidemie in Suid-Afrika is. Ons kan egter nie

die potensiële impak van nie-C subtipes ignoreer nie. Volgorde analises bevestig dat die meerderheid van die pasiënte wat virale lading toetse ondergaan weerstand het teen RT-middels wat in die eerste lyn behandeling gebruik word. Beter toesig stelsels is nodig om optimale sukses cART te verseker.

Acknowledgements

I would like to extend my sincere acknowledgment and thanks to the following individuals and institutions for their contributions and support towards the completion of this study:

Dr. Graeme Brendon Jacobs, my promoter for giving me the opportunity and platform to learn and develop as a young scientist under his superior supervision and guidance.

Professor Susan Engelbrecht, my co-promoter for her advice and troubleshooting the problems that I encountered on a daily basis of my research work.

Mr. Njenda Duncan, our research assistant, for teaching me laboratory techniques and skills, as well as giving me advice and encouragement throughout the course of my M.Sc. project.

Mr. Emmanuel Obasa, Dr. Lerato Sikhosana and Mr. Josiah Gichana, for their encouragement and support towards my study.

Ms. Cynthia Tamandjou, Ms. Karmista Poovan, Ms .Olivete Varathan and Mrs. Danelle van Jaarsveldt for their support, helping me with my thesis draft corrections and comments until the submission of the complete thesis version. My thesis is the reflection of their support.

I am thankful for the Poliomyelitis Research Foundation (PRF), National Research Fund (NRF), NRF-DAAD, Stellenbosch University Merit Bursary, Harry Crossley, National Health Laboratory Services (NHLS) Research Trust and Stellenbosch University research fund for their financial assistance

To all the staff and Colleague in the Division of Medical Virology, thank you for your encouragement and moral support throughout the duration of my master's degree.

I wish to thank my wife, Joy and my family for their constant support and encouragement throughout my academic career.

Finally, I wish to thank God for making everything possible throughout my studies in God I trust.

List of Scientific conference presentations and research visit

Conference presentations

1. **Sello Given Mikasi**, Susan Engelbrecht, Graeme Brendon Jacobs. HIV-1 resistance analyses of the Cape Winelands and Overberg districts, South Africa. Virology Africa. Cape Town, South Africa, 31 November – 3 December 2015 (**Poster presentation**).
2. **Sello Given Mikasi**, Susan Engelbrecht, Graeme Brendon Jacobs. HIV-1 resistance analyses of the Cape Winelands districts, South Africa. Pathology research day, 9 June 2016 (**Oral presentation**).
3. **Sello Given Mikasi**, Susan Engelbrecht, Graeme Brendon Jacobs. HIV-1 resistance analyses of the Cape Winelands districts, South Africa. Stellenbosch university 60th annual academic day.18 August 2016 (**Poster presentation**).
4. **Sello Given Mikasi**, Susan Engelbrecht, Graeme Brendon Jacobs. HIV-1 resistance analyses of the Cape Winelands districts, South Africa. Retrovirology conference 2016, Germany (Erlangen). 12-14 September 2016 (**Poster presentation**).

Research visit

Deutscher Akademischer Austauschdienst (DAAD) short-term Research scholarship visit.

1. Research visit to the institute of virology and immunobiology at the University of Wuerzburg, Germany. September to November 2016.

List of abbreviations

®	Registered
^{o}c	Degrees celcius
3TC	Lamivudine
AA	Amino acid
ABC	Abacavir
AIDS	Acquired Immuno Deficiency Syndrome
APOBEC3G	Apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3G
ART	Antiretroviral therapy
AZT	Zidovudine
Blast	Basic Local Alignment Search Tool
bp	Base Pair
bPI	boosted protease inhibitors
CA	Capsid
cART	combination ART
CCR5	Chemokine core-receptors 5
cDNA	Complementary deoxynucleic acid
CNS	Central nervous system
COMET	Context-based Modeling for Expeditious Typing
CPR	Calibrated Population Resistance
CRFs	Circulating recombinant forms
CTLs	Cytotoxic T-lymphocytes
CXCR4	Chemokine X core-receptors
d4T	Stavudine
DC	Dendritic cells
ddI	Didanosine
ddNTPs	Dideoxyribo-nucleoside triphospahte
DRC	Democratic Republic of Congo
DTG	Dolutegravir
EFV	Efavirenz
Env	Envelope
ETV	Etravirine
EVG	Elvitegravir
FDA	Food and Drug Administration
FTC	Emtricitabine
Gag	Group-specific antigen
gp	Glyco protein
HAART	Highly active Antiretroviral therapy
HBD	Human beta defensins
HHV 8	Human Herpesvirus 8
HIV	Human Immunodeficiency virus
HIVDB	HIV drug resistance database
HLA	Human Leukocyte Antigen
HREC	Human Research Ethics Committee
HTLV	Human T-lymphotropic virus

IDUs	Intravenous drugs users
IN	Integrase
InSTIs	Integrase strand transfer inhibitors
kb	Kilo base pair
KZN	Kwazulu-Natal
LANL	Los Alamos National Laboratory HIV Sequence Database
LAS	Lymphadenopathy syndrome
LAV	Lymphadenopathy virus
LTNP	Long –term non-progressers
LTR	Long terminal repeats
MA	Matrix
MEIA	Microparticle enzyme immunoassay
MHC	Major Histocompatability Complex
MRC	Medical Research Council
mRNA	messenger RNA
MSM	Men having sex with men
NC	Nucleocapsid
Nef	Negative factor
NHLS	National Health Laboratory Service
NNRTI	Non-nucleoside reverse transcriptase inhibitor
NRTIs	Nucleoside reverse transcriptase inhibitors
NTD	N-terminal domain
NVP	Nevarapine
PB	Phosphate buffer
PBMCs	Peripheral blood mononuclear cells
PCR	Polymerase chain reaction
PMCT	Prevent mother to child trans-mission
pol	Polymerase gene
PR	Protease
PrEP	Pre-exposure prophylaxis
RAL	Raltegravir
RAMs	Resistance-associated mutations
Rev	Regulator of expression of virion proteins
RIP	Recombinant Identification Program
RNA	Ribonucleic acid
RPV	Rilpivirine
RT	Reverse transcriptase
RTV	Ritonavir
SAMHD1	SAM domain and HD domain-containing protein 1
SCUEAL	Subtype classification using evolutionary algorithms
SDRM	Surveillance drug resistant mutations
SIVs	Simian immunodeficiency viruses
SLPI	Secretory leukocyte protease inhibitor
SU	Surface protein
Taq	Thermus aquaticus
TasP	Treatment as prevention

Tat	Transcriptional transactivator protein
TDF	Tenofovir
Tfl	Thermus flavus
TM	Transmembrane
TRIM	Tripartite motif
TRIM5α	Tripartite motif-containing 5α
UNAIDS	United Nation AIDS
URFs	Unique recombinant forms
USA	United States of America
Vif	Virion infectivity factor
Vpr	Viral protein R
Vpu	Viral protein U gene
WHO	World Health Organization
αH	alpha H
αI	alpha L

List of figures

List of tables

CHAPTER ONE

1. INTRODUCTION AND LITERATURE REVIEW

CHAPTER ONE

1. Introduction and literature review

1.1. Introduction

The human immunodeficiency virus (HIV) is the causative agent of the Acquired Immunodeficiency Syndrome (AIDS) epidemic we currently face (Hatziioannou *et al*., 2014, Sharp & Hahn, 2011). It is estimated that 36.7 million people worldwide are currently living with HIV/AIDS (UNAIDS 'Fact Sheet', 2016). Of these, approximately 19.0 million are living in Eastern and Southern Africa (Figure 1.1), about 1.1 million deaths were recorded in 2015 and 2.1 million people were newly infected (UNAIDS gap report, 2016). South Africa remains the region that is the most affected by HIV-1, with 7.0 million people living with the virus (UNAIDS gap report, 2016). In South Africa, the highest prevalence of HIV/AIDS is recorded in Kwazulu-Natal (KZN) Province (40.1%), with the lowest prevalence recorded in the Western Cape Province (18.0%) (NDOH, 2014). HIV prevalence in the Western Cape Province is significantly lower than the national estimates. However, there is a lack of data regarding the HIV-1 diversity and drug resistance, especially in remote communities found within the district of the Cape Winelands, the Westcoast and Overberg. The majority of research on HIV diversity and drug resistance testing in South Africa has been done in major centers, such as Cape Town (Western Cape), Johannesburg (Gauteng) and Durban (Kwazulu-Natal), where academics have easier access to clinics and patient samples.

In this study, we investigated the diversity of HIV-1 and drug resistance-associated mutations (RAMs) in cohorts from three remote districts of the Western Cape province in South Africa: This includes the West Coast with a prevalence of 9.6%, the Overberg region with a prevalence of 13.9% and Cape Winelands with a prevalence of 15.0%; which collectively accounts for 38.5% of HIV/AIDS prevalence cases in the Western Cape Province (NDOH, 2014). In the 1980s, HIV-1 subtype B and D viruses were the driving force of the South African epidemic (Sher *et al.,* 1989; Engelbrecht *et al*., 1995; Williamson *et al*., 1995). Currently, the majority of South African individuals are infected with HIV-1 (group M) subtype C (Gordon *et al.,* 2003; Jacobs *et al*., 2009). A relatively small number of Subtype A, D, CRF01_AE and other recombinant viruses have also been identified in the country (Hemelaar *et al*., 2006; Bredell *et al*., 2002; Papathanasopoulos *et al*., 2010). Among 7.0 million HIV-1 positive South Africans, only 3.4 million are currently receiving combination

antiretroviral therapy (cART). The new recommendations state to start cART immediately after being diagnosed with HIV-1, regardless of viral load or CD4 cell counts (WHO, 2015). The World Health Organization (WHO) cART guidelines recommend the use of a non-nucleoside reverse transcriptase inhibitor (NNRTI) and two nucleoside reverse transcriptase inhibitors (NRTIs) as the first-line cART regimen and two NRTIs, a ritonavir (RTV)-boosted protease inhibitors (bPI) as the second line therapy and Integrase strand transfer inhibitor (INSTs) as the third line therapy (WHO, 2016). One of the major drawbacks that hampers` the efficacy of Highly Active ART (HAART) is the emergence of drug resistance mutations conferred in the HIV-1 genome (Hu & Kuritzkes, 2014).

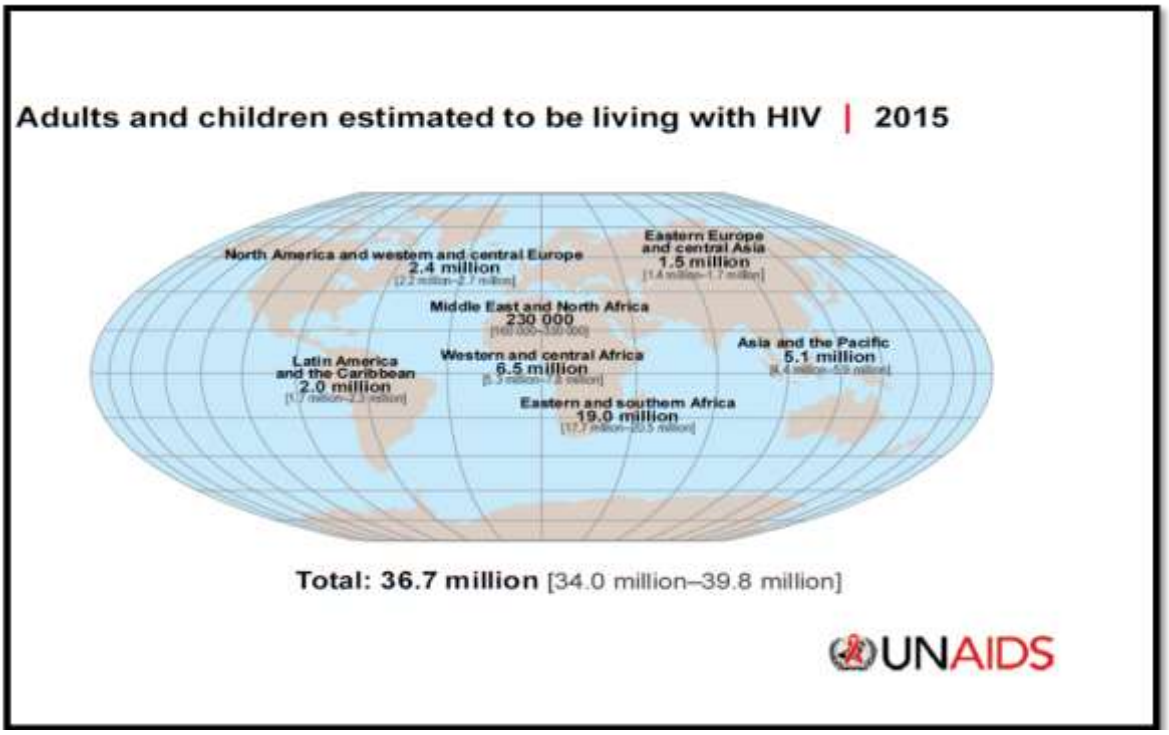


Figure 1.1: Estimated number of people living with HIV/AIDS at the end of 2015.

Sub-Saharan Africa, Asia and the Pacific are the geographical areas with the largest number of HIV-1 infections. The continents with the lowest number of infections are the Middle East and North Africa as well as in Eastern Europe and Central Asia (UNAIDS, 2015).

1.2 Literature Review

A brief literature review relevant to my study is given below. This includes the history of HIV, the origin of how HIV was introduced to the human population, the structural genome of HIV-1, HIV diversity, the HIV life cycle, the mechanisms by which cART inhibit the viral replication and the mechanisms that HIV-1 uses to reduce the susceptibility to cART.

1.2.1. History of HIV infection

By the early 1980s, AIDS was established as a new and distinct clinical entity among men having sex with men (MSM) in the United States of America (USA), when they presented with unusual opportunistic infections, such as pneumonia caused by *Pneumocystis carinii*, which is a rare disease causing lung infections in humans with compromised / downregulated immune system. In addition, the occurrence of a rare type of skin cancer caused by Human Herpesvirus 8 (HHV 8) referred to as Kaposi's Sarcoma was also observed in some of the homosexual men in the USA (Gottlieb *et al*., 1981; Friedman-Kien *et al*., 1981). AIDS cases were subsequently observed in other groups of patients presenting with similar symptoms, such as intravenous drugs users (IDUs), hemophiliacs and blood transfusion recipients (Barre-Sinoussi *et al*., 1983; Curran *et al*., 1984). In 1982 various published articles reported the emergence of the disease in infants who were born to mothers who were IDUs (MMWR – December, 1982b). Molecular epidemiology investigations in Africa reported similar clinical symptoms among homosexual men and the heterosexual population of Kinshasa, the capital city of the Democratic Republic of Congo (DRC) (Worobey *et al*., 2008). The first indication that a retrovirus is the causative agent of AIDS was discovered by Barré-Sinoussi and colleagues in 1983. The virus was isolated from a serum sample and a lymph node biopsy specimen collected from a homosexual man who presented with lymphadenopathy syndrome (LAS), a disease that affects the lymph node, leading to the name lymphadenopathy virus (LAV) (Barré-Sinoussi *et al*., 1983). In 1984, the virus was independently isolated by Levy and co-workers who called it AIDS-associated retrovirus (ARV) (Levy *et al*., 1984). The same virus was investigated by Robert Gallo and his colleagues who hypothesized that a variant of the human T-lymphotropic virus (HTLV) might be the causative agent of AIDS (Gallo *et al*., 1984). By 1985 the same virus had three names: LAV, ARV and HTLV-III, which all accounted for causing AIDS (Ratner *et al*., 1985a; 1985b). In order to prevent confusion the International Committee on the Taxonomy of viruses proposed to rename the AIDS causing virus Human Immunodeficiency Virus (HIV), as it is known today. (Coffin *et al*., 1986a; 1986b).

1.2.2. Origin of HIV

From a number of studies focusing on the origin of HIV, many revealed that HIV might have originated through cross-species infection (zoonotic transmission) from infected African primates to the human population, particularly in West and Central Africa (Apetrei *et al*., 2005; Nahmias *et al*., 1986). This was probably through hunting non-human primates as a

source of food (bushmeat), trade and keeping non-human primates as pets (Hahn *et al*., 2000; Sharp *et al*., 2010; Sharp *et al*., 2011). Phylogenetic analyses of HIV sequences clearly indicate that HIV is derived from simian immunodeficiency viruses (SIVs), commonly found in non-human primates. Chimpanzees, *Pan troglodytes troglodytes* was found to be the reservoir for HIV-1 (Gao *et al*., 1999; Hahn *et al*., 2000), while the sootey mangabey, *Cercocebus atys,* are the likely reservoir for HIV-2 (Gao *et al*., 1992; Hahn *et al*., 2000; Hirsch *et al*., 1989). More evidence that supports the spill-over of the virus into the human population is that the non-human primates that have been infected by SIVs are naturally resistant to the virus and this virus fails to induce any AIDS- like a disease in natural infected non-human primates (Paiardini *et al*., 2009*;* Pandrea *et al.,* 2008). This illustrates that the virus has evolved with their host for a long period of time and acquired the ability to adapt to a host that they infect or with whom they co-exist for commensalism (Rey-Cuille *et al*., 1998; Cichutek & Norley, 1993). Molecular clock and phylogenetic methodology dates the first HIV transmission into the human population back to the early 1930s with a ± 20 year confidence gap, before anyone becomes aware of the existence of the virus (Hahn *et al*., 2000; Korber *et al*., 2000; Lemey *et al*., 2003; Lemey *et al*., 2004; Worobey *et al*., 2008; Wertheim & Worobey, 2009; Silvestri, 2007).

1.2.3. HIV-1 genome structure

This thesis focusses on the PR, RT, and IN genes for further characterization of HIV-1 diversity. The *pol* gene is routinely sequenced in the clinical use of drug resistance testing context. HIV-1 subtypes can be determined based on the *pol* gene, as the fragment is long enough to determine HIV genotypes (Pasquier *et al*., 2001; Kessler *et al*., 2001; Yahi *et al*., 2001).

A structure of the HIV-1 genome is presented in Figure 1.2. HIV is a retrovirus, composed of double-stranded genomic RNA that is approximately 9kb in size (Zhuang *et al*., 2002). The virus encodes nine open reading frames of which three of these (*gag*, *pol* and *env*) are found in all retrovirus (Gallo *et al*., 1988). The structural components of the virion compose of six proteins that form the building block of the viral core and the outer membrane envelope. These are the four Gag proteins Capsid (CA), Matrix (MA), Nucleocapsid (NC) and p6, and the two Env proteins, Surface (SUsurface or gp120) and Transmembrane (TM or gp41). The polymerase gene (*pol)* encodes three of the major enzymatic components, Protease (PR), Reverse Transcriptase (RT) and Integrase (IN) that plays unique roles in other retroviruses.

HIV encodes at least six additional proteins that are not necessary for HIV replication, but do play a role in the viral replication cycle. These proteins are called the accessory proteins and regulatory proteins. Three of the accessory proteins, [Virion infectivity factor (Vif), Viral protein R (Vpr), and Negative factor (Nef)] are packed in the viral particle core and they play a role in increasing production of the HIV proteins. The *vif* gene increases the production of the HIV particle in the peripheral blood lymphocytes (Strebel *et al*., 1987). Vpr facilitates the infection of non-dividing cells by HIV (Heinzinger *et al*., 1994), while Nef plays a role in downmodulation of the CD4 and MHC class I (Schwartz *et al*., 1996). Two other regulatory proteins, Transcriptional transactivator protein (Tat) and Regulator of expression of virion proteins (Rev) are essential for regulating the production of HIV *in vitro* (Kim *et al*., 1989) and the last Viral protein U gene (Vpu), helps in the assembly of the virion indirectly (Estrabaud *et al*., 2007).

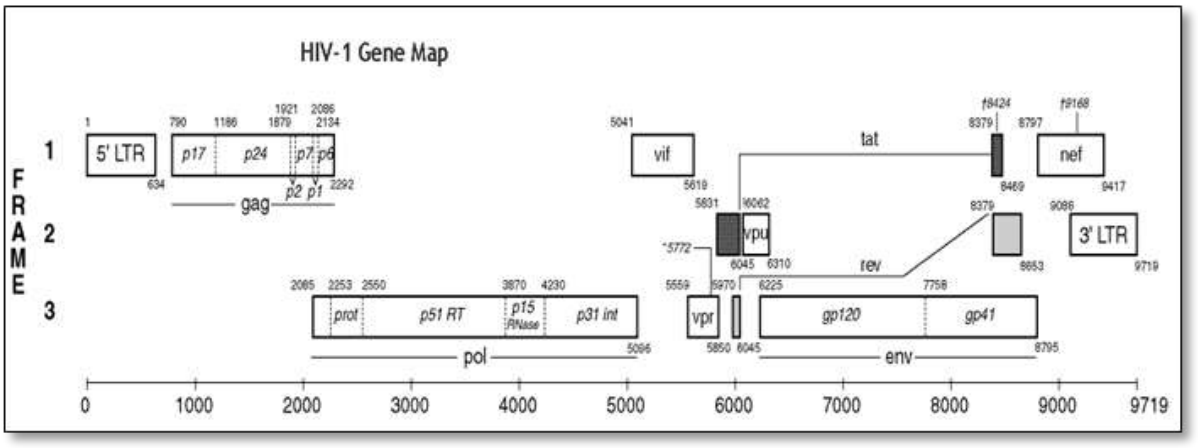


Figure 1.2: The HIV-1 genome structure.

The Structural genes (*gag*, *pol* and *env*) regulatory genes (*tat* and *rev*) and accessory genes (*nef*, *vif*, *vpr* and *vpu*) are indicated in Figure 1.2, as well as the proteins encoded by each genomic region (Source: www.hiv.lanl.gov).

1.2.4. HIV diversity

During this study we investigated the HIV-1 diversity of PR, RT and IN within our samples. Most of the therapeutic inhibitor drugs are designed to target these regions of the HIV-1 genome to inhibit HIV replication. They have proven to be efficient and successful to control viral replication to undetectable levels in the majority of HIV infected patients. The *pol* region is the most conserved part of the HIV-1 genome (Armstrong *et al*., 2009; Rhee *et al*., 2006). However, genetic diversity and HIV-1 resistance pose a major threat to the success of

therapy. There have been reports where RAMs are subtype- specific (Carr *et al.*, 2006; Taylor *et al.*, 2008; Grossman *et al.*, 2004). A study conducted by Grossman *et al.*, 2004, showed that the majority of the patients infected with subtype B viruses rarely develop V106→M mutations, whereas patients Infectected with subtype C develop RAMs against NNRTIs through either V106→M or K103→N mutations.

A schematic figure of HIV diversity is presented in Figure 1.3, while the global HIV subtype distribution is depicted in Figure 1.4. HIV diversity and its complexity in the worldwide epidemic is characterized by enormous genetic variation. Several factors contribute to the HIV genetic variation. This includes rapid replication of the virus, spontaneous turnover of HIV-1 *in vivo* (Roberts *et al.*, 1988), host selective pressure, the high mutation rate of the RT enzyme, which lacks a proofreading function and recombination of strains within infected individuals (Roberts *et al.*, 1988; Hodd *et al.*, 1995; Spira *et al.*, 2003; Rambau *et al.*, 2004; Perelson *et al.*, 1996). Insertions and deletions in the viral genome are also frequent occurrences (Korber *et al.*, 2000). Based on the phylogenetic analyses, HIV-1 strains can be classified into four main groups, namely M (major group), O (outlier), N (non-M non-O viruses) and the most recently P virus (Plantier *et al.*, 2009).

Group M viruses account for 52% of the global HIV-1 burden and can be further subdivided into discrete subtypes (A-D, F-H, J and K), circulating recombinant forms (CRFs) and many unique recombinant forms (URFs) (Hemelaar *et al.*, 2011). HIV-1 subtype B accounts for more than 95% of HIV-1 in the USA (Bennett, 2005). From 2004 to 2007, nearly one-half of all global infections were caused by subtype C (48%), followed by subtypes A (12%) and B (11%), CRF02_AG (8%), CRF01_AE (5%), subtype G (5%), and subtype D (2%) (Hemelaar *et al.*, 2011). Subtypes F, H, J and K cause less than 1% of infections worldwide. Other recombinant forms (CRFs and URFs) together cause 4% of global infections and combined with the burden of CRF02_AG and CRF01_AE infections, 20% of known HIV infections worldwide are currently caused by all recombinants (Hemelaar *et al.*, 2011).

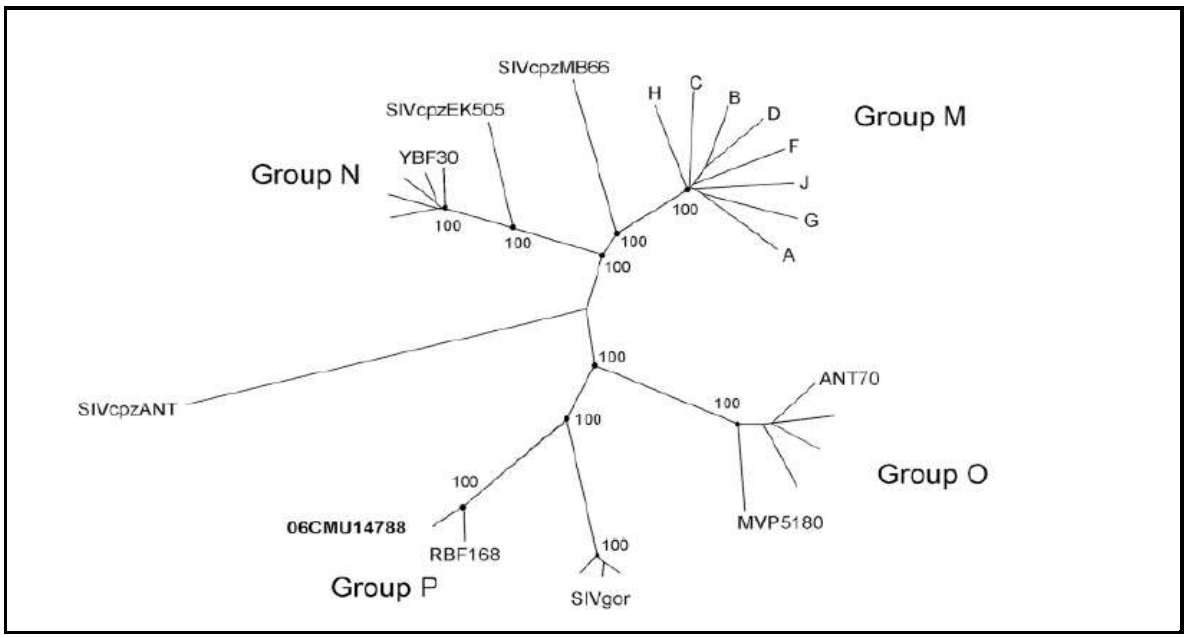


Figure 1.3: HIV-1 Phylogenetic tree derived from nucleotide alignment of genome sequences.

The different HIV-1 groups indicated, rooted with SIVcpzANT. The group M subtypes (A-D, F-H, and J) are shown, while reference sequences for groups N, O and P are also marked (Vallari *et al*., 2011).

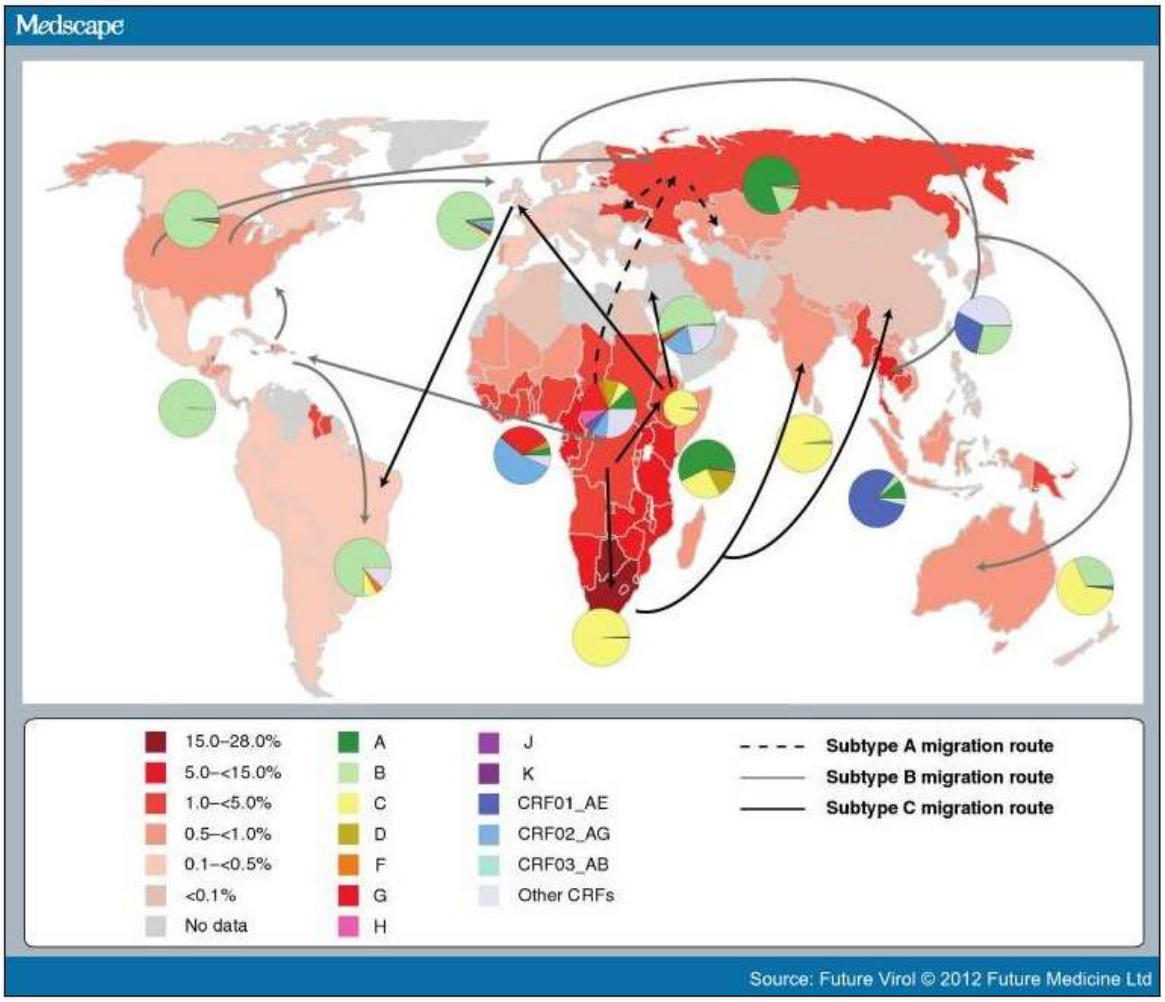


Figure 1.4: Global distribution of HIV-1 group M-subtypes and recombinant forms.

Countries are color-coded according to their last reported prevalence (Stack *et al*., 2010). Pie charts on Figure 1.4 represent the distribution of HIV-1 subtypes and recombinants over the globe. The colours representing the different HIV-1 subtypes and recombinants are indicated in the legend below of the figure.

1.2.4.1. HIV diversity in South Africa

By the late 1980s HIV-1 (group M) subtype C was identified among heterosexual individuals as the dominant subtype causing infection (Van Harmelen *et al.*, 1997). Since then, the South African epidemic has been dominated by HIV-1 subtype C (Hemelaar *et al*., 2011).

Today, the HIV epidemic in South Africa alone has affected more than 7.0 million people (UNAIDS gap report, 2016). There is an increasing number of HIV-1 recombinants and non-C subtypes strains being identified in South Africa (Hemelaar *et al*., 2006; Bredell *et al*., 2002; Papathanasopoulos *et al*., 2010; Jacobs *et al*., 2014). In the Westen Cape Province of South Africa approximately 1 - 10% HIV-1 infections has been ascribed to non-C strains (Engelbrecht *et al*., 1995; Van Harmelen *et al*., 1997; Bredell *et al*., 2002; Jacobs *et al*., 2009; Jacobs *et al*., 2014; Wilkinson *et al*., 2015). A study conducted by Jacobs *et al*., 2014 in our

laboratory identified circulating BC strains in at least 4.6% of their study population, followed by another recent study in our laboratory conducted by Wilkinson *et al*., 2015 that also identified other non-C subtypes: B, A, G, URF_AD and URF_AC circulating in the Western Cape (Wilkinson *et al*., 2015). The genetic diversity of HIV in several regions of south Africa have been studied (Williamson *et al.,* 1995; Hemelar *et al.,* 2006; Bredell *et al.,* 1998; Pillay *et al*., 2002). Thus far, no reports have been documented in the remote regions of the Western Cape Province. The role that HIV-1 diversity plays in the South African pandemic need to be investigated, because some study reported that HIV-1 diversity poses a major challenge on a wide spectrum of fields such as vaccine development and diagnostics (Buonaguro *et al.,* 2007; Peeters *et al*., 2003) and cART outcomes (Spira *et al*., 2003).

1.2.5. HIV life cycle

A summary of the HIV life cycle is illustrated in Figure 1.5. HIV mainly infects the white blood cells (leukocytes) of the human immune system. This is mainly CD4+ T-cells, Dendritic cells (DC), CD4+ monocytes, cytotoxic T-lymphocytes (CTLs) and macrophages (Stebbing *et al*., 2004). After the onset of infection the virus can be present in bodily fluids such as blood plasma, in peripheral blood mononuclear cells (PBMCs) as well as in primary and secondary organs of the immune system such as the lymph nodes and central nervous system (CNS) (Pierson *et al*., 2000; Stebbing *et al*., 2004).

HIV infection starts when the virus attaches its own glycoprotein (gp120) to the host CD4+ cell receptor and the chemokine core-receptors, either CCR5 and / or CXC4 of leukocytes and then penetrate the human host cell (Regoes & Bonhoeffer, 2005). The CCR5 and CXCR4 co-receptors are the main chemokine receptors that are used by HIV for entry *in vivo* (Clapham & McKnight, 2002). Following attachment, the viral lipid Envelope (Env) fuses with the cellular membrane and releases the HIV core, consisting of enzymes, proteins and genomic ribonucleic acid (RNA), into the host cytoplasm. Viral uncoating involves cellular factors and the viral proteins MA, Nef and Vif. (Hirsch & Curran, 1990; Harrich & Hooker, 2002). Once the HIV viral core is released into the cytoplasm, the RT enzyme, reverse-transcribes the viral RNA into a full-length double-stranded complementary deoxynucleic acid (cDNA). Once the proteins, enzymes and newly formed viral cDNA are transported to the nucleus of the host cell, the HIV IN incorporates the viral DNA into the human host cell genome (Craigie and Bushman, 2012). When the viral DNA is successfully integrated into the human genomic DNA, the provirus RNA can be transcribed with the help of Tat, which binds to the 5’ end long terminal repeats (LTR) region of the incorporated viral DNA, as the

host polymerase recognizes the integrated viral DNA as part of the host genomic DNA (Craigie and Bushman, 2012). The persistent of HIV within the reservoirs remains a major challenge for viral eradication. Thus, therapeutic interventions are urgently needed focusing on the strategies to eliminate the HIV reservoirs, because the persistence of latently infected cells in the reservoirs gives no hope of HIV cure, yet. (Siliciano *et al*., 2006).

Once the viral RNA is transcribed the viral messenger RNA (mRNA) is initially spliced and exported out of the nucleus, with the help of Rev, to the cytosol for translation with the help of cellular machinery (Briggs *et al*., 2003). Assembly of the HIV-1 viral occurs within the plasma membrane of the infected host cell. The HIV-1 *Gag* polyprotein play a major role in mediating all the necessary steps required for virion assembly (Ono & Freed, 2001). *Gag* oligomer polyproteins interact with the genomic DNA in the plasma membrane and multimerize further to form spherical immature capsids and prepares for packaging into new progeny virus. The newly formed virus leaves the host cell by budding, followed by the last step of virus maturation occurs with the viral proteins cleavage of the *Gag* and *Gag-Pol* polyprotein precursors to form the mature *Gag* and *Pol* proteins. The generated virus is free to infect other cells of the host cell (Roshal *et al*., 2001).

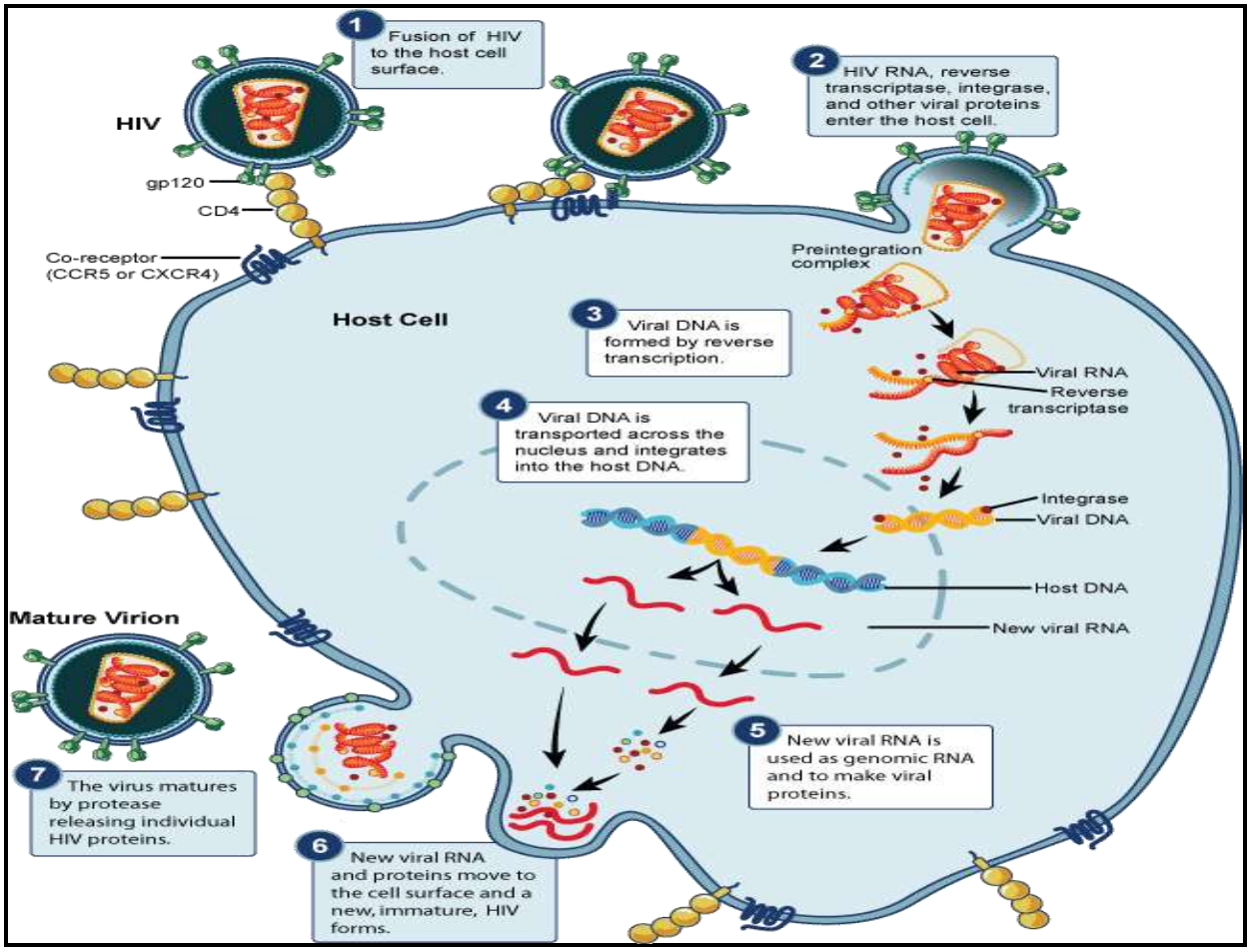


Figure 1.5: HIV life cycle.

The HIV life cycle steps consist of viral attachment, entry, reverse transcription and integration of the viral DNA into the host DNA. This is followed by export of the viral proteins as well as viral assembly, budding, and maturation of viral particles. The target sites of the three classes of inhibitors, reverse transcription, integration and maturation, are indicated (Pomerantz & Horn, 2003; Turner & Summers, 1999).

1.3. HIV-1 Enzymes of the *pol* gene

The HIV-1 *pol* gene encodes three viral enzymes namely PR, RT, and IN. Most of the drugs thus far were designed targeting the *pol* gene because this region was identified as the efficient target site for antiretroviral drugs to inhibit HIV replication. Our study focused on HIV diversity of the *pol* region because this region is commonly regarded as the gene that confers HIV-1 drug resistance, interfering with the efficacy of cART and most of the drugs targeting these regions have successfully proved to control the virus to undetectable level. The functions of these enzymes in the HIV-1 viral life cycles are explained in section 1.3.1 - 1.3.3 and the mechanisms of PI, RT and IN inhibitors are explained in section 1.4.3.

1.3.1.The Protease (PR) Enzyme

The HIV-1 Protease (PR) enzyme is a homodimer structure of two non-covalently identical monomers structures, which each is 99 amino acids long. A structural model of the HIV-1 PR homodimer is presented in Figure 1.6. The HIV-1 PR enzyme plays an important role in the life cycle of the retrovirus by post-translation processing of the precursor viral *Gag* and *Gag-Pol* polyproteins to yield essential mature viral structures and functional proteins (Oroszlan *et al*., 1990). Each monomer consists of a conserved triad, Aspartyl proteases on the active region site region 25-27 and forms the part of the catalytic site. The MA, CA, NC and p6 proteins are produced by the *Gag* polyprotein. The PR, RT and IN proteins are also produced from the *Gag* polyprotein when 9 different peptide sequences are recognised by the active site where it cleaves on the hydrophobic pockets (Erickson *et al*., 1999). Within each monomer, there is an extended β-sheet region (a glycine-rich loop) known as the flap that plays a role in the substrate-binding site by allowing the substrate to enter and bind to its binding site and closes down the active site upon substrate binding (James & Sielecki, 1985). Development of the drugs to fight against the HIV-1 life cycle have been designed to target the active site of the matured Protease enzyme (Vondrasek *et al*., 1997). The PR of subtype C is highly conserved and consists of different amino acidsequence in comparison with other subtypes A, B and D (de Oliveira *et al.*, 2003; Gordon *et al.*, 2003), which may potentially play a role in viral replication capacity and resistance acquisition (Kantor *et al.*, 2006; Ceccherini-Silberstein *et al*., 2009).

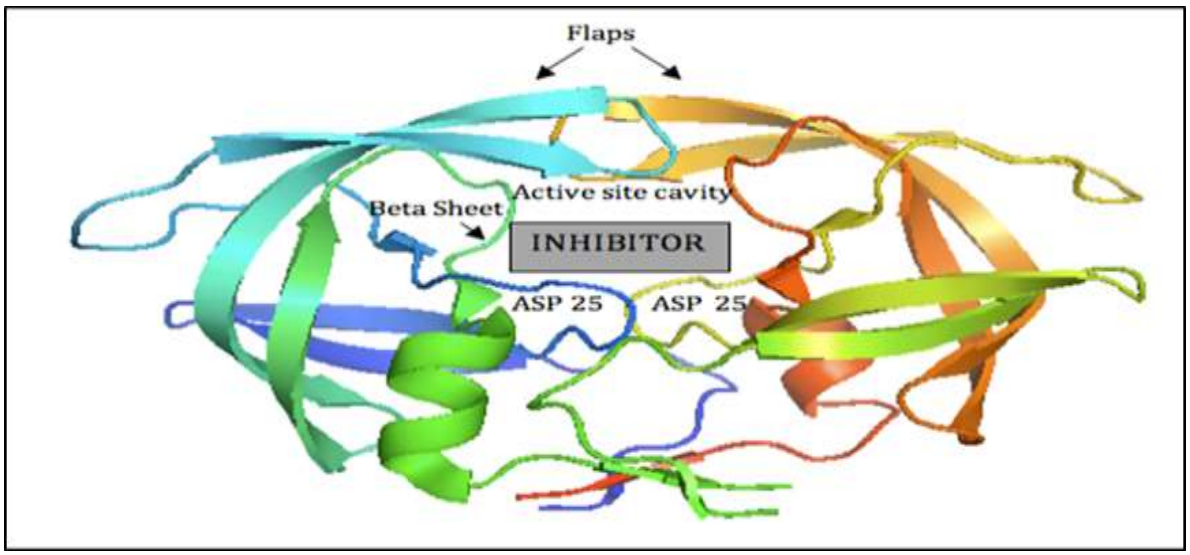


Figure 1.6: Structural model of HIV-1 PR.

Homodimer labelled with the protease inhibitor resistance mutations. The polypeptide backbone of both protease subunits (positions 1–99) is shown in Figure 1.6. The active site (positions 25–27) is also displaced. The Protease was co-crystallized with a protease inhibitor, which is displayed in space-fill mode.

1.3.2. The Reverse Transcriptase (RT) Enzyme

The RT enzyme contains three distinctive enzymatic activities namely; RNA template-dependent DNA polymerization, RNase H degradation of the RNA and DNA template-dependent DNA polymerization, enabling the synthesis of the linear double-stranded DNA. The latter is inserted into the human genome by the IN from a single-stranded viral RNA genome (Goff, 1990). HIV-1 RT exists as an asymmetric heterodimer consisting of two related Polypeptide subunits p66 and p51 (DiMarzo-Veronese *et al*., 1986). The p66 and p51 subunits are derived through cleavage by the viral PR, from a *Gag-Pol* polyprotein that is synthesized from a unspliced viral single-stranded RNA (di Marzo Veronese *et al*., 1986). The p51 subunit is composed of the first 450 amino acids of the RT gene. The p66 subunit consists of 560 amino acids in lengths and is encoded by RT gene, while the p51 subunits consists of the first 450 amino acids of the RT gene. Although p66 and p51 share the same 450 amino acids terminus, their relative organizations are significantly different. The longer subunit of the RT gene p66, contains the active site for carrying the enzymatic activities of the RT polymerase and RNase H and the DNA-binding "groove; the smaller p51 subunit carries no enzymatic activity and plays a role in the structural activity. The p66 is made of five subdomains namely: the "fingers", "palm", and "thumb" subdomains, playing a role in polymerization, and the "connection" and "RNase H" subdomains (Huang *et al*., 1998; Kohlstaedt *et al*., 1992). All of the five p66 subdomains mentioned form a nucleic-acid binding cleft, whereas the "connection" and the "thumb" of the small subunits form the floor of the binding cleft. The binding cleft is structured in a way that the nucleic acid touches both the polymerase and the RNase H active site. The "thumb" of the p66 subunit is made up of two helices: the alpha H (αH) and alpha L (αI) helices help to position the template strand to interact with the primer to the active site of the RT enzyme. The DNA "primer grip" is a structure that compromises of two hairpins (the p66 β12- β13) that help in positioning the 3-hydroxyl end of the primer terminus to the polymerase active site for the polymerization of the incoming nucleotide triphosphate (Jacobo-Molina *et al*., 1993, Xiong *et al*., 1990).

The HIV-1 RT is considered as the major molecular target of the several antiretroviral therapies against HIV-1 replication (Men´endez-Arias, 2002; Sarafianos *et al*., 2009; Men´endez-Arias, 2013). The RT sequence is relatively conserved region among HIV-1

subtypes, but differences exist between subtypes B, C and between other viral subtypes. This includes differences on the effect of virus replication, frequency and the location of polymorphisms and on the development of different resistant pathways mechanisms in response to treatment with the RT inhibitors (Armstrong *et al*., 2009; Rhee *et al*., 2006; Kantor *et al*., 2005; Grossman *et al*., 2001; Novitsky *et al*., 2007).

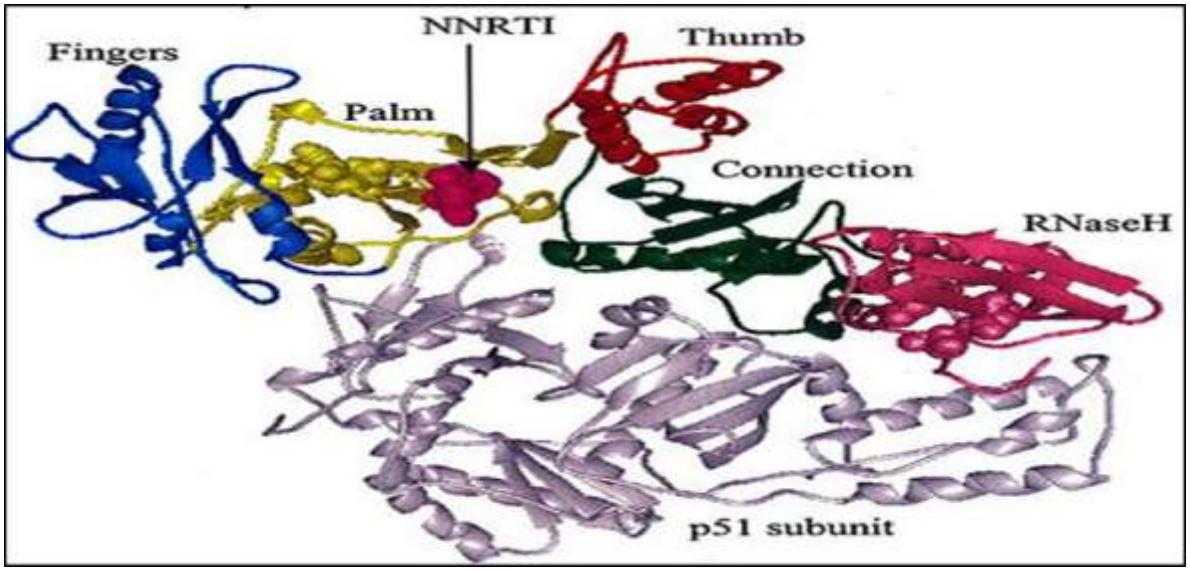


Figure 1.7: Structural model of HIV-1 (RT)

The polypeptide backbone of the complete p66 subunit (positions 1–560), and DNA primer and template strands are shown. This model is based on the structure provided by Kohlstaedt & Steitz, 1992 in which the RT is co-crystallized with nevirapine, which is displayed in space-fill mode. The positions associated with NNRTI resistance are shown surrounding the hydrophobic pocket to which nevirapine and other NNRTIs bind.

1.3.3. Integrase (IN) Enzyme

The HIV-1 IN is essential for the chromosomal integration of the newly synthesized double-stranded DNA into the host genomic DNA. The Intergrase enzyme is responsible for the insertion of the viral DNA into human genomic DNA using a two-step mechanism reaction. During the first mechanism, called 3' processing, two GT nucleotides bases are removed from each 3' end of the long terminal repeat of the viral double-stranded DNA synthesized by the RT enzyme. The second mechanism is called DNA strand transfer, whereby a pair of transesterification reactions integrate the 5' ends of the viral DNA into the backbone of the host genomic DNA (Craigie & Bushman, 2012). HIV-1 IN (32-kDa) is composed of 288-amino acid and is synthesised from the C-terminal portion of the *pol* gene by the HIV PR during maturation structure and function of HIV-1 IN (Chiu & Davies, 2004). HIV-1 IN is

comprised of three domains: A, the catalytic core domain (amino acids 50 – 212); B, the N-terminal domain (NTD) (amino acids 1 – 49) and the C-terminal domain (amino acides 213 – 288), Figure.1.8 (Chiu & Davies, 2004). The NTD enhances IN multimerization through zinc coordination (HHCC motif) and promotes the concerted integration of the two viral cDNA ends into the host cell chromosome. The C-terminal domain is responsible for sequence-independent, sequence-independent DNA binding to stabilize the IN–DNA complex (Marchand *et al*., 2006a). Each HIV-1 IN molecule consists of one catalytic site and has separate DNA binding sites for the ends of linear viral DNA and host cell DNA. The substrate of the enzyme is short double-stranded oligonucleotides corresponding to termini, called *att* sites. The activities of the *integrase* gene include the cleavage of deoxythymidylate-deoxythymidylate (TT) dinucleotide from the 3' of double-stranded *att* site substrate, the cleavage of double stranded target DNA to produce a staggered 5 base overhang and the transfer of activity in which the recessed 3' end in the substrate DNA is joined to the 5'phosphoryl end in the target DNA.

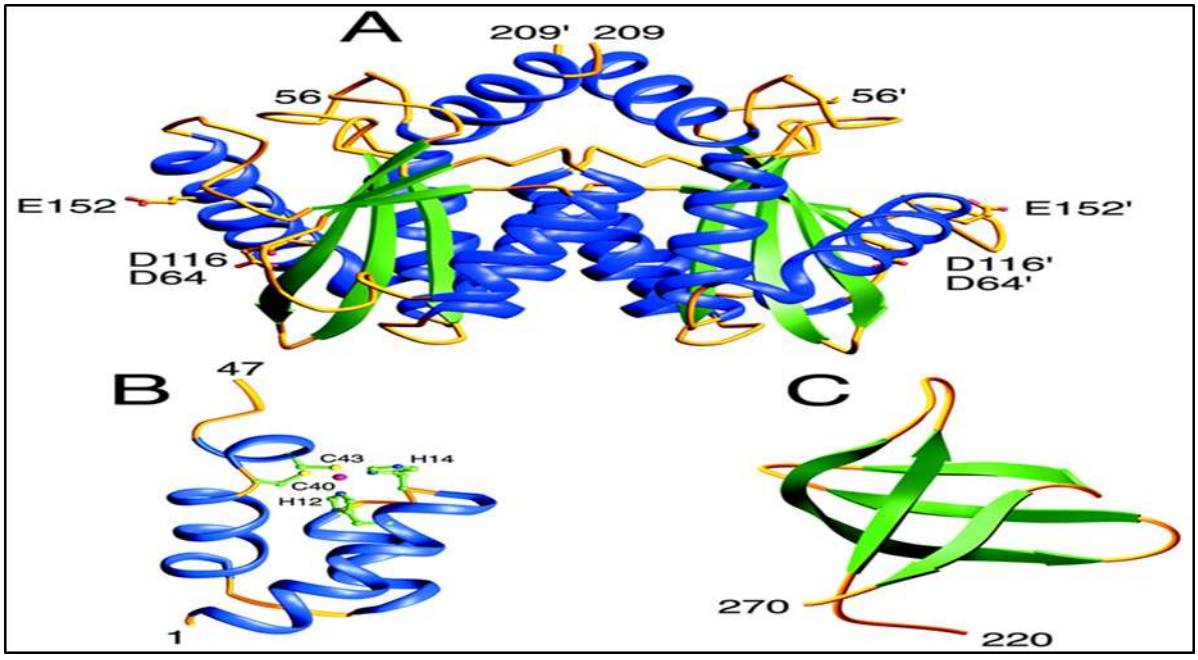


Figure 1.8: Structures of the three domains of HIV-1 integrase shown as ribbon diagrams.

A, the catalytic core domain; B, the N-terminal domain;C, the C-terminal domain. The drugs are designed to target the highly conserved Catalyic Core domain (A) (Robert, 2001).

1.4. Combination antiretroviral (cART)

Since the discovery of AIDS, tremendous efforts have been made in the development of antiretroviral compounds that target the different steps in the life cycle of HIV-1. cART has

transformed HIV-1 infection from an acute illness to a controllable chronic disease, if managed correctly. Dramatic reductions of morbidity and mortality have been achieved since the wide-scale use of cART, however no cure has been found (Severe *et al*., 2005; Egger *et al*., 2008). A major problem that interferes with the success of cART is the emergence of drug-resistant variants in patients under treatment, which complicates treatment options and hampers good prognosis (Quiros-Roldan *et al*., 2001; Rousseau *et al*., 2001; Winters *et al*., 2000; Lorenzi *et al.*, 1999; D'Aquila *et al.*, 2002; Yeni *et al*., 2002).

1.4.1. Natural resistance mechanisms to HIV-1.

In this section, I discuss the ability of how the human immune system is able to resist HIV-1 naturally. There are individuals who are resistant to HIV/AIDS without the use of cART, thus natural resistance. These individuals are divided into two groups: Individuals who are continuously exposed to HIV-1, but still remain uninfected (Fowke *et al*., 1996; Kaul *et al*., 2001; Kaul *et al*., 2004) and long–term non-progressers (LTNPs) (Easterbrook *et al*., 1999).

Three variants of chemokine-related genes associated with natural resistance to HIV infection have been studied extensively. These include CCR5-Δ32, CCR2-64I, and SDF1-3'A. CCR5-Δ32 is a naturally occurring mutation leading to the deletion of 32 (bp's) in the CCR5 gene (Nkenfou *et al*., 2013). Individuals who are homozygotes for CCR5-Δ32 confers natural resistance to HIV strains that use the CCR5 co-receptor (Galvani *et al*., 2005). The CCR5-Δ32 allele is found in approximately 5–15% European Caucasians, while it is not found in Africans and East Asians (de Silva, 2004). A mutation that exists on -2459 (A/G) in the CCR5 promoter has indicated a delay in progression to AIDS and seropositive individuals with the G/G genotype display slower progression to AIDS than those with the A/A genotype (Knudsen, 2001).

A polymorphism of CCR2-V64I creates a change in which a Valine within the first transmembrane segment of the receptor was substituted by an Isoleucine (Smith, 1997). The allele frequency does not vary much among populations, with an average frequency of 10–20%. CCR2-64I plays no role in the onset of HIV-1 infection, but heterozygotes for this allele were found to experience delayed progression to AIDS compared to those than homozygotes for the wild-type allele (Smith, 1997).

Genetic mutations in human beta defensins (HBD) are shown to be protective against HIV-1 infection. Homozygosity for the A692G polymorphism in *DEFB1* was reported to be significantly more frequent in ESNs than in SPs (Zapata *et al*., 2008). Polymorphism of SDF-

1-3' in the SDF-1 gene 21 showed a delayed progression rate to AIDS in homozygous individuals (Dezzutti *et al*., 2000; Winkler *et al*., 1998). A study conducted in individuals residing in Kenya showed that the extent of Human Leukocyte Antigen (HLA) discordance between mother and neonate plays a major role in natural resistance to vertical transmission of HIV-1 (MacDonald *et al*., 1998). Further studies also showed that saliva that contains a high level of SLPI were associated with low incidence of HIV-1 among neonates who acquired the disease during lactation after the first lactation month, meaning that SLPI can protect vertical transmission (Farquhar *et al*., 2002). Tripartite motif (TRIM) proteins (TRIM5α) are involved in protection against HIV-1 replication. TRIM5α found within the rhesus monkey (TRIM5αrh) blocks the HIV-1 by degrading the HIV before it undergoes reverse transcription (Reymond *et al*., 2001). Therefore, mutations within TRIM5α protect individuals against HIV-1 infection (Javanbakht *et al*., 2006). APOBEC3G was discovered to be associated with the inhibition of HIV-1 infection by reducing the fitness of Vif protein-deleted strains of HIV, thereby blocking the integration of the viral cDNA into the human genome (Sheehy *et al*., 2002). Further support for this concept stems was from data by Jin *et al., 2005,* showing correlations between APOBEC3G levels, CD4+ T cell counts and viremia in a cohort of LTNP indicating a role of this protein in the control of HIV disease progression (Jin *et al*., 2005).

1.4.2. Mechanisms of HIV-1 antiretroviral drugs

In the early 1990s, ART regimens were provided as monotherapy. In most cases Nevirapine was given for the prevention of Mother to Child Transmission (pMTCT), while Zidovudine (AZT) was given as monotherapy in patients with symptomatic HIV infection. However, monotherapy led to the quick development of RAMs. Thus, HIV-1 care changed the administration of monotherapy to use combinations of antiretroviral drugs known as highly active antiretroviral treatment (HAART), today referred as cART (Collier *et al*., 1996; D'Aquila *et al*., 1996). The development of cART has resulted in greater suppression of HIV from replicating, thus reducing mortality and morbidity (Hirsch *et al*., 2008) and prolonging the life span of HIV infected individuals (Fauci *et al*., 1996).

There are currently five classes of cART drugs that are available that target the HIV-1 life cycles. This includes nucleoside and nucleotide reverse transcriptase inhibitors (NRTIs and NtRTIs), non-nucleoside reverse transcriptase inhibitors (NNRTIs), protease inhibitors (PIs), integrase inhibitors (InSTIs) (also termed integrase strand transfer inhibitors) and entry inhibitors (Panel on Antiretroviral Guidelines for Adults and Adolescents, 2015). Most of the

designed cART work in a manner that attempt to exhibit activity that will inhibit viral replication by targeting RT, PR, IN or chemokine receptors used by the virus to enter the human host cell (Johnson *et al*., 2003).

The nucleoside/nucleotide reverse transcriptase inhibitors (NRTIs) were the first ART regimens to be approved by the Federal Drug Administration (FDA) for the use in the treatment of HIV-1 infection (Young, 1988). NRTIs inhibit the HIV replication by becoming incorporated into the viral DNA and interrupt viral replication via competitive with cellular deoxynucleoside triphosphates and result in DNA chain termination (Weller, 2001). Moreover, NRTIs play an important role in the pMTCT and their role in treatment as prevention (TasP) and pre-exposure prophylaxis (PrEP) is being extensively investigated (Gupta *et al*., 2012). Since 2001, pregnant mothers and infants were recommended to use Nevirapine (NVP) as part of the pMTCT program (Pillay *et al*., 2008a), while in the Western Cape province of South Africa , a combination of AZT and NVP was administered for pMTCT between 2004 and 2010 (Draper *et al*., 2008). The second generation RT inhibitors, the NNRTIs are a set of drugs that act as non-competitive inhibitors of HIV-1 RT-DNA by binding and physically block the the polymerase activity of the RT enzyme of HIV-1 (De Clercq, 1998). InSTIs are the new class of cART therapy that are designed specifically in blocking intergrase enzyme, a process whereby HIV proviral DNA is incorporated into the human host cell genome with the help of intergrase enzyme. Currently, there are only three INSTIs: [Raltegravir (RAL), Elvitegravir (EVG) and Dolutegravir (DTG)] that have been approved by the FDA (http://www.fda.gov). PIs are designed to target and actively bind the active site of the protease enzyme, thereby inhibiting the HIV-1 PR from cleaving the two *Gag* and *Gag-Pol* precursor proteins that are responsible for the production of new immature non-infectious virions (Karacostas *et al*., 1989). Fusion /entry inhibitor attempt to block the fusion of HIV with the host cell before it penetrates the host cell and start to replicate (Nagashima *et al*., 2001).

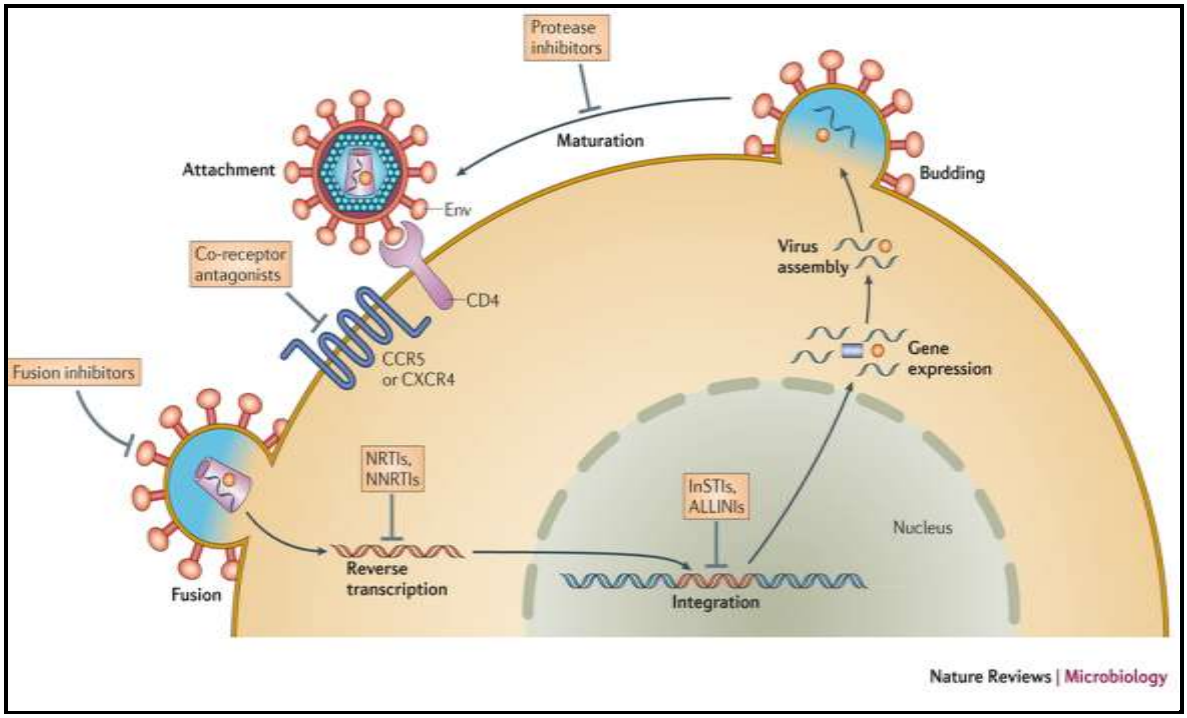


Figure 1.9: The figure illustrates the main steps in the HIV-1 replication cycle and the inhibitory site

The inhibitory antiretroviral drugs (green), and the step of the life cycle that they inhibit are indicated. Also shown are the key HIV restriction factors (tripartite motif-containing 5α (TRIM5α), APOBEC3G, SAMHD1 and tetherin; red) and their corresponding viral antagonist (Vif, Vpx and Vpu; blue). CCR5, CC-chemokine receptor 5; LTR, long terminal repeat; NRTIs, nucleoside reverse transcriptase inhibitors; NNRTIs, non-nucleoside reverse transcriptase inhibitors (Laskey *et al*., 2014).

1.4.3. HIV-1 drug resistance testing.

The emergence of HIV-1 drug resistance has been an important limiting factor that interferes with the success of cART. The accumulation of retrospective and prospective data has led three expert panels to recommend the routine use of HIV-1 drug resistance testing in the treatment of HIV-infected patients (EuroGuidelines Group for HIV Resistance, 2001; Hirsch *et al*., 2000; USA Department of Health and Human Services Panel on Clinical Practices for Treatment of HIV Infection, 2000).

South African cART guidelines recommend the use of NNRTI and two NRTIs as the first-line cART regimen and two NRTIs, a ritonavir (RTV)-boosted PI (bPI), usually Kaletra, as the second line therapy (NDOH, 2015). The South African cART drug resistance testing guidelines recommend that all persons (children and adults) who are experiencing failure on first-line therapy with a viral load of two measurements >1 000 RNA copies/ml as an indicator of therapy failure, to undergo resistance testing, combinedwithadherence

councelling (Cozzi-Lepri, 2009; Cozzi-Lepri., 2012). HIV-1 drug resistance testing can be assessed as either genotypic or phenotypic (Hirsch *et al*., 2003; Haubrich *et al*., 2001). Both assays are recommended as an integral part of HIV clinical care, such as selection of active drugs when changing cART regimens of the patient or individuals experiencing virological failure, monitoring, and surveillance the prevalence of drug resistance within individuals who are on cART. HIV-1 drug resistance surveillance and monitoring depend predominantly on sequences analysis of the *pol* gene to detect drug RAMs known to limit the success of the particular regimens (Clavel and Hance, 2004).

HIV-1 genotypic resistance testing assay involves sequencing of the PR and RT part of the *pol* gene to detect the presence or absence of the mutations that are known to confer resistance to cART (Eshleman *et al*., 2004). Genotypic resistance testing is commonly used in the diagnostic setting because of wider availability, affordability and it is less time consuming. After sequencing the patient HIV genome sequence is uploaded to a database with known algorithms of resistance detection associated with resistance mutations in the RT, PR, IN and envelope genes provides a useful guidance for interpreting genotypic resistance test results for HIV-1. The Stanford HIVDR database (http://hivdb.stanford.edu) is the most commonly used and up to date to interpret resistance results, but there are other available databases such as CPR (http://cpr.stanford.edu) and the SATuRN RegaDB Clinical Resistance Database to analyse and interpret resistance results.

Phenotypic resistance testing is an *in vitro* assay that measures the capability of HIV-1 isolates to replicate in the presence of different drugs at different concentrations. The concentration that block HIV-1 isolate growth by 50% [IC50] is calculated and results can be reported as fold-change in concentration of an inhibitor (IC50) (i.e. Fold resistance) for each HIV-1 (Lanier *et al*., 2004; Miller *et al*. 2004).

1.4.4. HIV drug RAMs.

HIV-1 drug RAMs in Africa are likely to increase in parallel with the wide scale-up of cART (Bennet *et al*., 2009). Several studies conducted in South Africa showed that many individuals who are on cART have high and / rapid frequency accumulation of NNRTI and other mutations, even while on first-line regimens (Marconi *et al*., 2008; Sunpath *et al*., 2008; Hoffmann *et al*., 2009; Barth *et al*., 2008). HIV transmitted drug resistance has also been documented in recent studies conducted in South Africa where a prevalence appears to be <5% (Barth *et al*., 2008; Jacobs *et al*., 2008; Pillay *et al*., 2008a; Pillay *et al*., 2008b).

There have been reports documented based on the development and evolution of antiretroviral drug resistance between HIV-1 B subtypes and several non-B subtypes major (M) group. M group HIV-1 subtypes are known to possess naturally occurring polymorphisms that causes change on several codons of the PR and RT and because of the change in the sequences of codons might predispose viral isolates to encode amino acids that are implicated in development of antiretroviral drug resistance (Holguin *et al*., 2002; Kantor *et al*., 2002). HIV-1 diversity can potentially influence the different level of treatment efficacy among individual infected by non-B subtypes. This is supported by the fact that HIV subtypes naturally differs by as much as 35% of the genetic content (Brenner *et al*., 2007; Kantor *et al*., 2006). Thus, a high degree of HIV-1 genetic diversity poses a greater challenge of producing the ART that is capable of inhibiting the enzymes of the *pol* gene among various subtypes of infected patients (Kantor & Katzenstein, 2003).

1.5. Aim

Thus, the aim of this study was to investigate the HIV-1 diversity in the remote areas of the Western Cape Province of South Africa (Cape Winelands, West Coast and Overberg).

1.6. Objectives

The objectives of this study were:

- To obtain recent HIV-1 viral load (VL) samples for HIV-1 genotyping.
- To identify HIV-1 subtypes through phylogenetic analysis in order to characterize HIV-1 diversity in the Western Cape Province of South Africa.
- To determine the prevalence of HIV-1 RAMs from our sequences obtained.

CHAPTER TWO

CHAPTER TWO

2. Materials

2.1. Introduction

A total of 205 HIV-1 Viral Load (VL) samples were used for HIV-1 genotyping and screening for RAMs during this study. The Western Cape Province have well-established consolidated guidelines for HIV treatment that have specific objectives, one of which is to use viral load testing as a preferred approach method for monitoring the success of cART. Viral load testing is also used as a tool for diagnosis of treatment failure, with consecutive high viral loads, after three months on cART, meaning the patient is likely failing therapy. Briefly, the materials used in this study are listed in this chapter.

2.2. Ethical permission

Ethical permission for this study was obtained and approved by the Human Research Ethics Committee at the Faculty of Medicine and Health Sciences at Stellenbosch University (Tygerberg Campus), South Africa. For ethical clearance, ethics permission was granted in 2015 and the project was registered under the application protocol number N15/08/071. It was renewed in 2016 (APPENDIX A). We obtained a waiver for the collection of diagnostic samples. All our work was conducted according to the ethical guidelines and principles of the international Declaration of Helsinki 2013, South African Guidelines for Good clinical Practice and the South African Medical Research Council (MRC) Ethical Guidelines for Research.

2.3. Equipment, commercial assays, enzymes and chemicals

The list of all the reagents, equipment, chemicals, software applications and assays used during the course of the study to characterize the Cape Winelands, Overberg and West Coast cohort samples are given in this chapter. In Table 2.1 the lists of chemicals, kits, and enzymes used are shown. In Table 2.2 the equipment needed to perform the necessary assays and analysis are presented, while the commercial packages used are given in Table 2.3. Buffers and additional media are listed in Table 2.4. Miscellaneous products used are listed in Table 2.5. The suppliers and lot numbers are provided where applicable. The ® and ™ designations

indicate that the products are either registered or trademark properties of the suppliers they are provided from.

Table 2.1: List of chemicals and kits used in the study

Chemical or commercial products and kits used	Company	Lot number
1 Kb DNA Ladder	Promega, USA	0461505
5x Gelpilot® Loading Dye	Qiagen, Germany	151051304
Agarose	Lonza Rocland, USA	0000381826
BigDyeTM Terminator cycle sequence ready Kit	Applied Biosystems, USA	1508055
dNTP´s	Promega, USA	002547
GoTaq® DNA Polymerase	Promega, USA	0000106979
Nuclease free water	Promega, USA	145045078
QIAamp Viral RNA Mini Kit	Qiagen, Germany	151017169
QIAquick PCR Purification Kit	Qiagen, Germany	151026478
Sam Solution	Applied Biosystems, USA	1410046
SuperScsript ® III one -step RT with Platinum Taq kit DNA Polymerase	Invitrogen, USA	1257-035
BigDye X-Terminator® purification kit	Applied Biosystems, USA	1509067

Table 2.3: Equipment used to perform sample analysis

Equipment	Supplier
ABI prism® 310 Genetic Analyzer	Applied Biosystems, USA
UVItec Gel Documentation System	UVItec Limited, UK
Beckman Coulter Allegra ™ 6R centrifuge	Beckman Inc, USA
geneAMP® 9700 PCR system	Applied Biosystems, USA
Microfuge® 16	Beckman Inc, USA
Labcon® FSIE shaking incubator	Labcon Ltd, RSA
Nanodrop™ ND 1000	Nanodrop Technologies Inc, USA
Vortex Mixer VM 300	Gemmy Industrial Corp, Taiwan

Table 2.4: Software programs and online tools used in sequence analysis

Software package	References and/or licensed companies
Sequencher v 4.8	Gene Codes Corporation, Ann Arbor, USA
ClustalW v 2.1	Thompson© *et al*., 1997
MEGA 6.0	Tamura *et al.*, 2011
Geneious	Biomatters Ltd, New Zealand
BioEdit v5.0.9	Hall, 2003. Ibis Biosciences, USA

Table 2.5: Databases used in the analysis of sequences

Databases	Web Address
Los Alamos National Laboratory HIV Sequence Database	http://www.hiv.lanl.gov
Stanford HIV RT and Protease Sequence Database	http://hivdb.stanford.edu
BioAfrica	http://www.bioafrica.net
Calibrated Population Resistance (CPR)	http://cpr.stanford.edu
Sequence-Specific subtype analysis (REGA)	http://hivdb.stanford.edu
Subtype classification using evolutionary algorithms (SCUEAL)	http://www.datamonkey

Table 2.6: Miscellaneous products used

Material	Supplying Company
Glassware	Labotec, RSA
Laboratory liquids (Ethanol, Methanol)	Sigma-Aldrich, USA
Laboratory wear	Labotec, RSA
Parafilm	Sigma-Aldrich, RSA
Pipette tips	Thermo Fisher Scientific, RSA
Pipettes	Thermo Fisher Scientific, RSA
Plastic material	Thermo Fisher Scientific, RSA
Sterile filters and filter paper	Thermo Fisher Scientific, RSA

CHAPTER THREE

Page

3. Methodology

CHAPTER THREE

3. Methodology

This study was a retrospective study. The samples were previously used for HIV-1 viral load quantification to monitor the success of cART and to diagnose the treatment failure through analysis of the *pol* gene of the HIV-1 viral genome. In this study, we only selected samples from the Western Cape Province districts for further analyses. Briefly, after HIV-1 viral RNA extraction from the plasma samples, the viral nucleic acid was used as a template for the reverse transcription of the PR, RT and IN genomic regions of the *pol* gene. Viral RNA was converted into cDNA during reverse transcription. This was followed by PCR amplification. PCR products were visualized on an agarose gel, and then purified for DNA cycle sequencing. Positive PCR products were purified for analysis with the capillary electrophoresis from which DNA sequence data are acquired. The latter were further analyzed with online subtyping tools (REGA HIV Subtyping tool version 3; Recombinant Identification Program (RIP) 3.0; subtype classification using evolutionary algorithms (SCUEAL) software and through the Stanford University HIV Drug Resistance Database (HIVdb) 8.1. Multiple alignments formed of query sequences and reference sequences were created and used to infer phylogenetic trees (http://www.hiv.lanl.gov). Figure 3.1 shows a brief overview of the techniques used in this study.

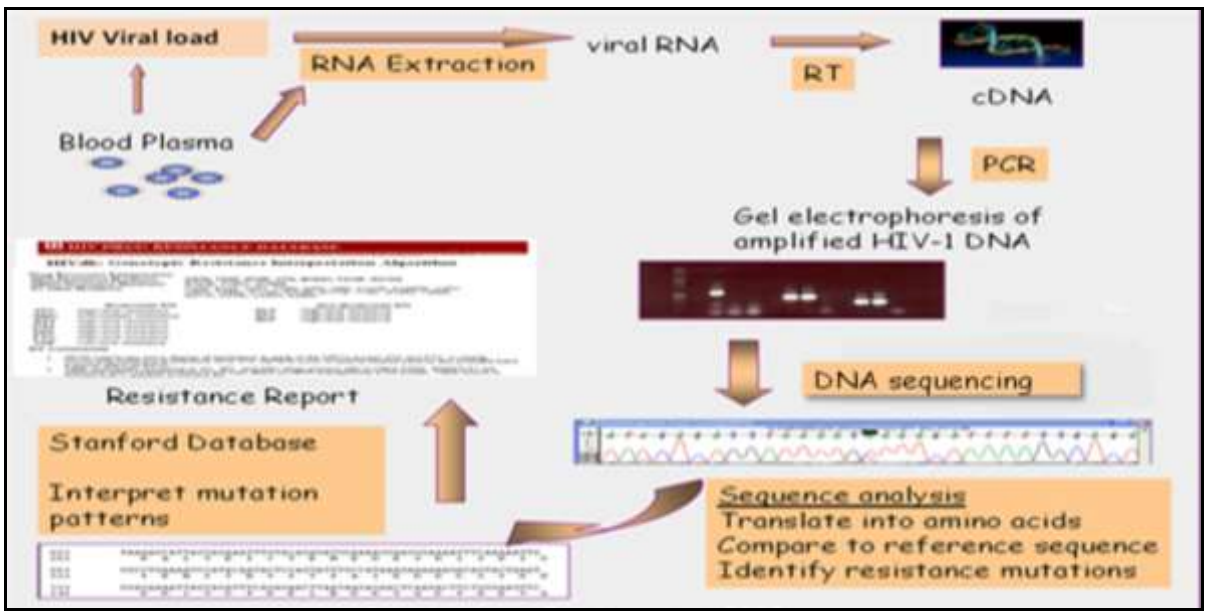


Figure 3.1: Flow diagram of molecular techniques used during the study.

HIV-1 RNA extraction for viral loads was done using the automated Abbott m2000sp, while the RealTime Abbott m2000rt analyzer was used to determine the HIV-1 viral load (Abbott laboratories, Abbott Park, Illinois,

USA). This was done by the National Health Laboratory Services (NHLS) diagnostic section of the Division of Medical Virology. The PR, RT and IN fragments of the HIV-1 viral genome were amplified by nested Polymerase Chain Reaction (PCR). The amplified gene fragments were visualized by agarose gel electrophoresis. The positive PCR products were purified and the nucleotide sequences of the HIV-1 genomes were directly determined. Detailed sequence verification and phylogenetic analysis were performed on all sequences obtained.

3.1 Smple preparation for viral load

HIV-1 patients were sent to and received at the diagnostic Virology Laboratory of the National Health Laboratory Services (NHLS), Tygerberg Hospital, Stellenbosch University in Cape Town, South Africa between January 2011 and December 2015. Samples were sent in Plasma Preparation Tubes). Approximately 2-3 ml plasma was stored from the 5 ml PPT blood samples by separating it from the red blood cells, white blood cells, and blood platelets. The plasma was collected after centrifugation (Beckman Coulter Allegra™ 6R, Beckman Inc., USA) at 2000 revolutions per minute (rpm) for 10 minutes at 4°C. Sample selection was based on patient district location and patient viral load data of 2000 copies/ml. Gall *et al*., 2012 showed that a sensitivity of at least 2000 copies/ml of the HIV-1 viral load is a limit of detection for the Superscript ® III one -step RT with Platinum Taq kit DNA Polymerase kit (Gall *et al*., 2012). In addition, this sensitivity of 2000 copies/ml was sufficient for PCR amplification of (99.3%) samples in their study. Only samples with sufficient plasma volumes left after HIV-1 viral load measurement were used for genotyping. Viral RNA was extracted from 1ml plasma using the high throughput M1000 robot extractor (Abbott Diagnostics, USA) in 200 uL volume of COBAS®AmpliPrep/COBAS® TaqMan® Wash Reagent buffer used for viral load determination. The sample RNA were stored at -20°C for subsequent analysis.

3.2 HIV-1 viral load assay

The Abbott RealTime ® HIV-1 RNA Quantitative assay (Abbott Laboratories, Illinois, USA) was used according to the manufacturer's instructions. Briefly, the assay is based on Reverse Transcription PCR (RT-PCR) technology with homogeneous real-time fluorescent detection and targets the *pol* IN gene to determine the viral load concentration. Labelled probes, specific for HIV-1, detect the DNA products in the sample after RT-PCR. The amplified product-probe hybrids were quantitated with a microparticle enzyme immunoassay (MEIA) on an automated Abbott Real Time® Analyzer (Abbott Laboratories, Illinois, USA). A volume of 1000 µl was used to determine the number of HIV-1 RNA copies per ml in each plasma sample. The assay has a range of 50 to 1 million copies per ml if 1 ml sample volume

is used. If a 200 µl sample is used, the detection limit ranges from 178 to 5 million copies per ml. The assay can detect and quantitate HIV-1 group M subtypes A-G, as well as HIV-1 group O (Johanson *et al*., 2001; Zanchetta *et al*., 2000). The assay results were reported in copies/mL or international Units/mL (IU/mL).

3.3. Viral nucleic acid extractions

For the purposes of my study and for fresh RNA extraction, Viral RNA was extracted from 140µl of plasma using the QIAamp viral mini kit (Qiagen GmbH, Germany) using the vacuum extraction system. The kit protocol was followed and RNA was eluted in 60µl of low salt elution buffer AVE containing 0.04% sodium azide. The Qiagen RNA extraction protocols are based on the silica-gel membrane principles developed by Vogelstein and Gillespie (1979). Briefly, lysing buffer is added to the patient sample to inactivate RNases and to ensure isolation of intact viral RNA. The sample is loaded onto the QIAamp Mini spin column, which binds the RNA to a silica membrane. The buffering conditions are adjusted to provide efficient binding of the RNA to the QIAamp membrane of the Spin Column during centrifugation or a vacuum step. Contaminants are efficiently washed away in two steps using two different wash buffers that ensure that residual contaminants are removed and improve the purity of the samples. High-quality RNA is eluted in a special RNase-free buffer, ready for direct use or safe storage. The purified RNA is free of protein, nucleases, and other contaminants and inhibitors. Isolated RNA was used for the molecular characterization of PR, RT, and IN, the region of the *pol* gene fragments.

3.4. Polymerase chain reaction (PCR)

The PCR method was used for the cycle sequencing reactions and HIV-1 genotyping reactions. The reactions were performed on the geneAMP® 9700 PCR system (Applied Biosystems, USA) with cycling parameters for the *pol* fragments described in the tables below. The ramp speed is calculated as the speed at which the PCR machine (thermocycler) switches between heating steps. A brief overview of the steps used to perform the PCR experiments are described as follows: A total volume of 5µl of RNA was used for pre-nested (first round) *pol* RT-PCR. The RNA was converted to generate cDNA by using the SuperScript® III one-step RT with Platinum Taq DNA polymerase (Invitrogen, USA) and then used as the template in a one-step PCR reaction. The SuperScript ® III one -step RT with Platinum Taq DNA polymerase kit is a one-step method where primers targeting a specific gene of interest are being used for both the cDNA synthesis and PCR amplification,

in a single reaction tube. The primer pairs used for pre-nested and nested *pol* amplification are given in Table 3.1, Table 3.3 and Table 3.5, whereas cycling conditions are presented in Table 3.2, Table 3.4 and Table 3.6. The primers were used at a final concentration of 5µM/µL. The final concentrations of the remaining reagents were as follows: 2X Reaction Mix (a buffer containing 0.4 Mm of each dNTP, 3.2 mM $MgSO_4$), SuperScript ® III RT/Platinum ® *Taq* Mix and nuclease free water was added to a make a final volume of 50µl. A second nested PCR was performed, using 5µl of the pre-nested PCR product as template. The final concentrations of reagents for nested PCR reaction were as follows 0.2mM dNTPs, 1.5mM $MgCl_2$, 5 x GoTaq Flexi buffer, 2.5 U/µlTaq DNA polymerase, water was added to a final volume of 50µl per reaction. All the PCR mixtures were prepared in 0.2 ml thin wall PCR® tubes (QSP, Porex BioProducts Inc., USA).

Table 3.1: Primers used in the amplification of the partial HIV-1 Protease gene products

PCR STEPS	Primer	F/R	Oligonucleotide sequence	HXB2 Location	Size (bp)
cDNA and first round	outer 30'prot1	R	GCAAATACTGGAGTATTGTATGGATTTTCAGG	2734-2703	653
	outer 50'prot1	F	TAATTTTTTAGGGAAGATCTGGCCTTCC	2085-2109	
Nested PCR	inner 30'prot2	R	AATGCTTTTATTTTTTCTTCTGTCAATGGC	2529-2558	507
	inner 50'prot2	F	TCAGAGCAGACCAGAGCCAACAGCCCCA	2136-2163	

References are listed in section 3.8 along with the primers.

Table 3.2: Cycling parameters of Protease gene products

Cycling conditions for the Pre-Nested *PR* PCR assay			
Step	**Temperature**	**Duration**	**Cycles**
Reverse transcription	50^0C	30 min	X 1
Initial denaturation	94^0C	2 min	
Denature	94^0C	15 sec	X 40
Anneal	60^0C	30 sec	
Extend	68^0C	1 min	
Final extension	68^0C	5 min	X 1
Cycling conditions for the Nested PR PCR assay			
Step	**Temperature**	**Duration**	**Cycles**
Initial denature step	94^0C	2 min	X 1
Denature	94^0C	20 sec	X 40
Anneal	60^0C	30 sec	
Extend	68^0C	90 sec	
Final extension	68^0C	5 min	X 1

Sec – seconds; min – minutes; °C – Degrees Celsius

Table 3.3: Primers used in the amplification of the partial HIV-1 Reverse Transcriptase gene products.

PCR steps	Primer	F/R	Oligonucleotide sequence	HXB2 Location	Size (bp)
Pre-Nested PCR	outer Mj3	R	AGTAGGACCTACACCTGTCA	2480 -2499	941
	outer Mj4	F	CTGTTAGTGCTTTGGTTCCTCT	3399-3420	
Nested PCR	inner A35	F	TTGGTTGCACTTTAAATTTTCCCATTAGTCCTATT	2530- 2564	798
	inner NE135	R	CCTACTAACTTCTGTATGTCATTGACAGTCCAGCT	3300-3334	

References are listed in section 3.8 along with the primers

Table 3.4: Cycling parameters of Reverse Transcriptase gene products

Cycling conditions for the Pre-Nested *RT* PCR assay			
Step	**Temperature**	**Duration**	**Cycles**
Reverse transcription	50^0C	30 min	X 1
Initial denaturation	94^0C	2 min	
Denature	94^0C	15sec	X 40
Anneal	60^0C	30sec	
Extend	68^0C	1 min	
Final extension	68^0C	5min	X 1
Cycling conditions for the Nested PR PCR assay			
Step	**Temperature**	**Duration**	**Cycles**
Initial denature step	94^0C	2 min	X 1
Denature	94^0C	20 sec	X 40
Anneal	60^0C	30 sec	
Extend	68^0C	90 sec	
Final extension	68^0C	5 min	X 1

Sec – seconds; min – minutes; °C – Degrees Celsius

Table 3.5: Primers used in the amplification of the partial HIV-1 Integrase gene products.

PCR Steps	**Primer**	**F/R**	**Oligonucleotide sequence**	**HXB2 location**	**Size (bp)**
Pre-nested PCR	Poli8	R	TAGTGGGATGTGTACTTC TGAAC	5210-5232	1056
	Poli5	F	CACACAAAGGRATTGGAG GAAATG	4177-4200	
Nested PCR	Poli6	R	ATACATATGRTGTTTTACT AARCT	5122-5145	945
	Poli7	F	AACAAGTAGATAAATTAG TCAGT	4201-4223	

References are listed in section 3.8 along with the primers

Table 3.6: Cycling parameters of Integrase gene products.

Cycling conditions for the Pre-Nested *PR* PCR assay			
Step	Temperature	Duration	Cycles
Reverse Transcription	50^0 c	30 min	X1
Initial denaturation	94^0 c	2 min	
Denature	94^0 c	15sec	X40
Anneal	55^0 c	30sec	
Extend	68^0 c	90 sec	
Final Extension	68^0 c	10 min	X1
Cycling conditions for the Nested PR PCR assay			
Step	Temperature	Duration	Cycles
Initial Denature Step	95^0 c	2 min	X1
Denature	95^0 c	20 sec	X40
Anneal	55^0 c	30 sec	
Extend	70^0 c	90 sec	
Final Extension	72^0 c	10 min	X1

Sec – seconds; min – minutes; °C – Degrees Celsius

3.5. Agarose gel electrophoresis

Gel electrophoresis is the standard laboratory procedure of separating and analyzing nucleic acids based on the molecular weight size. During agarose gel electrophoresis DNA molecules are separated by applying a constant electrical field to the electrophoretic apparatus. Negatively charged DNA migrates from anode to cathode through an agarose gel matrix, shorter molecules migrate faster and easily than longer molecules through the pore of the agarose gel matrix (Sambrook *et al*., 1989).

All PCR products were visualized by agarose gel electrophoresis on a 1.0% agarose (Seakem LE® agarose; FMC BioProducts, USA) gel in 1 x TAE buffer (0.04 M Tris-acetate, 0.001 M EDTA). After the gel solution cool to about 60^0C, GRGreen (Qiagen, Germany) was added to the solution to 1x final concentration to stain the gel. Samples were mixed with 6 x DNA loading dye (0.125% Bromophenol blue, 40% Sucrose) and approximately 5µl were loaded per lane on the agarose gel. Electrophoresis reactions were run at 75 V for a 50 ml agarose gel. A 1 kb DNA molecular weight marker (Promega, USA) was used as an estimation of DNA band size also included was a negative and a positive control. The DNA bands were visualized under an ultraviolet light at a wavelength between 280 and 320 nm and photographed with the Syngene™ GeneGenius computer system (Synoptics Ltd, United

Kingdom). The expected fragment sizes for PR, RT and IN gene fragments were 507 bp, 798 bp, and 985 bp, respectively.

3.6. Purification of nucleic acids

To remove the excess dNTPs, primers, enzymes and buffers that might interfere with the sequencing reactions the above mentioned-PCR products were purified with the QiaQuick PCR purification kit, according to the manufacturer's instructions (Qiagen, Germany). The purification protocols are based on silica membrane spin protocols that allow for binding of nucleic acids to a silica membrane inside a spin column (Vogelstein & Gillespie, 1979; Sambrook *et al*., 1989). Briefly, 5 volumes of phosphate buffer (PB) was added to 1 volume of the PCR sample (example, 47 µl was added to 235 µl of buffer PB). DNA was bound to the spin column membrane by centrifugation using microfuge® 16 at a maximum speed of 13 000 rpm (Beckman Inc) and washed once to remove any impurities that may still be present. DNA was eluted with 50µl of buffer EB (10 mM Tris-Cl, pH 8.5). The purified aliquot was stored at -20°C until the sequencing reactions were performed.

3.7. DNA concentration determination

The NanoDropTM system (NanoDrop technologies, USA) was used to determine the DNA concentration spectrophotometrically, using a 1µl input sample. DNA measurements were read at a wavelength of 260 nanometers (nm), while purity was determined by dividing the absorbance at 260 nm with the absorbance at 280 nm (Sambrook *et al*., 1989). Pure DNA has a value between 1.7 and 2.3. Lower values are indicative of protein contamination, while higher values indicate the presence of RNA in the sample.

3.8. DNA cycle sequencing reactions

DNA cycle sequencing reaction is performed in order to linearize the DNA fragment of interest and amplification of the labeled DNA. Briefly, each reaction was performed as follows: Approximately 50ng of the purified PCR products was used in the sequencing reactions. The BigDye^{TM} Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems, USA) was used for the PCR-based sequencing reactions according to the manufacturer's instructions. Briefly, 1µl of the Big Dye terminator enzyme mix (BigDye® Terminator Cycle Sequencing Kits, Applied BioSystems, USA), 5 pmol of primer, ~ 50ng of sample DNA, 2µl of 5x Sequencing Buffer (Applied BioSystems, USA) and Nuclease-Free water (Promega, USA) was added to each reaction to make up a final volume of 10µl. Each

sequencing reaction was performed under the following conditions on a GeneAmp PCR System 9700 thermal cycler (Applied BioSystems, USA): Denaturation at 96^{o}C for 10 seconds, primer annealing for 5 seconds, annealing temperature according to sequencing primers, Table 3.7 and Table 3.8, and an elongation step at 60^{o}C for 4 minutes. The cycles were repeated 25 times after which the samples were cooled down to 4^{o}C. The latter reaction mixture was purified using the BigDye X-Terminator Purification Kit and analyzed on an ABI 3130XL genetic analyzer (Applied Biosystems, USA).

The trace data files of each sequencing run were subsequently retrieved and imported into Sequencer v 4.8 (Gene Codes Corporation, Ann Arbor, USA). The quality of each chromatogram fragment was assessed manually. Each patient's chromatogram fragments were assembled into a single contigue and manually proofread in order to assure good sequence quality. After the fragments were proofread, they were exported in a text file (.txt) format and labeled.

Table 3.7: Summary of primers and cycling conditions used for the sequencing of pol PCR fragments

Primers	F/R	Annealing Temperature	Oligonucleotide sequence	References
JA217	R	60°C	CTTTTATTTTTTCTTCTGTCAATG	Plantier *et al*., 2005
30 PROT 2	R	60°C	AATGCTTTTTCTTCTGTCAATGGC	Plantier *et al*., 2005
*POL b	F	60°C	CACTCTTTGGCARCGACC	Jacobs personal communication
Pol 1D	F	60°C	TCCCTCAAATCACTCTTTGGC	Loxton *et al*., 2005
AK 10	R	60°C	TYCCCACTAAYTTCTGTATRT	Plantier *et al*., 2005
AK 11	F	60°C	GTACCAGTAAAATTAAARCCAG	Plantier *et al*., 2005
*pol 3D	F	60°C	AATGTTGCACTTTAAATTTATCC	Loxton *et al.*, 2005
NE 135	R	60°C	CCTACTAACTTCTGTATGTCATTGT	Plantier *et al*., 2005
Poli5	F	55°C	CACACAAAGGRATTGGAGGAAATG	Engelbrecht personal communication
Poli6	R	55°C	ATACATATGRTGTTTTACTAARCT	Engelbrecht personal communication
Poli7	F	55°C	AACAAGTAGATAAATTAGTCAGT	Engelbrecht personal communication
Poli8	R	55°C	TAGTGGGATGTGTACTTCTGAAC	Engelbrecht personal communication
50 PROT 2	F	60°C	TCAGAGCAGACCAGCAGCCCCA	Plantier *et al*., 2005

POL B was designed by Dr Jacobs, POL3D was designed by André Gareth Loxton, within the Division of Medical Virology

Table 3.8: Sequencing cycle parameters

PCR steps	Temperature	Time	Cycles
Denaturation	96 ºC	10 sec	25 X
Annealing	ºC*	5 sec	
Extension	60 ºC	4 min	

*Annealing temperatures for sequencing primer sets were indicated in Tables above.

3.9. Sequence quality control

The online Quality control programme was used to examine the sequence sets for common problems: http://www.hiv.lanl.gov/content/sequence/QC/index.html. Results from the tool

include: subtype (from RIP) 3.0, similarity (from HIV BLAST), phylogenetic (from Neighbour TreeMaker), stop codons and frameshifts (from GeneCutter) and hypermutation (from HyperMut).

3.10. HIV-1 characterization using online subtyping tools

All Sequence-Specific subtype analysis was executed with the REGA HIV Subtyping tool version 3; Recombinant Identification Program (RIP) 3.0; subtype classification using evolutionary algorithms (SCUEAL) software and through the Stanford University HIV Drug Resistance Database (HIVdb) 8.1.

3.11. Sequence alignments and phylogenetic analysis

In order to study the evolutionary relationships between the different strains of HIV sequences obtained during the study, phylogenetic analyses were used. Phylogenetic analysis is the means of inferring or estimating the pedigree of the inherited relationships among organism o molecules looking at the molecular level of the organism such as DNA and / or protein sequences. Sequence contigs obtained were aligned using the ClustalW version 2.0 software package (Thompson *et al*., 2007) and manually edited using the BioEdit software package. Phylogenetic trees were infered with the Mega version 6.0 software package (Tamura *et al*., 2011). Reference sequences used in the phylogenetic analyses were retrieved from the LANL (http://www.hiv.lanl.gov/).

3.12. Drug resistance analysis

As we used *pol* sequences to study HIV-1 diversity, we were able to use the sequences to investigate the RAMs in our study cohort. For screening of the HIV-1 RAMs, the PR, RT, and IN fragments, FASTA formatted sequence files were deposited into the Stanford University HIV Drug Resistance database (HIVdb) online tools for analysis using the genotypic resistance interpretation algorithm version 8.1 (http://hivdb.stanford.edu/). Drug resistance mutations results were classified as either major or minor mutations.

CHAPTER FOUR

CHAPTER FOUR

4. Results

4.1. Patient demographics

Epidemiological and clinical data describing the population used in the study are presented in Table 4.1 and Figure 4.1. The study involved two hundred and five (205) HIV-1 Viral Load (VL) samples, 83 (40.49%) were male and 122 (59.51%) were female. The mean age of the patients was 34.41 years (range between 3 years 7 months to 63 years). The median viral load was 236317.2copies/mL (range between: 6166147 and 2035 copies/mL). Among the Western Cape Province HIV-1 samples analyzed in this study 129 (62.92%) live in Cape Winelands district, 29 (14.14%) in the Overberg district and 47 (22.92%) in the West Coast district. With our study samples, only 62/205 (30.2%) of patients indicated that they were receiving first-line cART at the time of sample collection. We are unable to indicate on which regimen the other patients were at the time of sample collection if any. The cART regimens are shown in APPENDIX B.

Table 4.1: Summary of epidemiological and clinical characteristics of the Western Cape Province HIV-1 viral cohort from the different geographic regions.

Parameters	Cape Winelands (n=129)	Overberg (n=29)	West Coast (n=48)
Age			
Oldest	63 years	59 years	52 years
Average	33.8 years	34.2 years	35.8years
Youngest	6 years	13 years	3 years 7 months
Gender			
Male	48	14	21
Female	81	15	26
Viral Load			
Highest	3249561	3272216	6166147
Average	201349.7	499256.6	212303.1
Lowest	2096	2252	2035

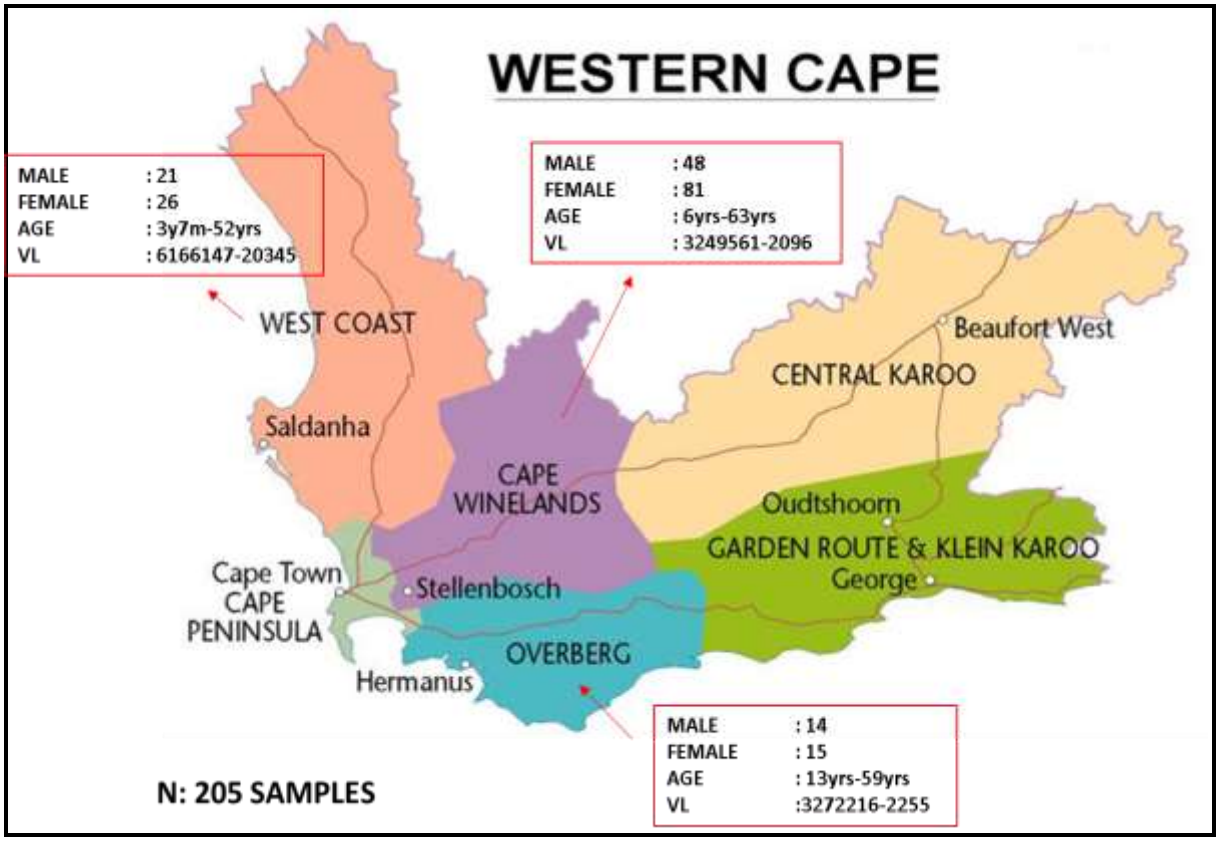


Figure 4.1: Map of the Western Cape region of South Africa, with the origin of the patients, indicated.

A total of 205 samples were analyzed during the study from the 3 districts: **Cape Winelands** (Bella Vista, n = 8; BrewelSkloof, n=2; Ceres, n=36 ;De Doorns , n=31; Mbekweni, n=16; N'duli Park, n=1; Paarl, n=3 ; Wolsely, n=2; Worcester, n=32; Zolani, n=1). **Overberg** (Bredasdorp, n = 4; Grabouw, n =23; Matjieskloof, n=1 ; Swellendam, n=1; Zwelihle, n=1.) and **West Coast** (ClanWilliam, n=5; Hanna Cotzee, n = 1; Malmesbury, n=27; Swartland, n=1; Vredenburg, n=3; Vredendal, n=6).

4.2. HIV Viral load assays

The gender, age, date of sample collection and VLs determined for each patient is given in APPENDIX C. The HIV RNA copies ranges between (6166147- 2035 cells per/ml).

4.3. Viral nucleic acid extractions

RNA was extracted from 205 HIV-1 viral cohort patient samples. A good quality RNA concentration level was obtained for all patient samples and ranged between 41.03 ng/µl and 160.89 ng/µl. The purity levels of all samples were good, being above 1.70 (260/280). APPENDIX D.

4.4. PCR and gel electrophoresis for the PR, RT and IN gene fragment

For the PR and RT gene fragments PCRs were performed on all 205 viral RNA extracted samples. For the IN region, only 176 samples had sufficient RNA and plasma left after the genotyping of the PR and RT region. All the amplified gene fragments were visualized by agarose gel electrophoresis and further analyzed.

4.4.1. PCR and gel electrophoresis for the PR and RT gene fragment

From a collection of 205 patient samples, PCRs were performed on the *pol* (PR and RT) gene fragments of the HIV-1 genome. One hundred and seventy (82.9%) and 166 (80.9%) of the *pol* PR and RT fragment region showed positive bands after agarose gel electrophoresis, respectively (Figure 4.2 and 4.3). The partial *pol* fragments (PR and RT) were visualized by agarose gel electrophoresis along with a positive control and Non-template control (NTC). From 170 PR and 166 RT fragment that positively amplified, only 82 (40%) of the samples were successfully amplified in both the RT and PR gene regions and 74 (36.1%) samples could not be amplified with any of the primers for both the PR/RT, following extensive methods to optimize the PCR conditions (e.g. changing the concentrations of $MgCl_2$ and/or slightly adjusting the annealing temperatures of primers as well as re-designing the primers). A summary of the PCR results is listed in Table 4.2.

Table 4.2: Summary of the PCR results for the different fragments

	PR	**RT**
n = Sample	205	205
Positive	170 (82.9%)	166 (80.9%)

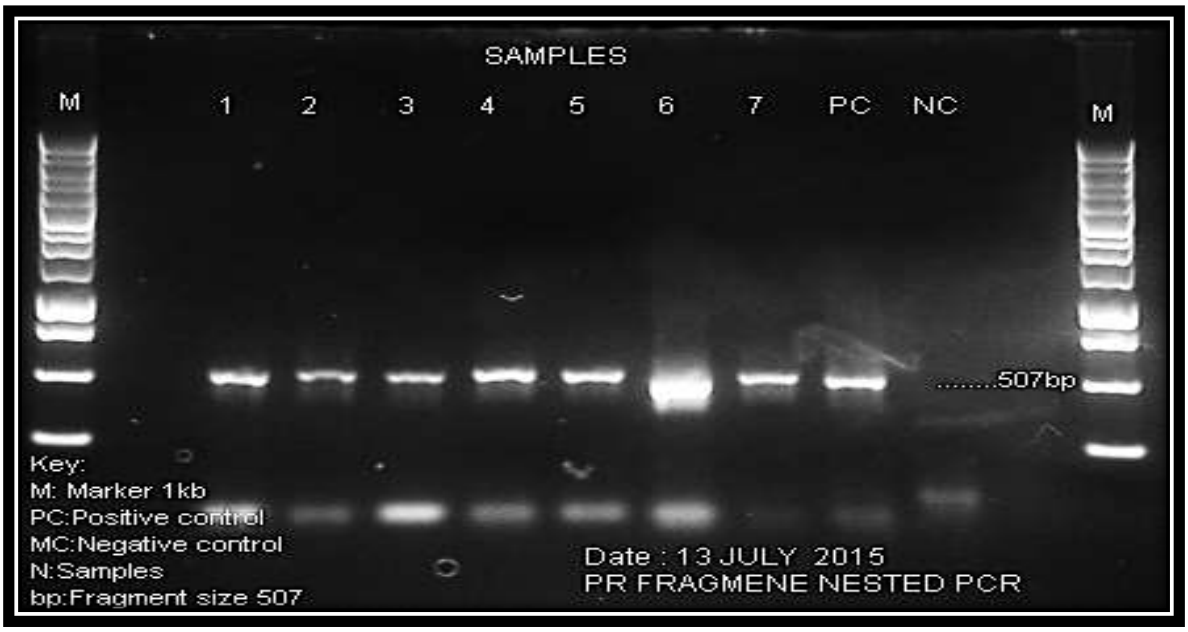


Figure 4.2: Partial *pol* PR gene fragments amplified in nested PCR assay.

After two rounds of amplification, the PR 507 bp was visualized on a 1% agarose gel. All samples were visualized on the agarose gel along with a 1kb molecular marker (M), Lane M: 1kb DNA Ladder, lane: 1-7 patient samples. Lane PC: Positive control, Lane NC: Negative control.

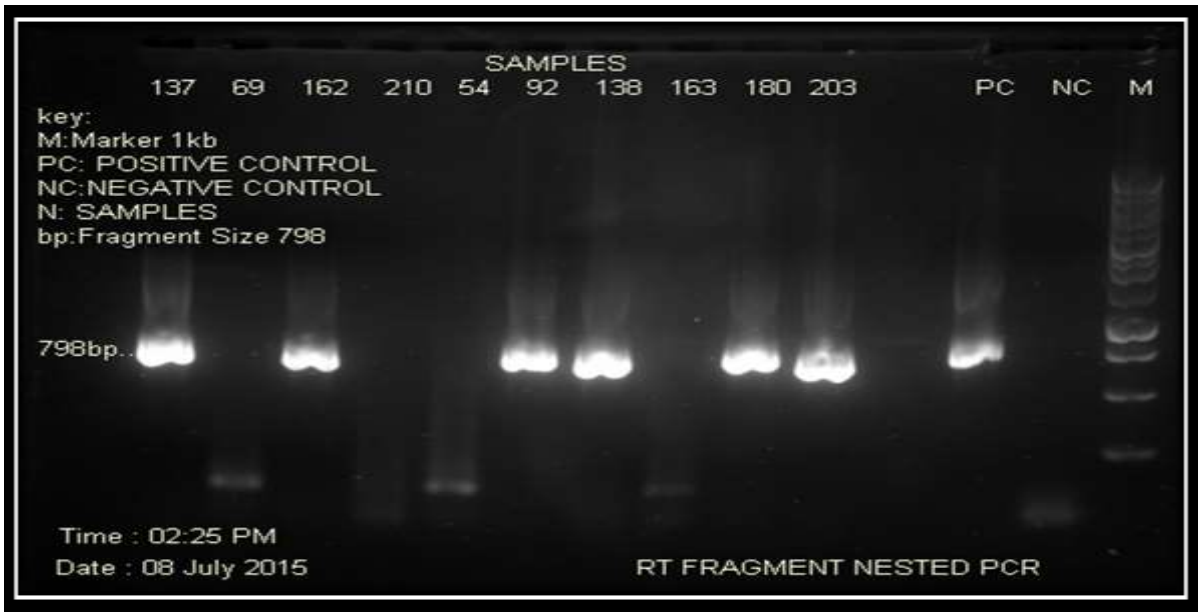


Figure 4.3: Partial *pol* RT gene fragments amplified in nested PCR assay

After two rounds of Amplification, the RT 798bp was visualized on a 1% agarose gel along with a 1kb molecular marker (M), Lane M: 1kb DNA Ladder, 137-203 patient samples. Lane PC: Positive control, Lane NC: Negative control.

4.4.2. PCR and gel electrophoresis for the IN gene fragment.

PCRs were performed on the *pol* IN gene fragment of the HIV-1 genome. One hundred and forty-three (81.3%) IN fragments showed positive bands after agarose gel electrophoresis (Figure 4.4). A positive control and Non-template control (NTC) is included on the gel. Thirty-three (18.8%) samples could not be amplified with any of the primers for the IN.

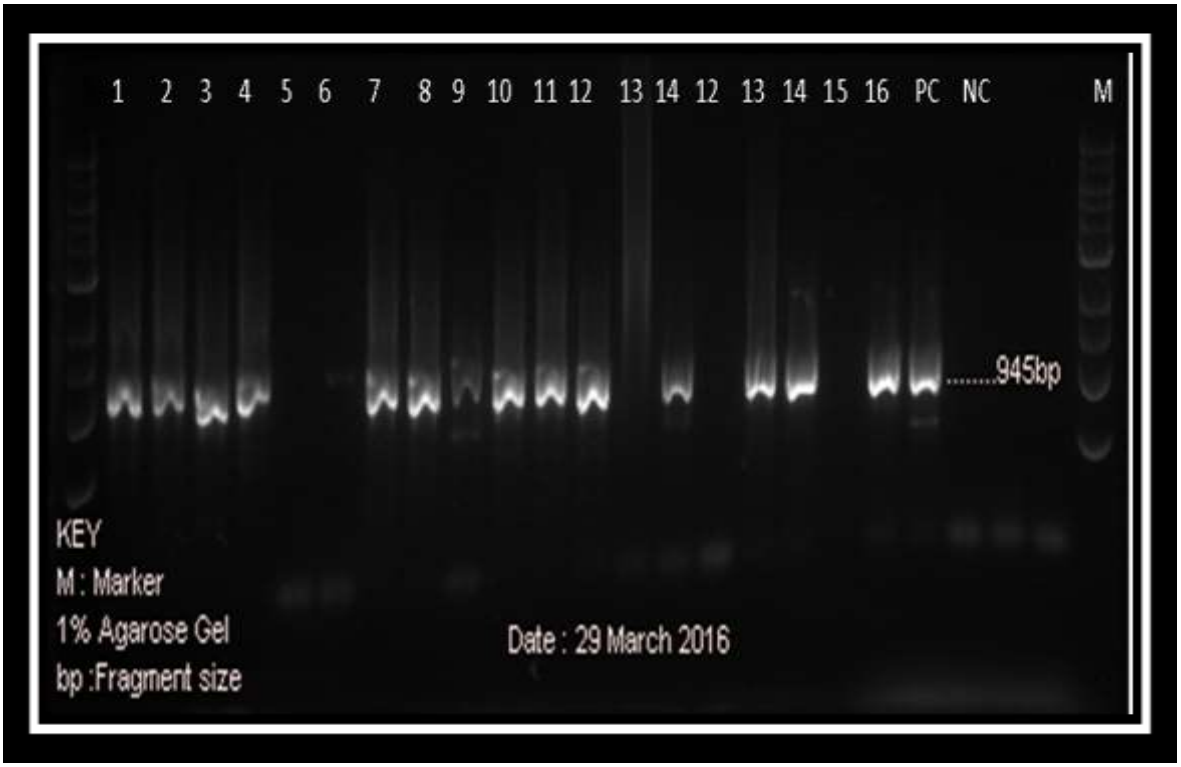


Figure 4.4: Partial *pol* IN gene fragments amplified in nested PCR assay.

After two rounds of amplification, the IN 945 bp was visualized on a 1% agarose gel along with a 1kb molecular marker (M), Lane M: 1kb DNA Ladder, lane: 1-16 patient samples. Lane NC: Negative control and Lane PC: Positive control.

4.5. Sequencing data quality.

All positive amplified PCR products were purified and sequenced for the PR, RT, and IN fragments. Of the 170 (82.9%) PR region PCRs 116 (68.2%) samples were successfully sequenced in both overlapping strands and 33 (19.4%), samples were positively sequenced in one strand only, followed by 21 (12.6%) samples that were not successfully sequenced with any of the primers. For the RT region, of 166 (81.0%) positively amplified fragments, 103 (62.0%) samples were positively sequenced in both strands, 38 (22.8%) samples were sequenced successfully in one strand only and 25 (15.1%) samples couldn't be sequenced for any of the primers used. Of the 143 (81.3%) IN region PCR products, 108 (75.5%) samples

were positively sequenced in both overlapping strands and 21 (14.7%) samples were positively sequenced in one strand only, followed by 14 (9.8%) samples that did not successfully sequenced with any of the primers used. All the sequences that were sequenced were checked for quality score and further edited manually with the Sequencer software to remove ambiguity nucleotides and sequences with quality less than 75.0% and fragments less than 200 bp. Further analysis was performed on the 62 (53.4%) PR, 103 (88.8%) RT and 93 (86.1%) IN sequences with quality ≥75%. The majority of the sequences that were sequenced on the PR fragment, generated shorter fragments sequences which is less than (200bp) and the possible reason for generating shorter fragment might be that the primers were not specific to the binding site of the protease gene. And this was followed by re-designing of the primer (PolB_fwd: 5'-CACTCTTTGGCARCGACC-3') that can sequence the partial PR fragment size of 507 bp.

4.6. Sequence quality control

All positive sequences were analyzed using the online Quality control tool to examine the sequence sets for common problems such as stop codons and frame shifts (from GeneCutter) and hypermutation (from HyperMut). When the online LANL QC tool indicated stop codons or frameshift, the sequences were re-read and verified. No Hyper- mutations was detected in all sequences after being re-read manually (http://www.hiv.lanl.gov).

4.7. HIV-1 characterization using online subtyping tools

Sequence-specific subtype analysis was executed with the REGA HIV Subtyping tool version 3; Recombinant Identification Program (RIP) 3.0; subtype classification using evolutionary algorithms (SCUEAL) software and also through the Stanford University HIV Drug Resistance Database (HIVdb) 8.1. The following tables below are the results summaries of the subtyping tools used to analyses the fragment of the PR, RT, and IN region of the *pol* gene, as described in sections 4.7. Figure 4.5, is an example of one of the sequence (12_ZA_WC_CW_20_RT) grouped as subtype A1, using the online subtyping tools and phylogenetic tree confirmed this sequence as subtype A1 tools

Table 4.3: PR fragment region analysis of non-subtype C sequences using various online tools

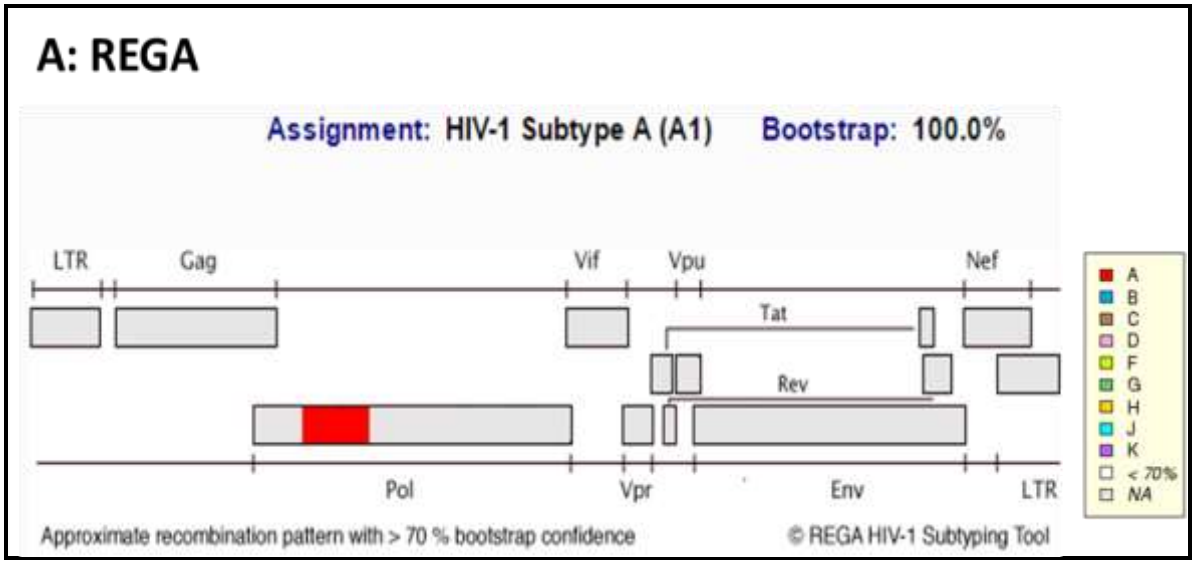
SEQUENCE CODE	REGA	RIP	SCUEAL	HIVDRD
15_ZA_WC_CW_22_PR	Unclassified	C	C	C
12_ZA_WC_OV_19_PR	Unclassified	C	C	C
13_ZA_WC_CW_66_PR	Unclassified	C	C	C

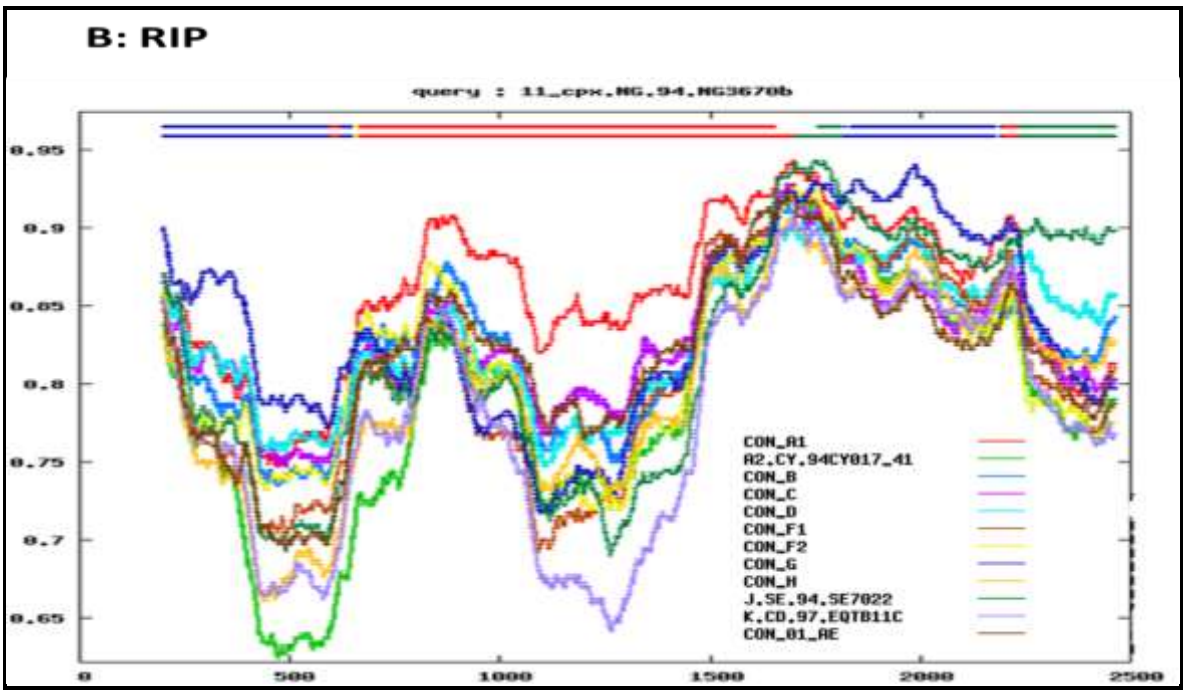
Table 4.4: RT fragment region analysis of non-subtype C sequences using various online tools

SEQUENCE CODE	REGA	RIP	SCUEAL	HIVDRD
14_ZA_WC_CW_49_RT	B	B	B	B
13_ZA_WC_CW_31_RT	A1	A1	A1	A1
12_ZA_WC_OV_5_RT	B	B	B	B
13_ZA_WC_CW_50_RT	B	B	B	B
11_ZA_WC_CW_7_RT	B	B	B	B
12_ZA_WC_CW_20_RT	A1	A1	A1	A1
13_ZA_WC_CW_52_RT	C	C	A1,C	C
14_ZA_WC_WC_11_RT	Unclassified	C	C	C
12_ZA_WC_CW_42_RT	Unclassified	C	C	C
14_ZA_WC_CW_43_RT	Unclassified	B	C	C
14_ZA_WC_CW_45_RT	Unclassified	C	C	C

Table 4.5: IN fragment region analysis of non-subtype C sequences using various online tools.

SEQUENCE CODE	REGA	RIP	SCUEAL	HIVDRD
13_ZA_WC_OV_13_IN	A1	A1	A1	A1
12_ZA_WC_WC_8_IN	B	B	B	B
12_ZA_WC_OV_15_IN	B	B	B	B
12_ZA_WC_OV_16_IN	B	B	B	B
13_ZA_WC_CW_54_IN	B	B	B	B
14_ZA_WC_WC_16_IN	B	B	B	B
13_ZA_WC_CW_20_IN	Unclassified	C	C	C
12_ZA_WC_CW_37_IN	Unclassified	B	B	B
14_ZA_WC_CW_12_IN	Unclassified	C	C	C
12_ZA_WC_CW_40_IN	A1	C	A1	A1
13_ZA_WC_OV_21_IN	Unclassified	C	C	C





C: HIVDRD

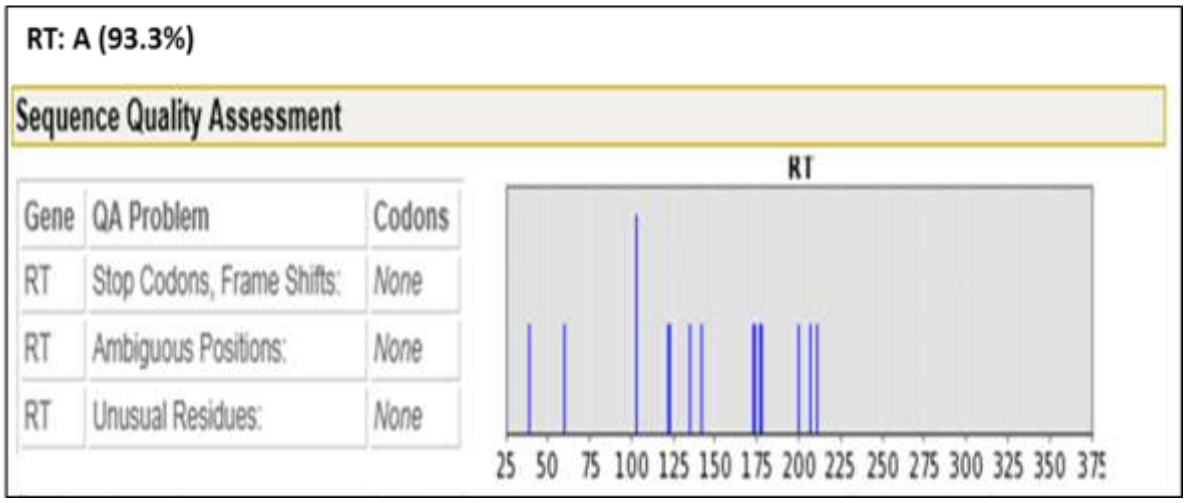


Figure 4.5: Analysis of sample 12_ZA_WC_CW_20_RT using three online HIV-1 subtyping tools.

Example of one of the sample (12_ZA_WC_CW_20_RT) from Cape Winelands district that was classified as subtype A1, using three different online tools; A: REGA, B: RIP, and C: HIVDRD. All the three HIV-1 online subtyping tools classified the sequence as subtype A1.

4.7.1. Recombinant Identification Program (RIP).

All PR, RT, and IN Sequences were analyzed for subtyping using RIP; last accessed 18/06/2016. The analysis of all 3 *pol* fragments region identified 14 Non-C subtypes. A summary of the results are shown in Table 4.3 to Table 4.5.

4.7.2. Rega 3.0 subtyping analysis.

All sequences were analyzed with the Rega 3.0 subtyping tool used to analyses the PR, RT, and IN gene fragment and classify all the sequences according to the HIV-1 subtypes, last accessed 03/07/2016 . Of 59/62 (95.2%) PR, 93/103 (90.2%) RT and 82/93 (88.2%) IN were identified as Subtype C respectively. However, three (4.8%) PR, four (3.9%) RT and four (4.3%) IN sequences designated with Unclassified report respectively, followed by one samples for PR identified as subtype A1. For the RT sequences four sample classified as subtype B, two sequences as subtype A1 and for the IN one as subtype A1, with five as subtype B.

4.7.3. SCUEAL analysis

SCUEAL identified three patient sequences for the RT gene as subtype A1 and four RT sequences as subtype B. For the IN two sequences were identified as subtype A1 and six sequences were characterized as subtype B. The remaining sequences for PR and RT were

identified as subtype C together with all the remaining sequences sample for IN were identified as subtype C, last accessed on the 13/06/2016.

4.7.4. Stanford University HIV Drug Resistance Database.

Stanford University HIV Drug Resistance Database, last accessed on the 18/07/2016, only identified two sequences for the RT as subtype A1 and two sequences for the IN as subtype A1. However, the database also identified four B subtype for the RT and six B subtypes for the IN. All the remaining sequences for RT, PR and IN classified as subtype C.

4.8. Sequence alignments and phylogenetic analysis

4.8.1. Sequence alignment

In order to study the evolutionary relationships of HIV sequences obtained during the course of the study, we aligned the study sequences with the group M reference sequences from the Los Alamos National Laboratory (LANL) HIV Sequence Database. Multiple alignments were generated to construct neighbour-joining trees. Phylogenetic analyses were used to validate and confirm the results obtained from different online subtyping tools.

4.8.2. Construction of partial *pol* neighbour-joining tree.

4.8.2.1. Construction of partial PR neighbour-joining tree

A neighbour-joining tree (Figure 4.6) of PR was constructed using sequences obtained from the LANL database (http://www.hiv.lanl.gov). The analysis included 62 nucleotides sequences from this study. The majority of the sequences (n = 61; 98.4%) clustered with subtype C reference sequences, except one sequence (15_ ZA _WC_ CW_ 36_ PR) that clustered with the subtype H reference sequence. This does not correspond with the online subtyping tools that identified this sequence as subtype C. Due to the short DNA sequence generated for PR in many of the cases (less than 200 bp), the majority of sequences (n =54; 48.3%) were excluded to construct a phylogenetic tree. The alignment was based on HXB2 position 2145 to 2635.

4.8.2.2. Construction of partial RT neighbour-joining tree

In Figure 4.7, the RT phylogenetic tree is shown, 101 (98.1%) of sequences grouped with subtype C reference sequences. While 1/103 (0.95%) sequence (12_ ZA_ WC_ CW _42 _RT) clustered with the subtype B reference sequences. In addition, another sequence (12_ ZA_ WC_ CW_ 20_ RT) clustered with subtype A1 reference sequences and corresponded

with the Rega 3.0, RIP, SQUEAL and Stanford HIVDRD analysis. The alignment was based on HXB2 position 2668 to 3434.

4.8.2.3. Construction of partial IN Neighbour-joining tree

The IN phylogenetic tree is depicted in Figure 4.8. The tree identified four cluster groups. The majority of sequences 101(92.4%) clustered with subtype C reference sequences. However, three sequences (2.7%) clustered with subtype B (11_ZA_WC_CW_16_IN), (13_ZA_WC_CW_15_IN), (15_ZA_WC_CW_26_IN) and another three sequences (12_ZA_WC_CW_22_IN), (12_ZA_WC_CW_30_IN) (14_ZA_WC_WC_1 IN) clustered with subtype A1 reference sequences. In addition, one sample (12_ZA_WC_CW_5_IN) grouped as subtype D. The alignment was based on HXB2 position 4351 to 5069.

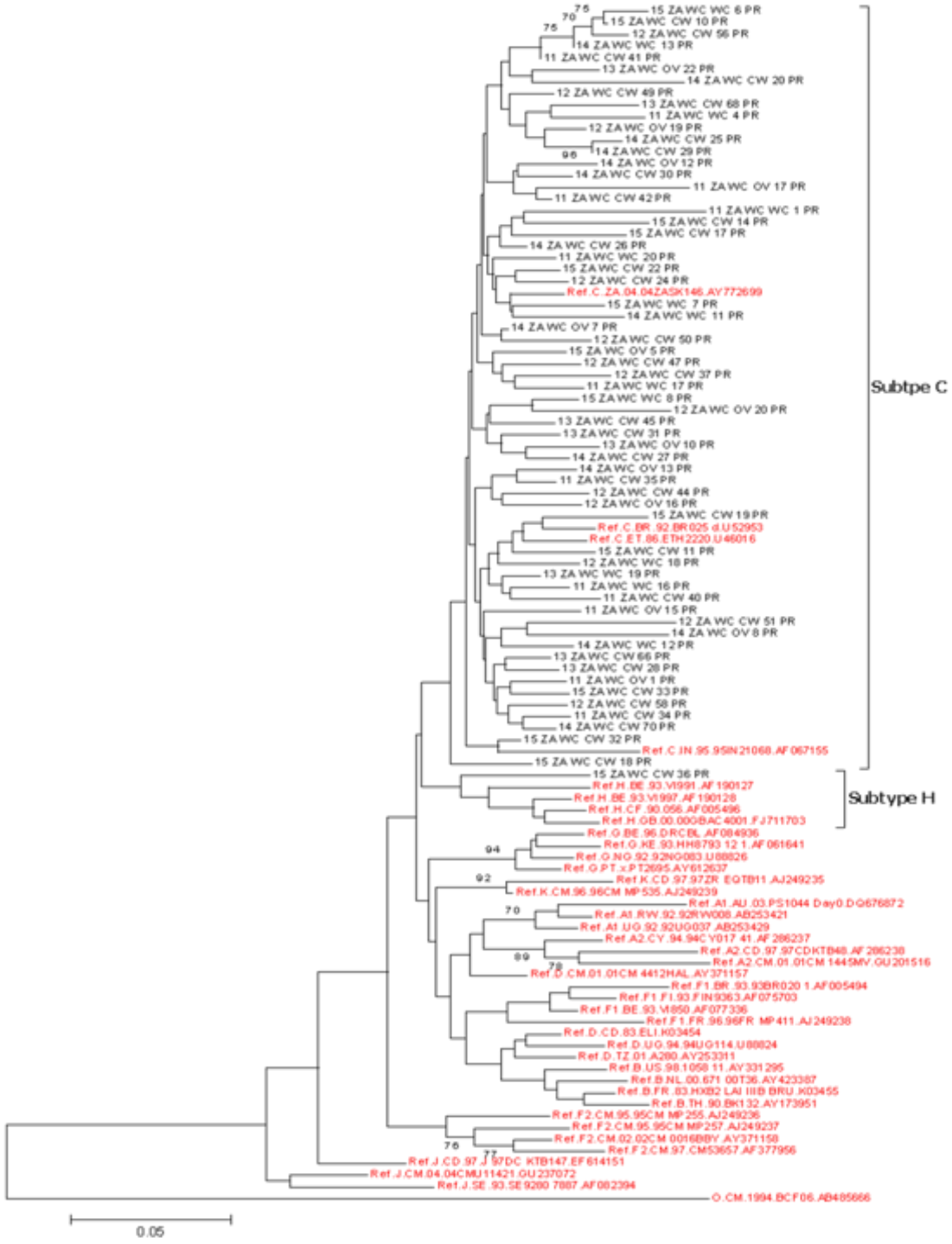


Figure 4.6: Neighbour Joining tree of PR reference and query sequences.

The phylogenetic tree was constructed in Mega 6v.1, with Kimura 2 parameter and consisted of cohort and query sequences approximately 205 bp in length. Bootstrapping was performed with 1000 replicates with significant values indicated on tree. Reference sequences are indicated in red colour and the study sequences are indicated in black colour. Majoriy of the sequence are classified as subtype C, with one sequence classified as subtype H.

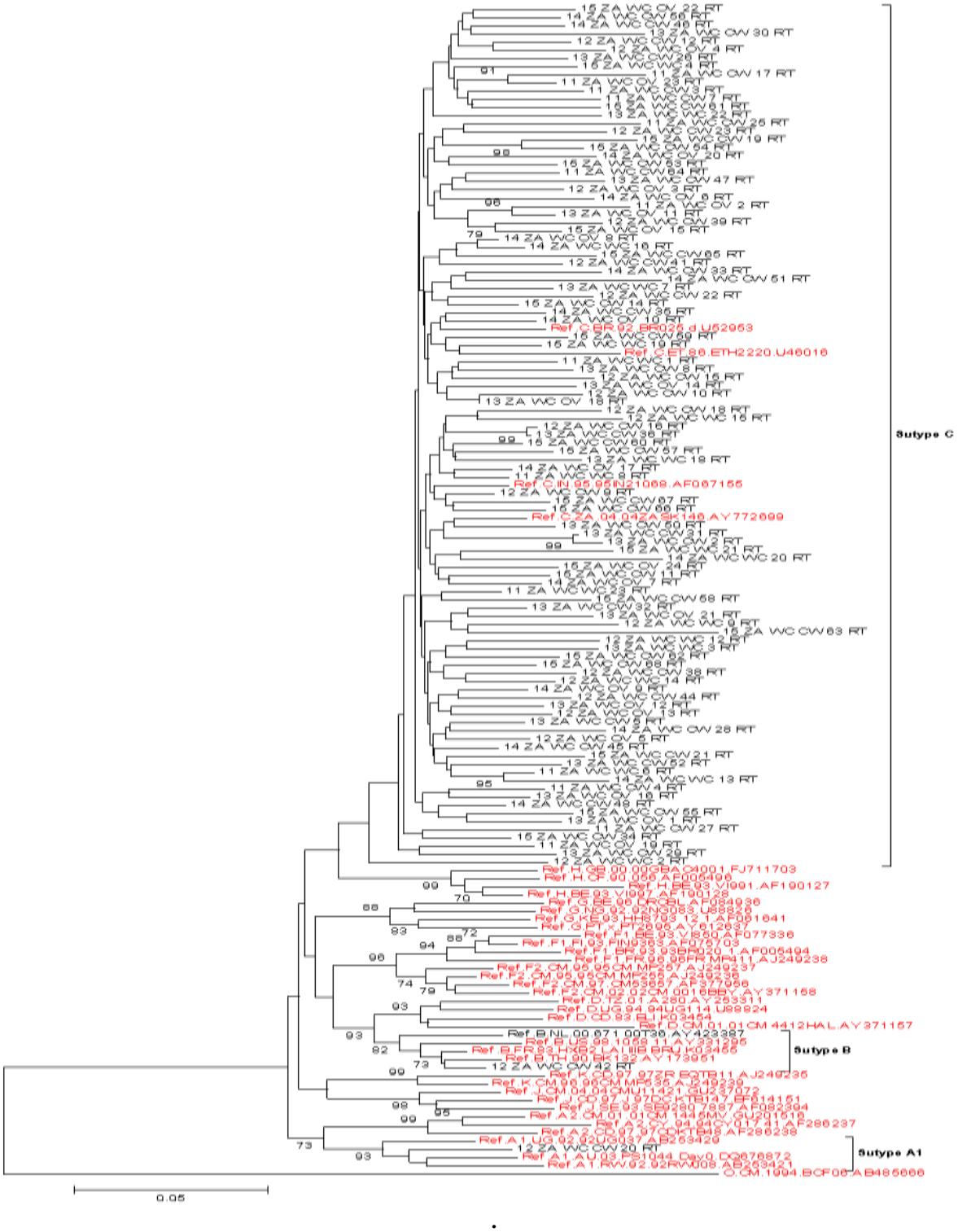


Figure 4.7: Neighbour Joining tree of RT reference and query sequences.

The phylogenetic tree was constructed in Mega 6v.1 ,with Kimura 2 parameter and consisted of cohort and query sequences approximately 520 bp in length. Bootstrapping was performed with 1000 replicates with significant values indicated on tree. Reference sequences are indicated in red colour and study sequences in black colour. Majority of sequence are classified as subtype C, with two sequences identified as non-C subtype.

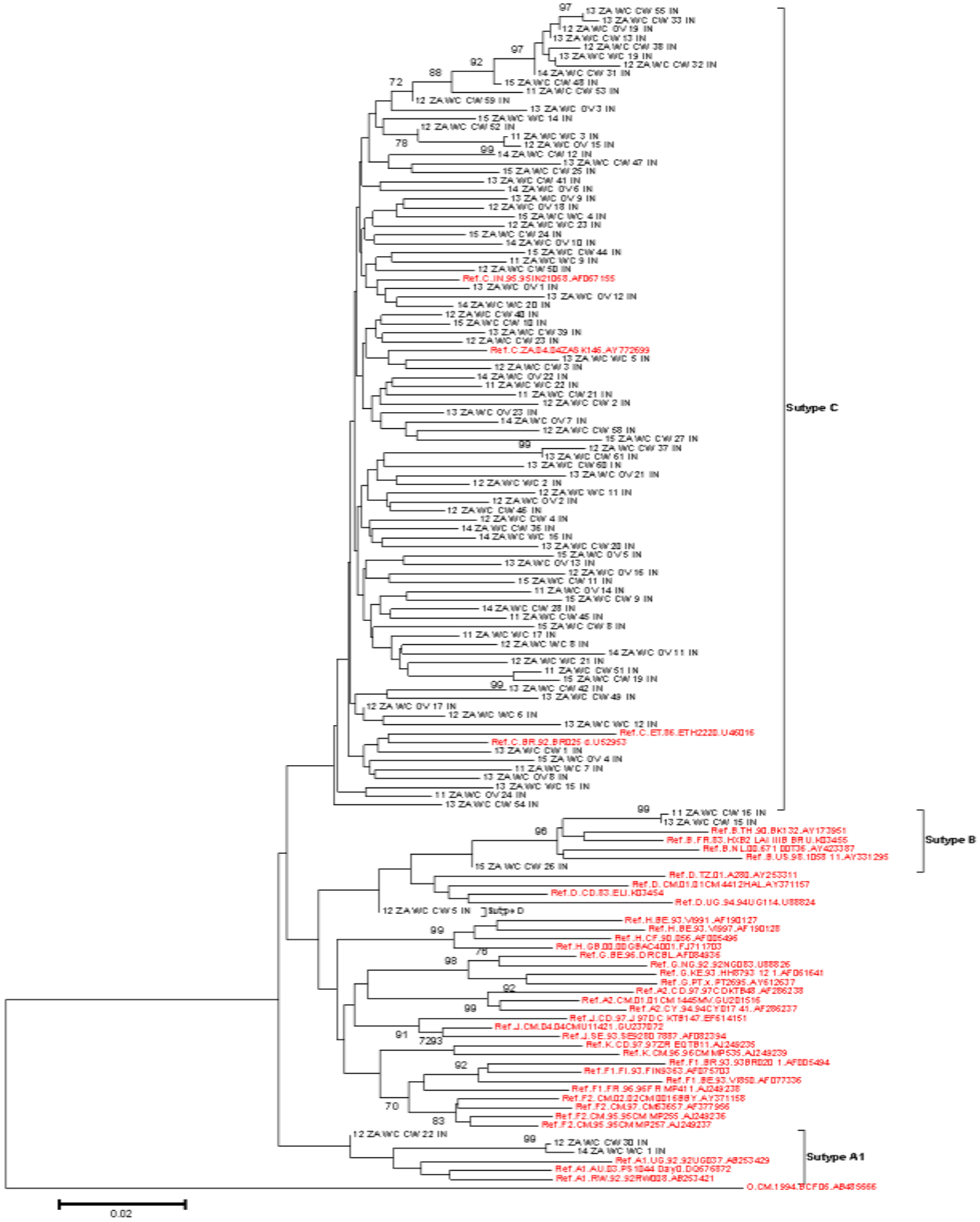


Figure 4.8: A Neighbour Joining tree of IN reference and query sequences.

Each query sequence are indicated in black colour. Bootstrapping was performed with 1000 replicates with significant values indicated on tree. Bootstrap value of greater than 70% is indicated on the condensed tree. Reference sequences are indicated in red colour and study sequences in black colour. Majority of sequence are classified as subtype C, with four sequences identified as non-C subtype.

4.9. Analyses of RAMs

As we used the *pol* sequences to study HIV-1 diversity, we could subsequently use the same sequences to look for RAMs amongst our sequences generated. The analyses of HIV-1 RAMs were based on the *pol* gene targeting the PR, RT, and IN gene fragment that are commonly known to be the regions that are conserved with mutations that act against the effectiveness of HIV-1 cART. The PR, RT, and IN sequences were used for the analysis of HIV-1 RAMs in the three antiretroviral classifications namely: PIs, RTIs (NRTIs and NNRTIs) and INSTs using the Stanford HIV-1 drug resistance database. The Stanford HIV-1 drug resistance database did not detect any major RAMs against PI and IN inhibitors. Minor RAMs detected was N88D, detected in 4 PR (3.7%) that cause low-level resistance especially against Nelfinavir, Figure 4.9 and Table 4.7. From the 93 IN sequences screened 15 (16.1%) had minor RAMs namely E157Q, G163EK, H51H, V151IV, L74I, L74IM, F121L, T97A, V151I, and V151IV, with only 9 (9.6%) of RAMs detected cause potential-low-level resistance against DTG, EVG and RAL and 6 (6.4%) of RAMs detected were found to cause low-level resistance against EVG and RAL, Figure 4.11 and Table 4.7. Of the 103 *pol*, RT sequences screened for RAMs 39 patients (37.8%) were confirmed on first line therapy and the remaining regimens were not indicated by the medical practitioner on the patient HIV-1 viral load request form. HIV-1 viral load cohort treatment is given in APPENDIX B. In 63 (61.7%) of the screened viral load patient sequences, major RTI RAMs were detected (Figure 4.10). 34 (48.7%), of which first line therapy regimens were indicated on the patient HIV-1 viral load request form the Stanford HIV-1 database detected no RAMs against those sequences. Mutations related with resistance to NRTI were observed in 39 (36.1%) patients, M184V 31 (39.7 %), and K65R 24 (30.7%) being the most common ones. Seventy-one (63.5%) with non-NRTIs (NNRTIs) RAMs with K103N 47 (36.1 %) and V106M 20 (15.4 %) being the most common ones, summarized in Table 4.6.

Table 4.6: List of all major drug resistance mutations detected against the NNRTI and NRTI

HIVDRM (NNRTI)	N=Mutations (127)	%	HIVDRM (NRTI)	N=Mutations (73)	%
E138A	7	5.3	D67N	3	3.8
E138K	3	2.3	K219E	2	2.6
F227L	4	3.1	K219EK	1	1.3
G190A	8	6.2	K219Q	2	2.6
G190Q	1	0.8	K65R	24	30.7
G190S	1	0.8	K70R	2	2.6
K101E	4	3.3	L74I	4	5.1
K103N	47	36.1	M184I	1	1.3
K103S	1	0.8	M184IM	1	1.3
L100I	3	2.3	M184V	31	39.7
V106M	20	15.4	T215F	2	2.6
V106MV	4	3.3			
V179D	2	1.5			
Y181C	5	3.8			
Y181V	2	0.8			
Y188CY	1	0.8			
Y188H	11	8.5			
Y188L	3	2.3			

Table 4.7 : List of all Minor drug resistance mutations detected against the PR and IN

HIVDRM (IN)	N=Mutations (30)	%	HIVDRM (PR)	N=Mutations (4)	%
E157Q	1	3.3	D30N	100	
G163	1	3.3			
H51HL	1	3.3			
V151IV	1	3.3			
L74I	11	36.6			
L74IM	1	3.3			
F121L	1	3.3			
L74LM	1	3.3			
T97A	2	6.6			
V151I	8	26.6			
V151IV	2	6.6			

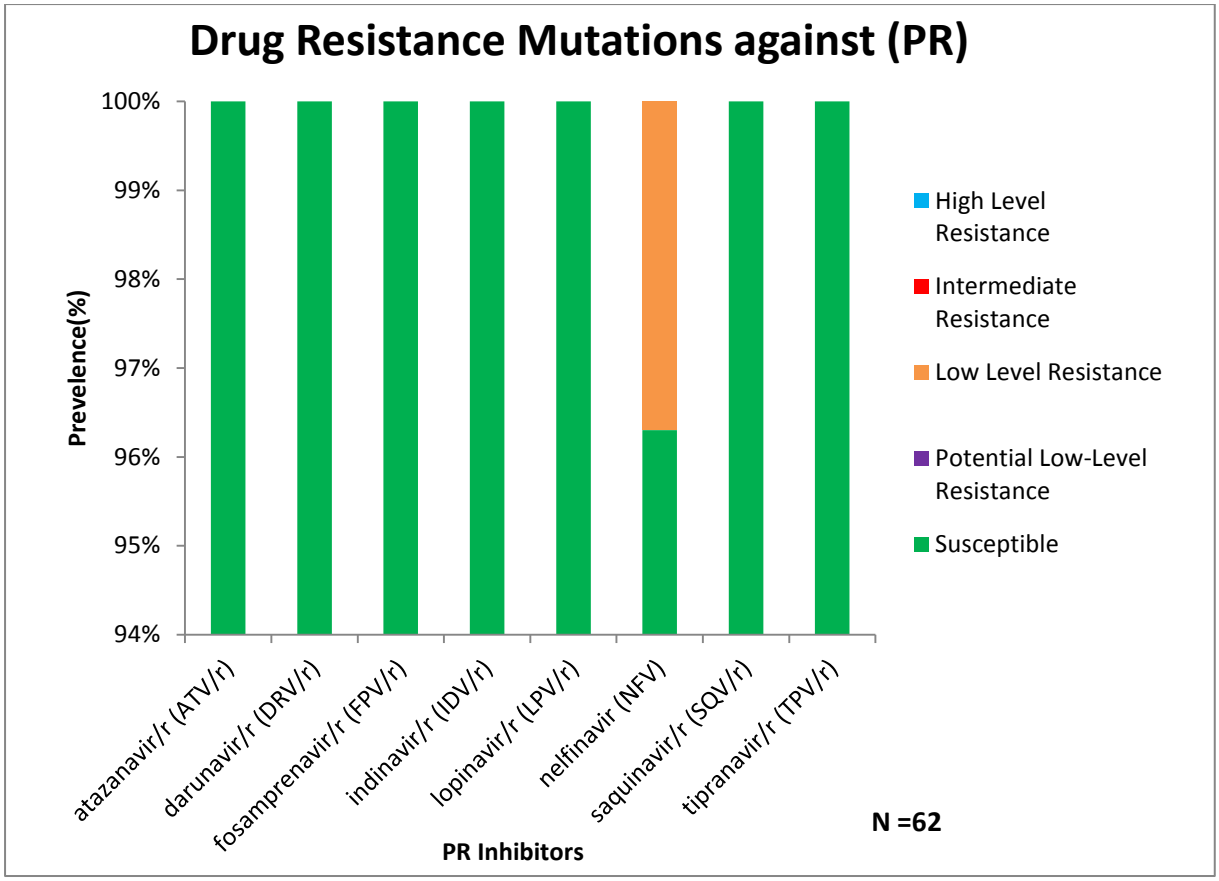


Figure 9: Observed drug resistance mutations against PR inhibitors.

Stanford University HIV Drug Resistance Database detected only minor RAMs (N88D) in 4 PR (3.7%), that that interfere with the efficacy of the cART (Nelfinavir).

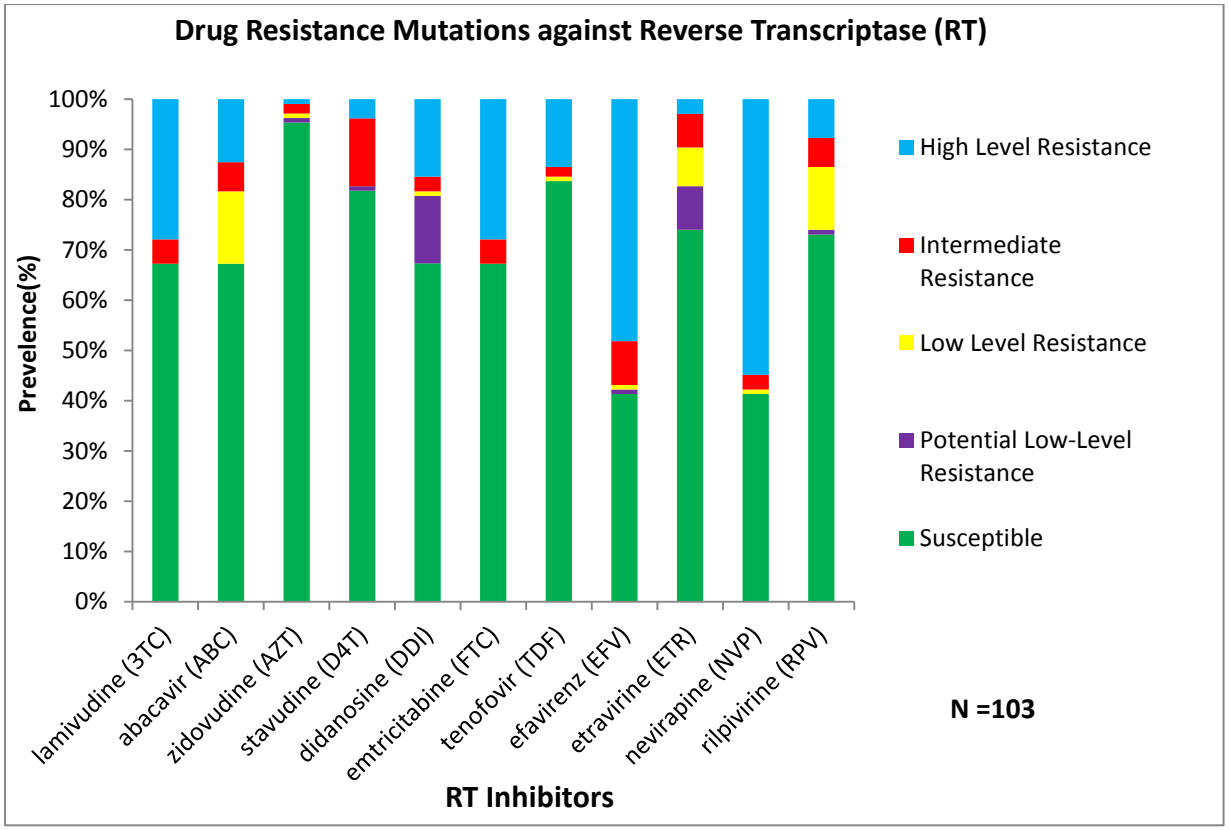


Figure 10: Observed drug resistance mutations against RT inhibitors.

Stanford University HIV Drug Resistance Database detected major RAMs in 63.5 % of the samples. The majority of mutations have HIVDRMs against EFV, NVP, FTC and 3TC, with M184V (39.7 %), K65R (30.7%) K103N (36.1 %) and V106M (15.4 %) mutations being the most common ones

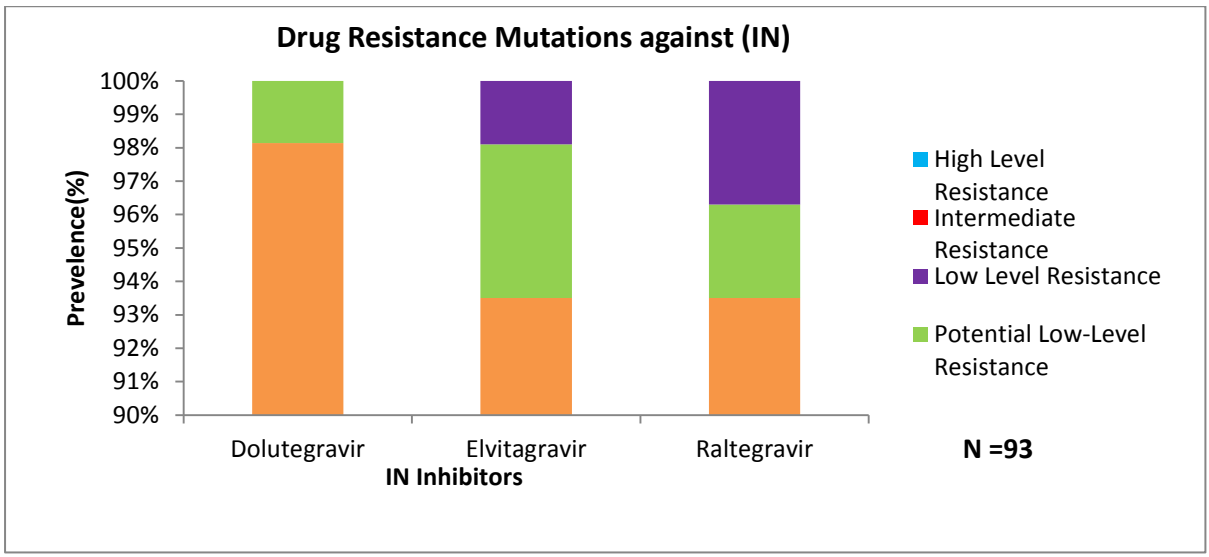


Figure 11: Observed drug resistance mutations against IN inhibitors.

In Figure 4.10, Minor mutations were detected in 16.1%, only potential-low level (9.6%) and low-level resistance mutations (6.4%) of samples had IN RAMs that interfere with the efficacy of the cART.

CHAPTER FIVE

Page

5. Discussion

CHAPTER FIVE

5. Discussion

5.1. Introduction

In this study, we investigated the HIV-1 *pol* PR, RT and IN regions for HIV-1 diversity and the prevalence of drug RAMs, from the Cape Winelands, West Coast and Overberg regions of the Western Cape Province of South Africa. The HIV-1 *pol* genome region is extremely valuable for researchers as it is known to harbor conserved areas needed for replication competence and has the regions with RAMs that interfere with the success of current cART regimens (De Oliveira *et al*., 2010). Most of the current therapeutic inhibitor drugs are designed targeting these regions of HIV-1 to inhibit HIV replication. HIV diversity, viral load monitoring and HIV drug resistance testing are important and essential tools to assess the success of cART. The findings of the study are discussed in this chapter.

5.1.1. HIV-1 in rural South Africa

There is a lack of data regarding HIV-1 diversity and the prevalence of HIV-1 RAM strains amongst treatment and naïve patients in remote and non-major metropolitan centers of South Africa, where resources, such as health clinics, are often limited. Most of the studies have been done in the major cities where accessibility of health care facilities by the researchers to collect samples is much convenience for them (Jacobs *et al*., 2008; Papathanasopoulos *et al*., 2010; Parboosing *et al*., 2011).

South Africa has the largest number (7.0 million) of HIV-1 infection worldwide with an estimated prevalence of 19.2 %, recorded in 2015 (UNAIDS Gap report, 2016). People who died from HIV-related illnesses were 180,000, 30.5% of all deaths in South Africa and 380,000 new HIV infections (HIV&AIDS statistics, 2015). High variations of HIV/AIDS are observed between the provinces of South Africa. Kwazulu-Natal has the highest prevalence of 40.1%, while HIV prevalence in the Western Cape Province is significantly lower (18.0%) than the national estimates (NDOH, 2014). The sexually active age group (males and females, aged 15 - 49 years) represents 55.7% of the Western Cape Province population. The HIV prevalence amongst this age group is 9.2%. Western Cape is one of the provinces that

showed high increases in the prevalence of HIV, from 5.0% in 2011 to 18.0% in 2014 (NDOH, 2014). The Cape Winelands, Overberg, and West Coast regions are remote district areas in the Western Cape Province. In the Province the Cape Winelands accounts for 15.0% of people living with HIV/AIDS, followed by Overberg (13.9%) and West Coast (9.6%), respectively. For many of the citizens residing in these districts access to clinical resources are still limited and only a few patients are able to access the health care facilities, as they are living in areas that are far from public health care centre.

A major primary mode of HIV transmission among this population is unprotected heterosexual contact (NDOH, 2011). Unemployment, lack of education, poverty, IDUs, alcohol abuse and sex workers are the other driving forces of HIV/AIDS epidemic in communities of South Africa. Despite the lowest prevalence of HIV in the Western Cape Province, of South Africa close monitoring and surveillance of HIV-1 is still crucial to control the epidemic. Many of the people from unmonitored remote areas who are living with HIV-1 migrate to major city Centre for employment and by doing so, the chances of HIV-1 spread increases. It was previously suggested that the labour system forced young men to abandon their communities and families, and by doing so may have been the route of HIV-1 transmission within South Africa (Hargrove, 2008; Wilkinson *et al*., 2015).

5.1.2. HIV-1 diversity in South Africa

In the early 1980s, HIV-1 subtypes B and D were responsible for the initial epidemic in South Africa, identified among homosexual contacts (Engelbrecht *et al*., 1995; Sher, 1989). Since the early 1990s, subtype C has been the driving force of the HIV-1 epidemic through heterosexual contact (van Harmelen *et al.,* 1997). HIV-1 Subtype C is the dominant cause of infection in South Africa and accounts for the majority of infections, with minor subtype B still circulating among the high-risk group of homosexuals, as well as among the heterosexual population (Jacobs *et al*., 2009; Jacobs *et al*., 2014).

In our study, we characterized the HIV-1 diversity from patients being monitored for the HIV-1 viral load from the Cape Winelands, West Coast and Overberg district of the Western Cape Province of South Africa. This was done using various HIV-1 online tools and phylogenetic tree analysis to validate and confirm the results obtained. Phylogenetic analysis of the three district sequences, Figures 4.6 - 4.8, reveals that HIV-1 subtype C (93.7%) is the most prevalent form in the Western Cape, with another ten non-C subtypes detected and

confirmed by phylogenetic trees, four subtype A1 (2.9%), four subtype B (2.9%), one subtype D (0.25%) and one subtype H (0.25%). To the best of our understanding, this might be the first time HIV-1 subtype H been identified in the Western Cape Province, of South Africa. Two recent studies by Jacobs *et al*., 2014 and Wilkinson *et al*., 2015, showed that HIV-1 subtype C is still the dominant subtype circulating in South Africa, with continuous introduction of non-C subtypes in the Western Cape Province of South Africa.

Other non-C subtypes have been sporadically identified (Loxton *et al*., 2005; Hemelaar *et al*., 2012, Bredell *et al*., 2002; Papathanasopoulos *et al*., 2010; Wilkinson *et al*., 2015, Wilkinson *et al*., 2009; Iweriebor *et al*., 2011, Jacobs *et al*., 2009). A follow-up study conducted by Jacobs *et al*., 2014 in our laboratory identified circulating BC strains in at least 4.6% of their study population, followed by a more recent study in our laboratory conducted by Wilkinson *et al*., 2015 that also identified other non-C subtypes: B, A, G, URF_AD and URF_AC circulating in the Western Cape (Wilkinson *et al*., 2015). Loxton *et al*., 2005, from our laboratory identified subtype D from full four length genome sequences. As the HIV-1 epidemic continues to increase in the country there is also a need to expand the research to rural communities and continue monitoring the current HIV-1 population to screen for circulating subtypes.

Thus, identification of subtype B sequences among the viral load patient sequences reveals that this subtype is still circulating in the community and probably to the rest of South Africa and has most likely entered the heterosexual community as well. The majority of non-C subtypes identified in the study came from the Cape Wineland , with the exception of one sequence that identified as Subtype A1 (14_ZA_WC_WC_1_IN) that originated from the West Coast (Figure 4.8). Two female patient sequences from the Ceres community of the Cape Winelands identified as A1 and clustered close to each other with a bootstrap support of 99%. In addition, two more female patient's sequences from the De Doorns community of Cape Winelands identified as subtype A1 and clustered close to each other, with a bootstrap support of 99%. A possible explanation might be that the sequences are from samples originating from epidemiologically linked patients that are possible related.

As a result of diverse viral strains and the identification of other non-C subtype strains, URFs and CRFs within our country and worldwide, it is important to regularly monitor the diversity and redistribution of viral strains. HIV diversity have a potential effects in clinical diagnosis, vaccine development, treatment and prevention (Peeters *et al.*, 2000; Lal *et al*., 2005).

Although genetic diversity of HIV-1 in South Africa has documented in several regions of South Africa (Gordon *et al.,* 2003; Hemelaar *et al.,* 2006; Bredell *et al.,* 1998; Pillay *et al.,* 2002), it is important to address the data regarding the development of drug resistance and HIV-1 diversity through periodic studies.

5.1.3. HIV-1 diagnostic in South Africa

VLs testing, as performed in our laboratory are used to monitor the treatment response and adherence problems early before the emergence of resistance. A subset of the patient with a detectable level of VL (>1000 copies/ml) after 3 months of initiation of therapy despite 100% adherence, additional adherence interventions is implemented and a repeat of VLs monitoring is done in 2-3 months (Hirsch *et al*., 2008). After 2-3 months of monitoring and the VLs remains above (>1000 copies/ml) such patients are referred for a drug resistance testing to identify and characterize drug resistance mutations and if the patients meet the criteria of switching the regimens, resistance will be a key to guide on the selection of new therapy regimen. During drug resistance testing, patient sequence are further used for the analysis of the HIV-1 diversity. HIV-1 diversity among subtypes it was reported to have a possible effect in yielding differential patterns of RAMs that interfere with the success of treatment (Ross *et al*., 2010) and associated with other clinical consequences such as escape of selected variants (Fischer *et al*., 2010) and progression of the disease (Lemey *et al.,* 2007; Salemi *et al*., 2001). New developed sequencing methods are focused on capturing low-frequency sequencing variants for characterization of intrahost diversity (Vrancken *et al*., 2010).

5.1.4. cART and the prevalence of drug RAMs

South Africa is the region most heavily affected by HIV/AIDS worldwide (UNAIDS gap report, 2016). Since the initiation of country wide cART in the country in April 2004, the HIV-1 infected population had benefited greatly from the cART program by reducing the mortality and morbidity rate in the country. South Africa has the largest antiretroviral program in the world with 3.4 million people on cART, this accounts for 48% of adults on cART (UNAIDS Gap Report, 2016). After the scale-up of cART in 2012, the mortality rate among children was reduced by 20% (HRSC, 2014). In the resource-limited settings of the Western Cape Province, the cART program was implemented as early as 2001, followed only in 2004 by the rest of the country (WHO, 2014). The margin of long-term cART success can be interrupted by the development of HIV-1 drug RAMs. The emergence of HIV-1 RAMs has been raised as the potential threat to the success of the cART worldwide (Bertagnolio *et*

al., 2013; Hamers *et al*., 2012). HIV-1 drug resistance testing is a useful tool for clinical care, such as the choice of active regimens before the patient's initiate therapy and also to help identify the presence of RAMs in patients prior to initiation of therapy. Resistance testing can be used to monitor the patients who are already on cART. HIV

The 205 HIV-1 positive patient samples selected for this study form part of a larger diagnostic monitoring viral load program in the Western Cape. All the patients receiving HIV-1 VLs are presumed on cART, as the Western Cape Consolidated Guidelines for HIV treatment recommend VLs monitoring only to patients who were on therapy for three months. However, only 58/205 (28.3%) of our study patients clearly indicated treatment regimens on patient data forms. Thus, we cannot confidently say if the patients were receiving cART, if at all. A high viral load assumes that cART is not being implemented properly, either through low adherence or suboptimal therapy that lead to viral resistance (Paredes *et al*., 2010). Our laboratory has previously developed an in-house HIV-1 resistance assay; however, these tests are not routinely done and only done on request (Claassen *et al*, personal communication – article in preparation).

In our, study we detected HIV-1 RT drug RAMs in 61.7% of the analyzed samples No major PI and IN mutations were detected. Of the *pol* RT sequences analysed seventy-one (61.7%) had major RTI RAMs, Table 4.6-4.7 and Figure 4.9 shows the results for RT, PR and IN analyzed. M184I/V causes high-level resistance to lamivudine (3TC) and emtricitabine (FTC) and low-level resistance to ddI and ABC. V106M is a nonpolymorphic mutation that causes high-level resistance to NVP and EFV while K65R causes intermediate/high-level resistance to TDF, ddI, ABC and d4T and low to intermediate-level resistance to 3TC and FTC. K65R increases susceptibility to AZT. However the detection of high level of K65R mutation in our study population might be attributed by more rapid in vivo selection of the K65R mutation in subtype C virus (Brenner BG *et al*., 2006). Thus far, in South Africa only few studies have investigated the antiretroviral RAMs in patients treated with TDF. One study by Sunpath *et al* 2012.showed that more than 65% of the patients on the TDF developed the K65R RAMs, whereas the second study found that K65R mutation was more prevalent in 46 % of the patients (Van Zyl *et al*., 2013). A third study which is substantially lower compared to the other two studies showed a K65R prevalence of 23% (Hoffman *et al*., 2013).A particular concern is that K65R RAMs has been found to cause cross-resistance mutations against most of the NRTIs (Johnson *et al*., 2011) and therefore interfere with the effectiveness of the

second line regimens.K103N causes high-level resistance to nevirapine (NVP) and efavirenz (EFV).

In addition detection of low-level RAMS to PR inhibitors and INSTIs in our study illustrates a possible threat to the future long-term management of HIV in South Africa contexts. The absence of major drug RAMs against Raltegravir have been documented in the recent studies conducted in South Africa (Gauteng) and in the other parts of the world. For example, Besong and Nwobegay, (Besong and Nwobegahay, 2013) observed minor RAMs and polymorphism in 7% of the sequences analysed. Fish *et al*. (Fish *et al*., 2010) detected polymorphism from 73 sequences obtained from South African HIV-1 infected patients. Similar results were also documented in Mozambique (Oliveira *et al., 2012*). In a previous study conducted in our laboratory, where 140 naïve -treated patients samples were received from 14 different clinics from the Cape Metropole area of South Africa during the period of 2002 to 2004 before the large scale increase of access to cART in the Western Cape, the prevalence of drug resistance mutations in the treatment naïve population was <5% (Jacobs *et al*., 2009). In another study in Cape Town, Western Cape the prevalence of major drug resistance mutations in naïve patients was 2.5% (Orrell *et al*., 2009) and in children born of drug-experienced mothers (Van Zyl *et al*., 2010). Our current study shows a prevalence of RAMs from small population size of three districts around the western Cape province of south Africa ,this indicate that the more the population size and expansion in samples collection from different regions within the country, true prevalence and reliability of results will be relied upon, because our results shows us not to ignore the resistance testing of HIV population from disadvantaged settings, since most of the studies are done in the large urban centres areas, where researchers have access to clinics and patient samples.

5.1.5. Strength of the study

The work describes the first data of HIV-1 diversity and RAMs of the VL sample sequences from three remote areas of the Western Cape Province of South Africa. In our laboratory, we have access to patient's VLs samples for further monitoring of cART outcome and characterization of HIV-1 for diversity using modern molecular techniques and analysis. We were able to identify the presence of other non-C subtypes causing infection in the Western Cape remote areas. With our analyses, we found a high prevalence of RTIs, despite having limited therapy treatment data for these patients we further included analysis of the IN genome region and we detected low-level RAMs against IN inhibitors drugs (DTG, EVG and

RAL) before third line therapy in South Africa. Thus, we already have optimized protocol for IN genotyping. Third line therapy treatment is not yet available in the low-resource setting and only available in high-income countries or through the private health sector. Therefore, our IN results give us insights and advantages knowledge of what to expect with regards to the emergence of RAMs in future when third line therapy is introduced in our country. This is a unique, important and valuable study as it bridges the lack of data regarding the emergence of RAMs and the increasing number of non-C subtypes in the remote areas of the Western Cape Province.

5.1.6. Limitations

A major limitation of this study is the lack of HIV-1 resistance data information available for all of the patients. In many cases the viral load request forms were filled incomplete, thus we could not make informative conclusions about RAMs and treatment outcomes. Due to ethical, reasons we could also not relay the findings of the study directly back to the patients or clinicians. However, the findings and recommendations from this study will be summarized for peer review publication. Due to time constraints and the high cost of laboratory techniques and lack of funding, we couldn't add more samples and district to the study and we could not as yet include patient follow-up the patients for additional samples. Another limitation was the fact that DNA amplification and sequencing was not successful in about 30% of the samples. This is most likely due to the high sequence variability seen in HIV, thus primer mismatches lead to inefficient PCR amplification and DNA sequencing.

5.1.7. Prospective work and recommendations

In future, we will include the use of Next Generation Sequencing (NGS) technology to help characterize our HIV strains. NGS could be especially useful to study non-Cs, CRFs and URFs in South Africa. NGS can also potentially be used for diagnostic HIV-1 resistance testing as being done in many first world laboratories already. NGS has been shown to detect minority variants that might be of potential clinical significant to hamper the success of the ARV treatment. In addition, resistance testing should be recommended before the initiation of therapy and to those who are already failing therapy.

5.1.8. Conclusion

With the continuous identification of non-C HIV-1 subtypes and drug RAMs in the country, there is a need to evaluate the effect of subtype diversity s on the success of cART. Studies have previously suggested that HIVDRM might be subtypes specific in certain cases. Optimally, a baseline resistance test should be performed for each patient before cART is initiated to ensure the chances of successful and complete virological suppression of HIV-1. This remains logistically challenging for the South African environment with so many people infected with HIV/AIDS. A resistance test should be available for patients failing therapy and we should continue to monitor the influence of RAMs and the diversity of subtypes on the HIV-1 epidemic in the rest of the country. It is important to understand the relation between the HIVDRM and geographic diversity of the epidemic.

CHAPTER SIX

6. REFERENCES

Apetrei C, Kaur A, Lerche NW, Metzger M, Pandrea I, Hardcastle J, Falkenstein S, Bohm R, Koehler J, Traina-Dorge V. Molecular epidemiology of simian immunodeficiency virus SIVsm in U.S. primate centers unravels the origin of SIVmac and SIVstm. J Virol. 2005; 79: 8991–9005.

Apetrei C, Robertson DL and Marx PA. The history of SIVS and AIDS: epidemiology, phylogeny and biology of isolates from naturally SIV infected non-human primates (NHP) in Africa. Front Biosci. 2004; 9: 225-254.

Armstrong KL, Lee TH, Essex M. Replicative capacity differences of thymidine analog resistance mutations in subtype B and C human immunodeficiency virus type 1. J Virol. 2009; 83: 4051-4059.

Badley AD. In vitro and in vivo effects of HIV protease inhibitors on apoptosis. Cell Death Differ. 2005; 12: 924-931.

Barre-Sinoussi F, Chermann JC, Rey F, Nugeyve MT, Chamaret S, Gruist J, Dauguet C, Axler-Blin C, Vezinet-Brun F and other authors. Isolation of a T- lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science. 1983; 220: 868-871.

Barré-Sinoussi F, Ross AL and Jean-François D. Past, present and future: 30 years of HIV research. Nature Reviews Microbiology. 2013; 11: 877–883.

Barth R, Wensing A, Tempelman H, Moraba R, Schuurman R, Hoepelman A. Rapid accumulation of nonnucleoside reverse transcriptase inhibitor-associated resistance: Evidence of transmitted resistance in rural South Africa. AIDS. 2008; 22: 2210.

Bennett DE, et al. Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. PLoS ONE. 2009; 4: e4724.

Bennett DE. HIV genetic diversity surveillance in the United States. J Infect Dis. 2005; 1929: 4–9.

Bertagnolio S, Perno CF, Vella S, Pillay D. The Impact of HIV Drug Resistance on the Selection of Firstand Second-Line ART in Resource-Limited Settings. J Infect Dis. 2013; 207 Suppl 2: S45–8.

Besong PO and Nwobegahay J. Genetic Analysis of HIV-1 Integrase Sequences from Treatment Naive Individuals in Northeastern South Africa. Int. J. Mol. Sci. 2013, 14: 5013-5024.

Bredell H, Hunt G, Casteling A et al. HIV-1 subtypes A, D, G, AG and unclassified sequences identified in South Africa. AIDS Res Hum Retroviruses. 2002; 18: 681-683.

Bredell H, Williamson C, Sonnenberg P, Martin DJ, Morris L. Genetic characterization of HIV type 1 from migrant workers in three South African gold mines. AIDS Res Hum Retroviruses. 1998; 14: 677–684.

Brenner BG. "Resistance and viral subtypes: how important are the differences and why do they occur?." Current Opinion in HIV and AIDS.200; 2: 94–102.

Brenner BG, Oliveira M, Doualla-Bell F, Moisi DD, Ntemgwa M, et al. HIV-1 subtype C viruses rapidly develop K65R resistance to tenofovir in cell culture. AIDS. 2006; 20: F9-13.

Briggs JA. Briggs T, Wilk R, Welker HG, Krausslich SD. Fuller Structural organization of authentic, mature HIV-1 virions and cores. EMBO J. 2003; 22: 1707-1715.

Buonaguro L, Tornesello ML, Buonaguro FM. Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications. Journal of Virology. 2007; 81: 10209–10219.

Carr J. viral diversity as a challenging to HIV-1 Vaccine development. current opinion in HIV AIDS. 2006; 1: 294-300.

Ceccherini-Silberstein F, Malet I, D'Arrigo R, Antinori A, Marcelin AG, Perno CF. Characterization and structural analysis of HIV-1 integrase conservation. AIDS Rev. 2009; 11: 17-29.

Chiu TK and Davies DR. Structure and function of HIV-1 integrase. Curr. Top. Med. Chem. 2004; 4: 965–977.

Cichutek K and Norley S. Lack of immune suppression in SIV infected natural hosts. AIDS 1993; 7: 25-35.

Clapham PR and McKnight A. Cell surfce receptors, virus entry and tropism of primate lentiviruses. J Gen Virol. 2002; 83: 1809–1829.

Clavel F, Hance AJ. HIV Drug Resistance. N Engl J Med. 2004; 350: 1023-1035.

Coffin JM, Haase A and Levy JA. What to call the AIDS virus?. Nature. 1986a; 321: 10.

Coffin JM, Haase A, Levy JA, Montagnier L, Oroszlan S, Teich N, Temin H, Toyoshima K, Varmus H, Vogt P and Weiss R. Human immunodeficiency viruses. Science. 1986b; 232: 697.

Collier AC, Coombs RW, Schoenfeld DA, Bassett RL, Timpone J, Baruch A, Jones M, Facey K, Whitacre C, McAuliffe VJ. Treatment of human immunodeficiency virus infection with saquinavir, zidovudine, and zalcitabine. AIDS Clinical Trials Group. N Engl J Med. 1996; 334: 1011–1017.

Cozzi-Lepri A, Phillips AN, Martinez-Picado J. Rate of accumulation of thymidine analogue mutations in patients continuing to receive virologically failing regimens containing zidovudine or stavudine: Implications for antiretroviral therapy programs in resource-limited settings. J Infect Dis .2009; 200: 687-697.

Cozzi-Lepri A, Phillips AN. The rate of accumulation of nonnucleoside reverse transcriptase inhibitor (NNRTI) resistance in patients kept on a virologically failing regimen containing an NNRTI. HIV Med. 2012; 13: 62-72.

Craigie R and Bushman FD. HIV DNA Integration. Cold Spring Harb Perspect Med. 2012 ; 2: a006890.

Craigie RJ. HIV Integrase, a Brief Overview from Chemistry to Therapeutics Biol. Chem. 2001; 276: 23213–23216.

Curran JW, Lawrence DN, Jaffe H, Kaplan JE, Zyla LD, Chamberland M, Weinstein R, Lui KJ, Schonberger LB and Spira TJ. Acquired immunodeficiency syndrome (AIDS) associated with transfusions. N. Eng J Med. 1984; 310: 69-75.

D'Aquila R, Schapiro J, Brun-Vézinet F, et al. Drug resistance mutations in HIV-1. Top HIV Med. 2002; 10: 21–25.

D'Aquila RT, Hughes MD, Johnson VA, Fischl MA, Sommadossi JP, Liou SH, Timpone J, Myers M, Basgoz N, Niu M, Hirsch MS. Nevirapine, zidovudine, and didanosine compared with zidovudine and didanosine in patients with HIV-1 infection. A randomized, double-blind, placebo-controlled trial. National Institute of Allergy and Infectious Diseases AIDS Clinical Trials Group Protocol 241 Investigators. Ann Intern Med. 1996; 124: 1019-30.

De Clercq E. The role of non-nucleoside reverse transcriptase inhibitors (NNRTIs) in the therapy of HIV-1 infection. Antiviral Res. 1998; 38: 153–179.

de Oliveira T, Engelbrecht S, Janse van Rensburg E, Gordon M, Bishop K, zur Megede J, Barnett SW, Cassol S. Variability at human immunodeficiency virus type 1 subtype C protease cleavage sites: an indication of viral fitness?. J Virol. 2003; 77: 9422-9430.

de Oliveira T, Pillay D, Gifford RJ, for the UK Collaborative Group on HIV Drug Resistance. The HIV-1 Subtype C Epidemic in South America Is Linked to the United Kingdom. PLoS ONE. 2010; 5: e9311.

de Silva E, Stumpf MP. HIV and the CCR5-Delta32 resistance allele. FEMS Microbiol Lett. 2004; 24: 1–12.

Department of Health. The National antenatal sentinel HIV and Syphilis prevalence survey, South Africa, 2014.

Dezzutti CS, Guenthner PC, Green TA, Cohen OJ, Spira TJ, Lal RB. Stromal-derived factor-1 chemokine gene variant is associated with the delay of HIV-1 disease progression in two longitudinal cohorts. AIDS. 2000; 14: 894–896.

di Marzo Veronese F, Copeland TD, DeVico AL, Rahman R, Oroszlan S, Gallo RC, Sarngadharan MG. Characterization of highly immunogenic p66/p51 as the reverse transcriptase of HTLV-III/LAV. Science. 1986; 231:1289-91.

Draper B, Abdullah F. A review of the prevention of mother-to-child transmission programme of the Western Cape provincial government, 2003–2004. S AfrMed J. 2008; 98: 431–4.

Easterbrook PJ. Long-term non-progression in HIV infection: definitions and epidemiological issues. J Infect. 1999; 38: 71–73.

Egger M, Boulle A. Population effect of scaling up ART in resource-poor settings. Lancet. 2008; 371: 1558-1559.

Engelbrecht S, Laten J, Smith TL, J. van Rensburg E. Identification of env subtypes in fourteen HIV type 1 isolates from South Africa. AIDS Res Hum Retroviruses. 1995; 11: 1269–1271.

Engelman A, Bushman FD, Craigie R. Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex. EMBO J. 1993; 12 :3269-3275.

Erickson JW, Gulnik SV, Markowitz M. Protease inhibitors: resistance, cross-resistance, fitness and the choice of initial and salvage therapies. AIDS. 1999;13 Suppl A:S189-204.

Eshleman SH, Guay LA, Mwatha A, Cunningham SP, Brown ER, Musoke P, Mmiro F, Jackson JB. Comparison of nevirapine (NVP) resistance in Ugandan women 7 days vs 6-8 weeks after single-dose nvp prophylaxis: HIVNET 012. AIDS Res. Hum. Retroviruses. 2004; 20: 595–599.

Estrabaud E, Le Rouzic E, Lopez-Vergès S, Morel M, Belaïdouni N, Benarous R, Transy C, Berlioz-Torrent C, Margottin-Goguet F. "Regulated degradation of the HIV-1 Vpu protein through a betaTrCP-independent pathway limits the release of viral particles". PLoS Pathog. 2007; 3: 7.

European guidelines on HIV-1 drug resistance testing have been published in the February edition of the authoritative peer review journal AIDS. 2001; 15: 3.

Evering TH, Markowitz M, Raltegravir. (MK-0518): an integrase inhibitor for the treatment of HIV-1. Drugs Today. 2007; 43: 865–77.

Farquhar C, VanCott TC, Mbori-Ngacha DA, et al. Salivary secretory leukocyte protease inhibitor is associated with reduced transmission of human immunodeficiency virus type 1 through breast milk. J Infect Dis. 2002; 186: 1173-6.

Fauci AS, Pantaleo G, Stanley S, Weissman D. Immunopathogenic mechanisms of HIV infection. Ann. Intern. Med. 1996; 124: 654–663.

Fischer W, Ganusov V, Giorgi E et al. Transmission of single HIV-1 genomes and dynamics of early immune escape revealed by ultra-deep sequencing. PLoS ONE. 2010; 8: e12303.

Fish MQ, Hewer R, Wallis CL, Venter WD, Stevens WS, Papathanasopoulos M.A. Natural polymorphisms of integrase among HIV type 1-infected South African Patients. AIDS Res. Hum. Retroviruses. 2010; 26: 489–493.

Flandre P, Chappey C, Marcelin AG, et al. Phenotypic susceptibility to didanosine is associated with antiviral activity in treatment-experienced patients with HIV-1 infection. J Infect Dis. 2007; 195: 392-398.

Fowke KR, Nagelkerke NJ, Kimani J, et al. Resistance to HIV-1 infection among persistently seronegative prostitutes in Nairobi, Kenya. RESISTANCE TO HIV INFECTION 13, Lancet. 1996; 348: 1347–1351.

Fox J ,Castro H, Kaye S, McClure M, Weber JN, Fidler S. Epidemiology of non-B clade forms of HIV-1 in men who have sex with men in the UK. Epidemiology and Social: CONCISE COMMUNICATION. 2010; 24: 2397–2401.

Friedman-Kien AE, Laubenstein L, Marmor M, Hymes K, Green J, Ragaz A, Gottleib J, Muggia F, Demopoulos R and other authors. Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men - New York City and California. MMWR. 1981; 30: 305-308.

Gall A, Ferns B, Morris C, Watson S, Cotten M, Robinson M, Berry N, Pillay D, Kellam P.. Universal amplification, next-generation sequencing, and assembly of HIV-1 genomes. J Clin Microbiol. 2012; 50: 3838–3844.

Gallo R, Wong-Staal F, Montagnier L, Haseltine WA, Yoshida M.HIV/HTLV gene nomenclature. Nature. 1988; 333: 504.

Gallo RC, Salahuddin SZ, Popovic M, Shearer GM, Kaplan M, Haynes BF, Palker TJ, Redfield R, Oleske J and Safai B. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science. 1984; 224: 500-3.

Galvani AP, Novembre J. The evolutionary history of the CCR5-Delta32 HIV-resistance mutation. Microbes Infect. 2005; 7: 302-9.

Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, Cummins LB, Arthur LO, Peeters M, Shaw GM, Sharp PM, Hahn BH. Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. Nature. 1999; 397: 436–441.

Gao F, Yue L, White AT, et al. Human infection by genetically-diverse SIVsm related HIV-2 in West Africa. Nature. 1992; 358: 495-499.

Goff SP. Retroviral reverse transcriptase: synthesis, structure, and function. J Acquir Immune Defic Syndr. 1990; 3: 817-31.

Gordon M, De Oliveira T, Bishop K, Coovadia HM, Madurai L, Engelbrecht S, Janse van Rensburg E, Mosam A, Smith A, Cassol S: Molecular characteristics of human immunodeficiency virus type 1 subtype C viruses from KwaZulu-Natal, South Africa: implications for vaccine and antiretroviral control strategies. J Virol. 2003; 77: 2587-2599.

Gottlieb MS, Schroff R, Schanker HM, Weisman JD, Fan PT, Wolf RA and Saxon A. Pneumocystis carinii pneumonia and mucosal cadidiasis in previously healthy homosexual men: Evidence of a new acquired cellular immune deficiency. N. Eng J. Med. 1981; 305: 1425-1428.

Grossman Z, Vardinon N, Chemtob D, Alkan ML, Bentwich Z, Burke M, Gottesman G, Istomin V, Levi I, Maayan S, et al. Genotypic variation of HIV-1 reverse transcriptase and protease: comparative analysis of clade C and clade B. AIDS. 2001; 15: 1453-1460.

Grossman Z, Istomin V, Averbuch D, et al. Genetic variation at NNRTI resistance-associated positions in patients infected with HIV-1 subtype C. AIDS. 2004; 18: 909-915.

Gupta RK, Jordan MR, Sultan BJ et al. Global trends in antiretroviral resistance in treatment-naive individuals with HIV after rollout of antiretroviral treatment in resource-limited settings: a global collaborative study and meta-regression analysis. Lancet. 2012; 380: 1250–1258.

Hahn BH, Shaw GM, De Cock KM and Sharp PM. AIDS as a Zoonosis: Scientific and public health implications. Science. 2000; 287: 607-614.

Hamers RL, Kityo C, Lange JM, Wit TF, Mugyenyi P: Global threat from drug resistant HIV in sub-Saharan Africa. BMJ. 2012; 344: e4159.

Harrich D, Hooker B. Mechanistic aspects of HIV-1 reverse transcription initiation. Rev Med Virol. 2002; 12: 31–45.

Hatziioannou T, Del Prete GQ, Keele BF, Estes JD, McNatt MW, Bitzegeio J, Raymond A, Rodriguez A, Schmidt F, Mac Trubey C, Smedley J, Piatak M Jr, KewalRamani VN, Lifson JD, Bieniasz PD. HIV-1-induced AIDS in monkeys. Science. 2014; 344: 1401–1405.

Haubrich R, Demeter L. International perspectives on antiretroviral resistance: clinical utility of resistance testing: retrospective and prospective data supporting use and current recommendations. J Acquir Immune Defic Syndr. 2001; 26: Suppl 1:S51-S59.

Heinzinger NK, Bukinsky MI, Haggerty SA, Ragland AM, Kewalramani V, Lee MA, Gendelman HE, Ratner L, Stevenson M, Emerman M. The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells. Proc Natl Acad Sci U S A. 1994; 91: 7311-5.

Hemelaar J, Gouws E, Ghys PD ,Osmanov S. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. AIDS. 2006; 20: W13–W23.

Hemelaar J, Gouws E, Ghys PD, Osmanov S. Global trends in molecular epidemiology of HIV-1 during 2000-2007. AIDS. 2011; 25: 679–89.

Hemelaar J. The origin and diversity of the HIV-1 pandemic. Trends in Molecular Medicine. 2012; 18: 182-92.

Hirsch MS, Brun-Vezinet F, Clotet B, et al. Antiretroviral drug resistance testing in adults infected with human immunodeficiency virus type 1: 2003 recommendations of an International AIDS Society-USA Panel. Clin Infect Dis. 2003; 37: 113-28.

Hirsch MS, Brun-Vezinet F, D'Aquila RT, et al. Antiretroviral drug resistance testing in adult HIV-1 infection. JAMA. 2000; 283: 2417±2426.

Hirsch MS, Curran J. Human immunodeficiency viruses. In: Fields, editor. Virology. 4th ed. Philadelphia: Lippincott-Raven Publishers; 1990. p.1953–75.

Hirsch MS,Gunthard HF, Schapiro JM, Brun-Vézinet F, Clotet B, Hammer SF, Johnson, VA, Kuritzkes DR, Mellors JW, Pillay D, Yeni PG, Jacobsen DM, Richman DD. 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.

Hirsch V, Olmsted R, Murphy-Corb M, Pyrcell, Johnson P. An African primate lentivirus (SIVsm) closely related to HIV-2. Nature. 1989; 339: 389-392.

Ho D, Neumann AU, Perelson AS. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature. 1995; 373: 123–26.

Hoffmann C, Ledwaba J, Li JF, Johnston V, Hunt G, et al. Resistance to tenofovir-based regimens during treatment failure of subtype C HIV-1 in South Africa. Antivir Ther. 2013; 18: 915–920.

Hoffmann C, Charalambous S, Sim J, Ledwaba J, Schwikkard G, Chaisson R, Fielding K, Churchyard G, Morris L, Grant A. Viremia, resuppression, and time to resistance in human immunodeficiency virus (HIV) subtype C during first-line antiretroviral therapy in South Africa. Clin Infect Dis. 2009; 49: 1928–35.

Holguin A, Soriano V. Resistance to antiretroviral agents in individuals with HIV-1 non-B subtypes. HIV Clin Trials. 2002; 3: 403–411.

Hu Z and Kuritzkes DR. Altered viral fitness and drug susceptibility in HIV- 1 carrying mutations that confer resistance to nonnucleoside reverse transcriptase and integrase strand transfer inhibitors. J. Virol. 2014; 88: 9268–9276.

Huang H, Chopra R, Verdine GL, and Harrison SC. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. Science. 1998; 282: 1669-1675.

Human Sciences Research Council (HSRC) (2014) 'South African National HIV Prevalence, Incidence and Behaviour Survey 2012'.

Iweriebor BC, Bessong PO, Mavhandu LG, Masebe TM, Nwobegahay J, Moyo SR, Mphahlele JM. Genetic analysis of the near full-length genome of an HIV type 1 A1/C unique recombinant form from Northern South Africa. AIDS Res Hum Retroviruses. 2011; 27: 911–915.

Jacobo-Molina A, Ding J, Nanni RG, Clark AD Jr, Lu X, Tantillo C, Williams RL, Kamer G, Ferris AL, Clark P, et al. Crystal structure of human immunodeficiency virus type 1 reverse transcriptase A. 1993; 90: 6320–6324.

Jacobs GB, Laten A, van Rensburg E, Bodem J, Weissbrich B, Rethwilm A, Preiser W, Engelbrecht S. Phylogenetic diversity and low level antiretroviral resistance mutations in HIV type 1 treatment-naive patients from Cape Town, South Africa. AIDS Res Hum Retroviruses. 2008; 24: 1009–1012.

Jacobs GB, Loxton AG, Laten A, Robson B, Janse Van Rensburg E, et al. Emergence and diversity of different HIV-1 subtypes in South Africa, 2000–2001. J Med Virol. 2009; 81: 1852–1859.

Jacobs GB, Wilkinson E, Isaacs S, Spies G, de Oliveira T, et al. HIV-1 Subtypes B and C Unique Recombinant Forms (URFs) and Transmitted Drug Resistance Identified in the Western Cape Province, South Africa. PLoS ONE. 2014; 9: e90845.

James MN and Sielecki AR. Stereochemical analysis of peptide bond hydrolysis catalyzed by the aspartic proteinase penicillopepsin. Biochemistry. 1985; 2414: 3701-13.

Javanbakht H, An P, Gold B, et al. Effects of human TRIM5alpha polymorphisms on antiretroviral function and susceptibility to human immunodeficiency virus infection. Virology 2006; 354: 15-27.

Jin X, Brooks A, Chen H, et al. APOBEC3G/GEM15 (HA3G) mRNA levels associate inversely with human immunodeficiency virus viremia. J Virol. 2005; 7: 11513-6.

Johanson J, Abravaya K, Caminti W, Erickson D, Flanders R, Leckie G, Marshall E, Mullen C, Ohhashi Y, Perry R, Ricci J, Salituro J, Smith A, Tang N, Vi M and Robinson JA. New ultrasensitive assay for quantification of HIV-1 RNA in plasma. J.Virol.Methods. 2001; 95: 81-92.

Johnson VA, Calvez V, Gunthard HF, Paredes R, Pillay D, et al. 2011 update of the drug resistance mutations in HIV-1. Top Antivir Med. 2011; 19: 156–164.

Johnson MO, Catz S, Remien R, Rotheram-Borus M, Morin S, Charlebois E. the NIMH Healthy Living Project Team. Theory guided, empirically supported avenues for intervention on HIV medication nonadherence: Findings from the Healthy Living Project. AIDS Patient Care & STDS. 2003; 17: 645–656.

Kantor R & Katzenstein D. Polymorphism in HIV-1 non-subtype B protease and reverse transcriptase and its potential impact on drug susceptibility and drug resistance evolution. AIDS Reviews. 2003; 19: 25–35.

Kantor R, Zijenah L, Shafer R, et al. HIV-1 subtype C reverse transcriptase and protease genotypes in Zimbabwean patients failing antiretroviral therapy. AIDS Res Hum Retroviruses. 2002; 18: 1407–1413.

Kantor R. et al. Impact of HIV-1 subtype and antiretroviral therapy on protease and reverse transcriptase genotype: results of a global collaboration. PLoS Med. 2005; 2: e112.

Kantor R: Impact of HIV-1 pol diversity on drug resistance and its clinical implications. Curr Opin Infect Dis. 2006; 19: 594-606.

Karacostas V, Nagashima K, Gonda MA, Moss B. Human immunodeficiency virus-like particles produced by a vaccinia virus expression vector. Proc Natl Acad Sci U S A. 1989; 86: 8964-8967.

Kaul R, Dong T, Plummer FA, et al. CD8(+) lymphocytes respond to different HIV epitopes in seronegative and infected subjects. J Clin Invest. 2001; 107: 1303–1310.

Kaul R, Rutherford J, Rowland-Jones SL, et al. HIV-1 Env-specific cytotoxic Tlymphocyte responses in exposed, uninfected Kenyan sex workers: a prospective analysis. AIDS. 2004; 18: 2087–2089.

Kessler HH, Deuretzbacher D, Stelzl E, Daghofer E, Santner BI, Marth E. Determination of human immunodeficiency virus type 1 subtypes by a rapid method useful for the routine diagnostic laboratory. Clin Diagn Lab Immunol. 2001; 8: 1018–1020.

Kim SY, Byrn R, Groopman J, Baltimore D.Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: evidence for differential gene expression. J Virol. 1989; 63: 3708-13.

Knudsen TB, Kristiansen TB, Katzenstein TL, Eugen-Olsen J. Adverse effect of the CCR5 promoter -2459A allele on HIV-1 disease progression. J Med Virol. 2001; 65: 441-4.

Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA. Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science. 1992; 256: 1783-90.

Korber B, Muldoon M, Theiler J, Gao F, Gupta R, Lapedes A, Hahn BH, Wolinsky S and Bhuttacharga T. Timing the ancestor of the HIV-1 pandemic strains. Science. 2000; 288: 1789-1795.

Lal RB, Sekher C, Chunfu Y. Impact of genetic diversity of HIV-1 on diagnosis, antiretroviral therapy & vaccine development. Indian J Med Res. 2005; 121: 287–314.

Lanier ER, Ait-Khaled M, Scott J, et al. Antiviral efficacy of abacavir in antiretroviral therapy-experienced adults harbouring HIV-1 with specific patterns of resistance to nucleoside reverse transcriptase inhibitors. Antivir Ther. 2004; 9: 37-45.

Laskey SB & Siliciano RF. A mechanistic theory to explain the efficacy of antiretroviral therapy. Nature Reviews Microbiology. 2014; 12: 772–780.

Lemey P, Pond SLK, Drummond AJ et al. Synonymous substitution rates predict HIV disease progression as a result of underlying replication dynamics. PLoS Comput. Biol. 2007; 3: 282–292.

Lemey P, Pybus OG, Rambaut A, Drummond AJ, Robertson DL, Roques P, Worobey M, Vandamme AM. The Molecular Population Genetics of HIV-1 Group O. Genetics. 2004; 167: 31059–1068.

Lemey P, Pybus OG, Wang B, Saksena NK, Salemi M, and Vandamme AM. Tracing the origin and history of the HIV-2 epidemic. Proceedings of the National Academy of Science. 2003; 100: 6588–6952.

Levy JA, Hoffman AD, Kramer SM, Landis JA, Shimabukuro JM, Oshiro LS. Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. Science. 1984; 22: 840-2.

Lorenzi P, Opravil M, Hirschel B, et al. Impact of drug resistance mutations on virologic response to salvage therapy. Swiss HIV Cohort Study AIDS. 1999; 13: 17–21.

Loxton AG, Treurnicht F, Laten A, van Rensburg EJ, Engelbrecht S. Sequence analysis of near full-length HIV type 1 subtype D primary strains isolated in Cape Town, South Africa, from 1984 to 1986. AIDS Res Hum Retroviruses. 2005; 21: 410–413.

MacDonald KS, Embree J, Njenga S, Nagelkerke NJ, Ngatia I, Mohammed Z, Barber BH, Ndinya-Achola J, Bwayo J, Plummer FA. Mother-child class I HLA concordance increases perinatal human immunodeficiency virus type 1 transmission. J Infect Dis. 1998; 177: 551-6.

Marchand C, Johnson AA, Semenova E and Pommier Y. Mechanism and inhibition of HIV integration. Drug. Discov. Today: Disease Mechanism. 2006a; 3: 253-260.

Marconi V, Sunpath H, Lu Z, Gordon M, Koranteng-Apeagyei K, Hampton J, Carpenter S, Giddy J, Ross D, Holst H. 2008.Clin Infect Dis.Prevalence of HIV-1 drug resistance after failure of a first highly active antiretroviral therapy regimen in KwaZulu Natal, South Africa. 2008; 46: 1589-97.

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

Menéndez-Arias L. Targeting HIV: antiretroviral therapy and development of drug resistance. Trends Pharmacol Sci. 2002; 23: 381-8.

Miller MD, Margot N, Lu B, et al. Genotypic and phenotypic predictors of the magnitude of response to tenofovir disoproxil fumarate treatment in antiretroviral-experienced patients. J Infect Dis. 2004; 189: 837-846.

MMWR – December. Centers for Disease Control (CDC).'Unexplained Immunodeficiency and Opportunistic Infections in Infants-New York, New Jersey, California'. MMWR – Morbidity and Mortality Weekly Report. 1982b; 49: 665-667.

Nagashima KA, Thompson DA, Rosenfield SI,Maddon PJ, Dragic T, Olson WC. Human immunodeficiency virus type 1 entry inhibitors PRO 542 and T-20 are potently synergistic in blocking virus-cell and cell-cell fusion. J Infect Dis. 2001; 183: 1121-1125.

Nahmias AJ, Weiss J, Yao X, Lee F, Kodsi R, Schanfield M, Matthews T, Bolognesi D, Durack D & other authors. Evidence for human infection with an HTLV III/ LAV like virus in central Africa, 1959. Lancet. 1986: 1279-1280.

Nkenfou CN, Mekue LC, Nana CT, Kuiate JR. Distribution of CCR5-Delta32, CCR5 promoter 59029 A/G, CCR2-64I and SDF1-3'A genetic polymorphisms in HIV-1 infected and uninfected patients in the west region of Cameroon. BMC Res Notes. 2013; 6: 288.

Novitsky V, Wester CW, DeGruttola V, Bussmann H, Gaseitsiwe S, Thomas A, Moyo S, Musonda R, Van Widenfelt E, Marlink RG, Essex M: The reverse transcriptase 67N 70R 215Y genotype is the predominant TAM pathway associated with virologic failure among HIV type 1C-infected adults treated with ZDV/ddI-containing HAART in southern Africa. AIDS Res Hum Retroviruses. 2007; 23: 868-878.

Oliveira MF, Ramalho DB, Abreu CM, Vubil A, Mabunda N, Ismael N, Francisco C, Jani IV, Tanuri A. Genetic diversity and naturally polymorphisms in HIV type 1 integrase isolates from Maputo, Mozambique: Implications for integrase inhibitors. AIDS Res. Hum. Retroviruses. 2012; 12: 368.

Ono A, Freed E. Plasma membrane rafts play a critical role in HIV-1 assembly and release. Proc. Natl. Acad. Sci. U.S.A. 2001; 98: 13925–13930.

Oroszlan s and Luftig RB. Retroviral Proteinases. Current Topics in Microbiology and Immunology. 1990; 157: 153-185.

Orrell C, Walensky RP, Losina E, Pitt J, Freedberg KA, Wood R. HIV-1 clade C resistance genotypes in treatment-naive patients and after first virological failure in a large community antiretroviral therapy programme. Antivir Therap. 2009; 14: 523-31.

Paiardini MPI, Apetrei C, Silvestri G. Lessons learned from the natural hosts of HIV-related viruses. Annu Rev Med. 2009; 60: 485–495.

Pandrea I, Sodora DL, Silvestri G, Apetrei C. Into the wild: simian immunodeficiency virus (SIV) infection in natural hosts. Trends Immunol. 2008; 29: 419–428.

Papathanasopoulos MA1, Vardas E, Wallis C, Glashoff R, Buttó S, Poli G, Malnati M, Clerici M, Ensoli B. Characterization of HIV type 1 genetic diversity among South African participants enrolled in the AIDS Vaccine Integrated Project (AVIP) study. AIDS Res Hum Retroviruses. 2010; 26: 705-9.

Paredes R, Lalama CM, Ribaudo HJ. Pre-existing minority drug-resistant HIV-1 variants, adherence, and risk of antiretroviral treatment failure. The Journal of Infectious Diseases. 2010; 201: 662-671.

Pasquier C, Millot N, Njouom R, Sandres K, Cazabat M, Puel J, et al. HIV-1 subtyping using phylogenetic analysis of pol gene sequences. J Virol Methods. 2001; 94: 45–54.

Peeters M, Toure-Kane C, Nkengasong JN. Genetic diversity of HIV in Africa: impact on diagnosis, treatment, vaccine development and trails. AIDS. 2003; 17: 2547–3256.

Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. Science. 1996; 271: 1582–1586.

Pierson T, Hoffman TL, Blankson J, Finzi D, Chadwick K, Margolick JB, Buck C, Siliciano JD, Doms RW, Siliciano RF.Characterization of chemokine receptor utilization of viruses in the latent reservoir for human immunodeficiency virus type 1. J Virol. 2000; 74: 7824-33.

Pillay V, Ledwaba J, Hunt G, Rakgotho M, Singh B, Makubalo L, Bennett DE, Puren A, Morris L. Antiretroviral drug resistance surveillance among drug-naive HIV-1-infected individuals in Gauteng Province, South Africa in 2002 and 2004. Antivir Ther. 2008a; 13:101–107.

Pillay V, Pillay C, Kantor R, Venter F, Levin L, Morris L. HIV type 1 subtype C drug resistance among pediatric and adult South African patients failing antiretroviral therapy.AIDS Res Hum Retroviruses. 2008a; 24:1449-54.

Plantier JC, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, Lemee V, Damond F, Robertson DL, Simon F. A new human immunodeficiency virus derived from gorillas. Nature Med.2009; 15: 871–872.

Plantiera JC, Dachraouia R, Leme´ea V, Gueudina M, Franc¸oise B, Franc¸ois C and Franc¸ois S. HIV-1 resistance genotyping on dried serum spots. AIDS. 2005; 19: 391–397.

Pomerantz RJ, Horn DL. Twenty years of therapy for HIV-1 infection. Nat Med. 2003; 9: 867-73.

Quiros-Roldan E, Signorini S, Castelli F, et al. Analysis of HIV-1 mutation patterns in patients failing antiretroviral therapy. J Clin Lab Anal. 2001; 15: 43–46.

Rambaut A, Posada D, Crandall KA, Holmes EC. The causes and consequences of HIV evolution. Nature Genetics Review. 2004; 5: 52-61.

Ratner L, Gallo RC & Wong-Staal F. HTLV-III, LAV, ARV are variants of same AIDS virus. Nature. 1985b; 313: 636-637.

Ratner L, Haseltine W, Patarca R, Livak KJ, Starcich B, Josephs SF, Doran ER, Rafalski JA, Whitehorn EA, Baumeister K, et al. Complete nucleotide sequence of the AIDS virus, HTLV-III. Nature. 1985a; 313: 277-84.

Regoes RR and Bonhoeffer S.The HIV coreceptor switch: a population dynamical perspective. Trends Microbiol. 2005; 13 : 269-77.

Rey-Cuille MA, Berthier JL, Bomsel-Demontoy MC, Chaduc Y, Montagnier L, Hovanessian LA and Chakrabarati LA. Simian immunodeficiency virus replicates to high levels in sooty mangabeys without inducing disease. J Virol. 1998; 72: 3872-3886.

Reymond A, Meroni G, Fantozzi A, et al. The tripartite motif family identifies cell compartments. EMBO J. 2001; 20: 2140-51.

Roberts JD. et al. The accuracy of reverse transcriptase from HIV-1. Science. 1988; 242: 1171–1173.

Roshal M, Zhu Y & Planelles V. Apoptosis in AIDS. 2001; 6: 103-116.

Ross LL, Weinberg WG, Dejesus E et al. Impact of low abundance HIV variants on response to ritonavir-boosted atazanavir or fosamprenavir given once daily with tenofovir/emtricitabine in antiretroviral-naive HIV-infected patients. AIDS Res. Hum. Retrov. 2010; 26: 407–417

Rousseau M, Vergne L, Montes B, et al. Patterns of resistance mutations to antiretroviral drugs in extensively treated HIV-1-infected patients with failure of highly active antiretroviral therapy. J Acquir Immune Defic Syndr. 2001; 26: 36–43.

Salemi M, Strimmer K, Hall WW et al. Dating the common ancestor of SIVcpz and HIV-1 group M and the origin of HIV-1 subtypes using a new method to uncover clock-like molecular evolution. FASEB J. 2001; 15: 276–278.

Sambrok J, Fritsch EF and Maniatis T. Molecular Cloning: A laboratory Manual second edition. Cold Spring Harbor Laboratory Press. New York, USA. 1989.

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

Schwartz O, Maréchal V, Le Gall S, Lemonnier F, Heard JM. Endocytosis of major histocompatibility complex class I molecules is induced by the HIV-1 Nef protein. Nat Med. 1996 2:338-42.

Severe P, Leger P, Charles M, Noel F, Bonhomme G, et al. Antiretroviral therapy in a thousand patients with AIDS in Haiti. N Engl J Med. 2005; 353: 2325- 2334.

Sharp PM and Hann BH. 'Origins of HIV and the AIDS pandemic' Cold spring Harbour Perspectives in Medicine. 2011; 1:a006841.

Sharp PM and Hann BH. The evolution of HIV-1 and the origin of AIDS. Philos Trans R Soc Lon B Biol Sci. 2010; 365: 2487- 94.

Sheehy AM, Gaddis NC, Choi JD, Malim MH. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature. 2002; 418: 646–650.

Sher R. HIV infection in South Africa, 1982–1988—a review. South African. Medical Journal. 1989; 76: 314–318.

Siliciano JD, Siliciano RF. The latent reservoir for HIV-1 in resting CD4+ T cells: a barrier to cure. Curr Opin HIV AIDS. 2006; 1: 121–128.

Silvestri G, Paiardini M, Pandrea I, Lederman MM, Sodora DL. Understanding the benign nature of SIV infection in natural hosts. J. Clin. Invest. 2007; 117: 3148-3154.

Smith MW, Dean M, Carrington M, Winkler C, Huttley GA, LombDA, et al. Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression. Science 1997;277: 959–65.

Spira S, Wainberg MA, Loemba H, Turner D, Brenner BG. Impact of clade diversity of HIV-1 virulence, antiretroviral drug sensitivity and drug resistance. Journal of Antimicrobial Chemotherapy. 2003; 51: 229–240.

Stack JC, Welch JD, Ferrari MJ, Shapiro BU, Grenfell BT. Protocols for sampling viral sequences to study epidemic dynamics. J. R. Soc. Interface. 2010; 7: 1119–1127.

Stebbing J, Gazzard B, Patterson S, Bower M, Nelson M, Epenetos A, Ogg G, Gotch F, Savage P. Simplified one-step antibody-HLA directed expansion of HIV-specific cytotoxic T lymphocytes: a system suited for use in vivo. AIDS. 2004; 18: 2099-101.

Strebel K, Daugherty D, Clouse K, Cohen D, Folks T, Martin MA The HIV "A" (sor) gene product is essential for virus infectivity Nature.1987; 328 : 728–730.

Sunpath H, Wu B, Gordon M, Hampton J, Johnson B, et al. High rate of K65R for antiretroviral therapy-naive patients with subtype C HIV infection failing a tenofovir-containing first-line regimen. AIDS. 2012; 26: 1679–1684.

Sunpath H, Kamihara J, Chelin N, et al. Prospectiveanalysis of HIV-1 drug resistance after virologic failure on antiretroviral therapy (ART): initial results from a paediatric cohort study from Kwazulu-Natal, SouthAfrica. In: Third South African AIDS Conference, Durban, South Africa, 2008.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011; 28: 2731–2739.

Taylor B,sobieszczyk.M.McCutchan F. Hammers S.The challenge of HIV-1 Subtype diversity.N Engl J Med. 2008; 358: 1590-1602.

Thompson JD, Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, , Gibson TJ, Higgins DG. Bioinformatics. 2007; 23: 2947-8.

Turner BG and Summers MF. Structural biology of HIV. J Mol Biol. 1999; 285: 1-32.

UNAIDS (2016) 'Fact Sheet 2016'. http://www.unaids.org/en/resources/fact-sheet. 2016/11/24.

UNAIDS (2016) 'Prevention Gap Report'. http://www.avert.org/global-hiv-and-aids-statistics. 2016/11/18.

Vallari A, Holzmayer V, Harris B, Yamaguchi J, Ngansop C, Makamche F, Mbanya D, Kaptue L, Ndembi N, Gurtler L, et al. Confirmation of putative HIV-1 group P in Cameroon. J Virol. 2011; 85: 1403–1407.

Van Harmelen J, Wood R, Lambrick M, Rybicki EP, Williamson AL, et al. An association between HIV-1 subtypes and mode of transmission in Cape Town, South Africa. AIDS .1997; 11: 81–87.

Van Zyl GU, Liu TF, Claassen M, Engelbrecht S, de Oliveira T, et al. 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. 2013.

Van Zyl GU, Cotton MF, Claassen M, Abrahams C, Preiser W.Surveillance of transmitted resistance to antiretroviral drug classesamong young children in the Western Cape Province of SouthAfrica. Pediatr Infect Dis J. 2010; 29:370–371.

Vogelstein B and Gillespie D. Preparative and analytical purification of DNA from agarose. Proc. Natl. Acad. Sci. USA. 1997; 76: 615-619.

Vondrasek J, van Buskirk CP and Wlodawer A. Database of three-dimensional structures of HIV proteinases Nat. Struct. Biol. 1997; 4: 8

Vrancken B, Lequime S, Theys K, Lemey P. Covering all bases in HIV research: unveiling a hidden world of viral evolution. AIDS Rev. 2010; 12: 89–102.

Weller IV, Williams IG. ABC of AIDS. Antiretroviral drugs. BMJ. 2001; 322: 1410-2.

Wertheim JO and Worobey M. Dating the Age of the SIV Lineages That Gave Rise to HIV-1 and HIV-2. PLoS Computational Biology. 2009; 5: 1-9.

Wilkinson E, Engelbrecht S, de Oliveira T. History and origin of the HIV-1 subtype C epidemic in South Africa and the greater southern African region. Sci. Rep. 2005; 5: 16897.

Wilkinson E, Engelbrecht S. Molecular characterization of non-subtype C and recombinant HIV-1 virusesfrom Cape Town, South Africa. Infection, Genetics and Evolution. 2009; 9: 840–846.

Wilkinson E, Holzmayer V, Jacobs GB, de Oliveira T, Brennan CA, Hackett J Jr, van Rensburg EJ, Engelbrecht S. Sequencing and phylogenetic analysis of near full-length HIV-1 subtypes A, B, G and unique recombinant AC and AD viral strains identified in South Africa. AIDS Res Hum Retroviruses. 2015; 31: 412-20.

Williamson C, Englebrecht S, Lambrick M et al. HIV-1 subtypes in different risk groups in South Africa. Lancet. 1995; 346:782.

Winkler C, Modi W, Smith, MW, Nelson GW, Wu X, Carrington M, Dean M, Honjo T, Tashiro K, Yabe D, Buchbinder S, Vittinghoff E, Goedert JJ, O'Brien TR, Jacobson LP,

Detels R, Donfield S, Willoughby A, Gomperts E, Vlahov D, Phair J, O'Brien SJ, Genetic restriction of AIDS pathogenesis by an SDF-1 chemokine gene variant. ALIVE Study, Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS). 1998.

Winters M, Baxter J, Mayers D, et al. Frequency of antiretroviral drug resistance mutations in HIV-1 strains from patients failing triple drug regimens. The Terry Beirn Community Programs for Clinical Research on AIDS. Antivir Ther. 2000; 1: 57–63.

Witvrouw M, Pannecouque C, Switzer WM, Folks TM, de Clercq E, Heneine W. Susceptibility of HIV-2, SIV and SHIV to various anti-HIV-1 compounds: implications for treatment and postexposure prophylaxis. Antivir. Ther. 2004; 9: 57-65.

World Health Organisation (WHO) (2015, 30 September) 'Treat all people living with HIV, offer antiretroviral as additional prevention choice for people at "substantial" risk.'

World Healh Organisation (WHO): Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection Recommendations for a public health approach - Second edition 9 June 2016.

Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, Bunce M, et al. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. Nature. 2008; 455: 661-664.

Xiong Y, Eickbush TH. Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J. 1990; 9: 3353–3362.

Yeni P, Hammer S, Carpenter C, et al. Antiretroviral treatment for adult HIV infection in 2002: updated recommendations of the International AIDS Society-USA Panel. JAMA. 2002; 288: 222–235.

Young FE. The role of the FDA in the effort against AIDS. Public Health Rep. 1988; 103: 242–245.

Zanchetta N, Nardi G, Tocalli L, Drago L, Bossi C, Pulvirenti FR, Galli C and Gismondo MR. Evaluation of the abbott LCx HIV-1 RNA quantitative, a new assay for quantitative determination of human immunodeficiency virus type 1 RNA. J. Clin. Microbiol. 2000; 38: 3882-3886.

Zapata W, Weber J, Estrada H, et al. Increased levels of human beta-defensins mrna in sexually HIV-1 exposed but uninfected individuals. Curr HIV Res. 2008; 6: 531-8.

Zhuang J, Jetzt AE, Sun G, Yu H, Klarmann G, Ron Y, Preston BD, Dougherty JP. Human immunodeficiency virus type 1 recombination: rate, fidelity, and putative hot spots. J. Virol. 2002; 11: 11273-11282.

APPENDIX

APPENDIX A: Ethical approval

UNIVERSITEIT•STELLENBOSCH•UNIVERSITY
jou kennisvennoot • your knowledge partner

Approval Notice
Response to Modifications- (New Application)

13-Oct-2015
Jacobs, Graeme GB

Ethics Reference #: N15/08/071
Title: Tracking the molecular epidemiology and resistance patterns of HIV-1 in South Africa.

Dear Dr. Graeme Jacobs,

The **Response to Modifications** - *(New Application)* received on **28-Sep-2015**, was reviewed by members of **Health Research Ethics Committee 1** via Expedited review procedures on **13-Oct-2015** and was approved.
Please note the following information about your approved research protocol:

Protocol Approval Period: **13-Oct-2015 -12-Oct-2016**

Please remember to use your **protocol number** (**N15/08/071**) on any documents or correspondence with the HREC concerning your research protocol.

Please note that the HREC has the prerogative and authority to ask further questions, seek additional information, require further modifications, or monitor the conduct of your research and the consent process.

After Ethical Review:
Please note a template of the progress report is obtainable on www.sun.ac.za/rds and should be submitted to the Committee before the year has expired. The Committee will then consider the continuation of the project for a further year (if necessary). Annually a number of projects may be selected randomly for an external audit.
Translation of the consent document to the language applicable to the study participants should be submitted.

Federal Wide Assurance Number: 00001372
Institutional Review Board (IRB) Number: IRB0005239

The Health Research Ethics Committee complies with the SA National Health Act No.61 2003 as it pertains to health research and the United States Code of Federal Regulations Title 45 Part 46. This committee abides by the ethical norms and principles for research, established by the Declaration of Helsinki, the South African Medical Research Council Guidelines as well as the Guidelines for Ethical Research: Principles Structures and Processes 2004 (Department of Health).

Provincial and City of Cape Town Approval

Please note that for research at a primary or secondary healthcare facility permission must still be obtained from the relevant authorities (Western Cape Department of Health and/or City Health) to conduct the research as stated in the protocol. Contact persons are Ms Claudette Abrahams at Western Cape Department of Health (healthres@pgwc.gov.za Tel: +27 21 483 9907) and Dr Helene Visser at City Health (Helene.Visser@capetown.gov.za Tel: +27 21 400 3981). Research that will be conducted at any tertiary academic institution requires approval from the relevant hospital manager. Ethics approval is required BEFORE approval can be obtained from these health authorities.

We wish you the best as you conduct your research.
For standard HREC forms and documents please visit: www.sun.ac.za/rds

If you have any questions or need further assistance, please contact the HREC office at 0219399657.

UNIVERSITEIT·STELLENBOSCH·UNIVERSITY
jou kennisvennoot • your knowledge partner

Ethics Letter

06-Oct-2016
Jacobs, Graeme GB

Ethics Reference #: N15/08/071
Clinical Trial Reference #:
Title: Tracking the molecular epidemiology and resistance patterns of HIV-1 in South Africa.

Dear Dr. Graeme Jacobs

Your request for extension/annual renewal of ethics approval dated 23 September 2016 refers.

The Health Research Ethics Committee reviewed and approved the annual progress report you submitted through an expedited review process.

The approval of the research project is extended for a further year.

Approval date: 06 October 2016

Expiry date: 05 October 2017

Where to submit any documentation
Kindly submit **ONE HARD COPY** to Elvira Rohland, RDSD, Room 5007, Teaching Building, and **ONE ELECTRONIC COPY** to ethics@sun.ac.za

Please remember to use your **protocol number (N15/08/071)** on any documents or correspondence with the HREC concerning your research protocol.

Federal Wide Assurance Number: 00001372
Institutional Review Board (IRB) Number: IRB0005240 for HREC1
Institutional Review Board (IRB) Number: IRB0005239 for HREC2

The Health Research Ethics Committee complies with the SA National Health Act No.61 2003 as it pertains to health research and the United States Code of Federal Regulations Title 45 Part 46. This committee abides by the ethical norms and principles for research, established by the Declaration of Helsinki, the South African Good Clinical Practices Guidelines as well as the Guidelines for Ethical Research: Principles Structures and Processes 2004 (Department of Health).

Sincerely,
Francis Masiye
REC Coordinator
Health Research Ethics Committee 2

APPENDIX B: HIV-1 viral load cohort regimens

SAMPLE CODE	REGIMEN	SAMPLE CODE	REGIMEN
11_ZA_WC_CW_35	TDF+3TC+EFV	14_ZA_WC_OV_13	D4T+3TC+KLT
11_ZA_WC_CW_38	TDF+3TC+EFV	14_ZA_WC_OV_8	TDF+EFV+FTC
11_ZA_WC_CW_39	TDF+EFV+FTC	14_ZA_WC_WC_12	TDF+3TC+EFV
11_ZA_WC_CW_4	TDF+3TC+LPV/r	14_ZA_WC_CW_46	NEW
11_ZA_WC_CW_40	TDF+3TC+EFV	14_ZA_WC_CW_46	NEW
11_ZA_WC_CW_41	LPV +Truvada	14_ZA_WC_CW_48	TDF+3TC+EFV
11_ZA_WC_CW_42	TDF+3TC+LPV/r	14_ZA_WC_CW_49	TDF+3TC+LPV/r
11_ZA_WC_CW_7	TDF+3TC+EFV	14_ZA_WC_CW_51	FDC
11_ZA_WC_WC_23	TDF+EFV+FTC	14_ZA_WC_CW_56	NEW
11_ZA_WC_WC_5	FDC	14_ZA_WC_OV_10	TDF+3TC+EFV
11_ZA_WC_WC_6	TDF+3TC+EFV	14_ZA_WC_OV_17	TDF+3TC+EFV
11_ZA_WC_WC_8	TDF+3TC+EFV	14_ZA_WC_OV_20	TDF+3TC+KLT
12_ZA_WC_CW_24	TDF+EFV+FTC	14_ZA_WC_OV_6	FDC
12_ZA_WC_CW_44	TDF+3TC+EFV	14_ZA_WC_OV_7	TDF+EFV+FTC
12_ZA_WC_CW_46	TDF+3TC+EFV	14_ZA_WC_OV_8	TDF+3TC+EFV
12_ZA_WC_CW_47	D4T+3TC+NVP	14_ZA_WC_OV_9	TDF+3TC+EFV
12_ZA_WC_CW_48	NEW	14_ZA_WC_WC_11	TDF+EFV+FTC
12_ZA_WC_CW_49	AZT +3TC+EFV+RIF	14_ZA_WC_WC_13	ABC +TDF +3TC
12_ZA_WC_CW_5	TDF+3TC+EFV	14_ZA_WC_WC_16	TDF+EFV+FTC
12_ZA_WC_CW_50	AZT +3TC+KLT +RTF	14_ZA_WC_WC_20	ABC+3TC+EFV
12_ZA_WC_CW_51	AZT+3TC+NVP	15_ZA_WC_CW_15	D4T+3TC+EFV
12_ZA_WC_CW_53	TDF+3TC+EFV	15_ZA_WC_CW_16	TDF+EFV+FTC
12_ZA_WC_CW_54	TDF+3TC+EFV	15_ZA_WC_CW_17	NEW
12_ZA_WC_CW_55	ABC +TDF +EFV	15_ZA_WC_CW_18	AZT +3TC+EFV
12_ZA_WC_CW_56	TDF+EFV+FTC	15_ZA_WC_CW_19	TDF+3TC+EFV
12_ZA_WC_CW_57	TDF+3TC+EFV	15_ZA_WC_CW_21	EFV+LAMIVUDINE
13_ZA_WC_CW_66	TDF+3TC+EFV	15_ZA_WC_CW_22	TDF+EFV+FTZ
13_ZA_WC_CW_67	TDF+3TC+KLT+INH	15_ZA_WC_CW_23	TDF+3TC+EFV
13_ZA_WC_CW_68	TDF+3TC+EFV	15_ZA_WC_CW_32	AZT +3TC+NVP
13_ZA_WC_CW_69	D4T+3TC+KLT	15_ZA_WC_CW_33	TDF+3TC+EFV
13_ZA_WC_OV_10	AZT+3TC+LPV/r	15_ZA_WC_CW_36	TDF+3TC+EFV

APPENDIX C: HIV-1 viral load results

Appendix A.RNA Viral loads results using the Abbott LCx® HIV Quantitative assay

Sample Code	Age	Gender	Viral loads (RNA Copies per/ml)	Sample Date
11_ZA_WC_CW_3	39	M	4888	2011
11_ZA_WC_CW_34	36	F	82883	2011
11_ZA_WC_CW_35	26	F	36148	2011
11_ZA_WC_CW_38	37	M	7874	2011
11_ZA_WC_CW_39	52	F	255538	2011
11_ZA_WC_CW_4	15	F	303100	2011
11_ZA_WC_CW_40	49	F	43580	2011
11_ZA_WC_CW_41	41	M	13587	2011
11_ZA_WC_CW_42	27	F	25720	2011
11_ZA_WC_CW_7	50	M	815609	2011
11_ZA_WC_CW_8	27	F	35767	2011
11_ZA_WC_CW_9	28	F	11980	2011
11_ZA_WC_OV_1	31	F	6235	2011
11_ZA_WC_OV_15	11	F	78887	2011
11_ZA_WC_OV_17	57	F	254353	2011
11_ZA_WC_OV_18	13	M	8954	2011
11_ZA_WC_OV_3	25	M	437261	2011
11_ZA_WC_WC_1	29	F	6795	2011
11_ZA_WC_WC_14	50	F	6715	2011
11_ZA_WC_WC_15	50	F	97740	2011
11_ZA_WC_WC_16	26	M	198439	2011
11_ZA_WC_WC_17	35	F	8205	2011
11_ZA_WC_WC_2	31	F	21854	2011
11_ZA_WC_WC_20	35	F	470985	2011
11_ZA_WC_WC_3	47	F	436327	2011
11_ZA_WC_WC_4	20	F	8095	2011
11_ZA_WC_WC_5	27	M	548203	2011
11_ZA_WC_CW_1	24	F	7736	2011
11_ZA_WC_CW_17	26	F	7191	2011
11_ZA_WC_CW_25	27	F	660756	2011
11_ZA_WC_CW_27	39	M	5758	2011
11_ZA_WC_CW_3	34	M	286717	2011
11_ZA_WC_CW_4	31	F	3532	2011
11_ZA_WC_CW_6	35	M	588250	2011
11_ZA_WC_CW_64	36	M	51839	2011
11_ZA_WC_CW_7	26	F	3421	2011
11_ZA_WC_OV_19	57	F	4556	2011
11_ZA_WC_WC_8	25	M	148810	2011
11_ZA_WC_WC_2	26	M	55634	2011
11_ZA_WC_WC_33	35	F	4317	2011

Sample Code	Age	Gender	Viral loads (RNA Copies per/ml)	Sample Date
11_ZA_WC_OV_2	13	M	44065	2011
11_ZA_WC_OV_23	11	M	368492	2011
11_ZA_WC_WC_1	39	F	2173	2011
11_ZA_WC_WC_17	50	M	164867	2011
11_ZA_WC_WC_23	50	M	3994	2011
11_ZA_WC_WC_5	25	M	125810	2011
11_ZA_WC_WC_6	26	M	9334	2011
11_ZA_WC_WC_8	35	F	439818	2011
12_ZA_WC_CW_24	42	F	619718	2012
12_ZA_WC_CW_37	26	M	293993	2012
12_ZA_WC_CW_43	21	F	6766	2012
12_ZA_WC_CW_44	38	F	3137	2012
12_ZA_WC_CW_46	30	F	7081	2012
12_ZA_WC_CW_47	30	M	2909	2012
12_ZA_WC_CW_48	30	F	25698	2012
12_ZA_WC_CW_49	35	M	13415	2012
12_ZA_WC_CW_5	16	M	360840	2012
12_ZA_WC_CW_50	38	F	2528632	2012
12_ZA_WC_CW_51	32	F	983658	2012
12_ZA_WC_CW_53	32	M	2096	2012
12_ZA_WC_CW_54	43	F	17612	2012
12_ZA_WC_CW_55	13	F	56130	2012
12_ZA_WC_CW_56	52	F	69283	2012
12_ZA_WC_CW_57	32	F	116542	2012
12_ZA_WC_CW_58	36	F	4278	2012
12_ZA_WC_CW_59	46	F	9515	2012
12_ZA_WC_CW_6	54	F	44992	2012
12_ZA_WC_CW_60	26	F	23430	2012
12_ZA_WC_OV_16	38	M	37297	2012
12_ZA_WC_OV_19	36	F	19059	2012
12_ZA_WC_OV_20	33	M	3546	2012
12_ZA_WC_WC_10	45	F	20499	2012
12_ZA_WC_WC_18	31	F	11124	2012
12_ZA_WC_WC_21	29	F	1240706	2012
12_ZA_WC_WC_22	43	M	2416	2012
12_ZA_WC_CW_10	32	F	146528	2012
12_ZA_WC_CW_12	31	F	2252	2012

Sample Code	Age	Gender	Viral loads (RNA Copies per/ml)	Sample Date
12_ZA_WC_CW_13	24	M	2061393	2012
12_ZA_WC_CW_15	21	F	45884	2012
12_ZA_WC_CW_16	42	M	1309767	2012
12_ZA_WC_CW_18	36	M	366761	2012
12_ZA_WC_CW_20	29	F	320249	2012
12_ZA_WC_CW_22	26	F	29191	2012
12_ZA_WC_CW_23	32	M	2562	2012
12_ZA_WC_CW_24	23	F	89525	2012
12_ZA_WC_CW_37	36	F	22525	2012
12_ZA_WC_CW_38	46	F	6214	2012
12_ZA_WC_CW_39	33	M	36233	2012
12_ZA_WC_CW_40	40	F	590722	2012
12_ZA_WC_CW_41	26	M	882499	2012
12_ZA_WC_CW_42	38	F	221054	2012
12_ZA_WC_CW_44	46	M	603027	2012
12_ZA_WC_CW_9	34	F	14564	2012
12_ZA_WC_OV_13	33	F	12706	2012
12_ZA_WC_OV_3	36	M	3873	2012
12_ZA_WC_OV_4	28	M	2172905	2012
12_ZA_WC_OV_5	34	M	52992	2012
12_ZA_WC_WC_10	30	F	537448	2012
12_ZA_WC_WC_12	31	M	122971	2012
12_ZA_WC_WC_14	29	F	3272216	2012
12_ZA_WC_WC_15	30	F	14499	2012
12_ZA_WC_WC_2	40	M	225889	2012
12_ZA_WC_WC_9	9	F	47360	2012
13_ZA_WC_CW_1	35	F	2329	2013
13_ZA_WC_CW_2	42	F	7232	2013
13_ZA_WC_CW_28	30	M	2292	2013
13_ZA_WC_CW_31	33	M	9136	2013
13_ZA_WC_CW_45	25	F	11675	2013
13_ZA_WC_CW_52	49	M	211680	2013
13_ZA_WC_CW_61	43	M	228696	2013
13_ZA_WC_CW_62	29	M	12618	2013
13_ZA_WC_CW_63	52	F	14623	2013
13_ZA_WC_CW_64	39	M	775585	2013
13_ZA_WC_CW_65	51	M	775585	2013

Sample Code	Age	Gender	Viral loads (RNA Copies per/ml)	Sample Date
13_ZA_WC_CW_66	14	M	205860	2013
13_ZA_WC_CW_67	40	F	104991	2013
13_ZA_WC_CW_68	28	M	48548	2013
13_ZA_WC_CW_69	42	M	28843	2013
13_ZA_WC_OV_10	59	F	26389	2013
13_ZA_WC_OV_14	42	F	4660	2013
13_ZA_WC_OV_2	46	F	152537	2013
13_ZA_WC_OV_22	30	F	246511	2013
13_ZA_WC_OV_9	14	M	3862	2013
13_ZA_WC_WC_19	35	F	2350	2013
13_ZA_WC_WC_23	28	F	136932	2013
13_ZA_WC_CW_2	52	F	14945	2013
13_ZA_WC_CW_26	21	F	20162	2013
13_ZA_WC_CW_29	42	F	25684	2013
13_ZA_WC_CW_30	28	M	170852	2013
13_ZA_WC_CW_31	32	F	27760	2013
13_ZA_WC_CW_32	33	M	288729	2013
13_ZA_WC_CW_36	42	F	3397	2013
13_ZA_WC_CW_47	49	F	18073	2013
13_ZA_WC_CW_5	49	F	20499	2013
13_ZA_WC_CW_50	35	M	61674	2013
13_ZA_WC_CW_52	21	M	121276	2013
13_ZA_WC_CW_8	39	F	21637	2013
13_ZA_WC_OV_1	45	F	12897	2013
13_ZA_WC_OV_11	14	F	6992	2013
13_ZA_WC_OV_12	30	F	61739	2013
13_ZA_WC_OV_14	59	F	7892	2013
13_ZA_WC_OV_16	30	M	158229	2013
13_ZA_WC_OV_18	34	M	84975	2013
13_ZA_WC_OV_21	30	F	70893	2013
13_ZA_WC_WC_18	36	M	4131	2013
13_ZA_WC_WC_22	35	M	20238	2013
13_ZA_WC_WC_3	20	M	3394	2013
13_ZA_WC_WC_7	47	M	75033	2013
14_ZA_WC_CW_20	20	M	4944	2014
14_ZA_WC_CW_25	42	M	217243	2014
14_ZA_WC_CW_26	22	M	9428	2014
14_ZA_WC_CW_27	46	M	10961	2014

Sample Code	Age	Gender	Viral loads (RNA Copies per/ml)	Sample Date
14_ZA_WC_CW_29	37	M	2719	2014
14_ZA_WC_CW_30	6	F	35521	2014
14_ZA_WC_CW_70	49	F	19411	2014
14_ZA_WC_CW_71	32	F	45424	2014
14_ZA_WC_OV_11	32	F	3709	2014
14_ZA_WC_OV_12	25	M	5885	2014
14_ZA_WC_OV_13	39	M	344660	2014
14_ZA_WC_OV_21	14	F	14724	2014
14_ZA_WC_OV_7	29	F	3998	2014
14_ZA_WC_OV_8	22	F	201623	2014
14_ZA_WC_WC_11	37	F	53996	2014
14_ZA_WC_WC_12	43	M	538507	2014
14_ZA_WC_WC_13	31	F	164909	2014
14_ZA_WC_CW_28	42	M	19547	2014
14_ZA_WC_CW_33	28	M	58138	2014
14_ZA_WC_CW_35	6	F	705296	2014
14_ZA_WC_CW_43	49	M	103648	2014
14_ZA_WC_CW_45	32	F	8610	2014
14_ZA_WC_CW_46	30	M	95949	2014
14_ZA_WC_CW_48	22	M	5588	2014
14_ZA_WC_CW_49	32	F	301227	2014
14_ZA_WC_CW_51	37	F	4155	2014
14_ZA_WC_CW_56	43	F	12618	2014
14_ZA_WC_OV_10	22	M	70224	2014
14_ZA_WC_OV_17	14	F	15663	2014
14_ZA_WC_OV_20	41	F	92031	2014
14_ZA_WC_OV_6	39	F	2035	2014
14_ZA_WC_OV_7	36	M	166011	2014
14_ZA_WC_OV_8	39	F	5812	2014
14_ZA_WC_OV_9	32	F	31569	2014
14_ZA_WC_WC_11	39	F	12571	2014
14_ZA_WC_WC_13	34	M	45941	2014
14_ZA_WC_WC_16	31	M	45941	2014
14_ZA_WC_WC_20	52	M	11624	2014
15_ZA_WC_CW_10	31	F	14441	2015
15_ZA_WC_CW_11	14	F	319181	2015

Sample Code	Age	Gender	Viral loads (RNA Copies per/ml)	Sample Date
15_ZA_WC_CW_12	32	F	396299	2015
15_ZA_WC_CW_13	40	F	1328754	2015
15_ZA_WC_CW_14	24	M	3249561	2015
15_ZA_WC_CW_15	43	M	147609	2015
15_ZA_WC_CW_16	43	F	608665	2015
15_ZA_WC_CW_17	25	F	16138	2015
15_ZA_WC_CW_18	41	M	20097	2015
15_ZA_WC_CW_19	52	F	9709	2015
15_ZA_WC_CW_21	26	M	7219	2015
15_ZA_WC_CW_22	34	F	3706	2015
15_ZA_WC_CW_23	41	F	15875	2015
15_ZA_WC_CW_32	14	M	492505	2015
15_ZA_WC_CW_33	22	M	150341	2015
15_ZA_WC_CW_36	36	F	46309	2015
15_ZA_WC_OV_4	25	F	16580	2015
15_ZA_WC_OV_5	41	F	206291	2015
15_ZA_WC_OV_6	32	M	541438	2015
15_ZA_WC_WC_6	41	F	7232	2015
15_ZA_WC_WC_7	42	F	3661	2015
15_ZA_WC_WC_8	44	F	138983	2015
15_ZA_WC_WC_9	29	F	13715	2015

APPENDIX D: Nucleic acids concentration determination.

Sample Code	ng/ul	260/280
11_ZA_WC_CW_3	124.2	1.83
11_ZA_WC_CW_34	123.84	1.87
11_ZA_WC_CW_35	87.08	1.82
11_ZA_WC_CW_38	87.03	1.8
11_ZA_WC_CW_39	83.86	1.29
11_ZA_WC_CW_4	84.38	2.31
11_ZA_WC_CW_40	129.01	1.81
11_ZA_WC_CW_41	126.81	1.81
11_ZA_WC_CW_42	100.17	1.82
11_ZA_WC_CW_7	93.88	1.81
11_ZA_WC_CW_8	83.93	1.84
11_ZA_WC_CW_9	85.36	1.85
11_ZA_WC_OV_1	136.27	1.87
11_ZA_WC_OV_15	144.88	1.88
11_ZA_WC_OV_17	116.37	1.88
11_ZA_WC_OV_18	128.1	1.82
11_ZA_WC_OV_3	136.42	1.81
11_ZA_WC_WC_1	137.55	1.91
11_ZA_WC_WC_14	138.98	1.89
11_ZA_WC_WC_15	140.25	1.87
11_ZA_WC_WC_16	106.13	1.99
11_ZA_WC_WC_17	105.63	1.85
11_ZA_WC_WC_2	131.32	1.93
11_ZA_WC_WC_20	84.67	1.99
11_ZA_WC_WC_3	84.29	1.88
11_ZA_WC_WC_4	94.76	1.82
11_ZA_WC_WC_5	92.99	1.85
11_ZA_WC_CW_1	105.1	1.85
11_ZA_WC_CW_17	104.12	1.85
11_ZA_WC_CW_25	103.5	1.81
11_ZA_WC_CW_27	102.21	1.83
11_ZA_WC_CW_3	122.36	1.83
11_ZA_WC_CW_4	122.85	1.85
11_ZA_WC_CW_6	100.19	1.81
11_ZA_WC_CW_64	99.72	1.86
11_ZA_WC_CW_7	116.01	1.84
11_ZA_WC_OV_19	113.06	1.84

Sample Code	ng/ul	260/280
11_ZA_WC_OV_2	94.45	1.89
11_ZA_WC_OV_23	94.59	1.82
11_ZA_WC_WC_1	105.59	1.87
11_ZA_WC_WC_17	103	1.85
11_ZA_WC_WC_23	120.28	1.84
11_ZA_WC_WC_5	120.65	1.88
11_ZA_WC_WC_6	100.48	1.91
11_ZA_WC_WC_8	103.16	1.86
12_ZA_WC_CW_24	121.11	1.83
12_ZA_WC_CW_37	120.9	1.85
12_ZA_WC_CW_43	80.98	2.15
12_ZA_WC_CW_44	80.51	2.19
12_ZA_WC_CW_46	128.79	1.82
12_ZA_WC_CW_47	125.05	1.83
12_ZA_WC_CW_48	121.4	1.83
12_ZA_WC_CW_49	123.42	1.86
12_ZA_WC_CW_5	89.1	1.83
12_ZA_WC_CW_50	79.27	1.83
12_ZA_WC_CW_51	86.18	1.61
12_ZA_WC_CW_53	86.91	1.63
12_ZA_WC_CW_54	109.92	1.86
12_ZA_WC_CW_55	110.41	1.9
12_ZA_WC_CW_56	40.56	1.88
12_ZA_WC_CW_57	85.1	1.92
12_ZA_WC_CW_58	89.96	1.9
12_ZA_WC_CW_59	32.82	1.95
12_ZA_WC_CW_6	106.24	1.82
12_ZA_WC_CW_60	106.63	1.82
12_ZA_WC_OV_16	106.45	1.82
12_ZA_WC_OV_19	101.58	1.8
12_ZA_WC_OV_20	106.41	1.82
12_ZA_WC_WC_10	105.14	1.86
12_ZA_WC_WC_18	112.31	1.84
12_ZA_WC_WC_21	111.11	1.87
12_ZA_WC_WC_22	82.52	1.92
12_ZA_WC_CW_10	81	1.91
12_ZA_WC_CW_12	65.92	1.8

Sample Code	ng/ul	260/280
12_ZA_WC_CW_13	65.64	1.87
12_ZA_WC_CW_15	97.6	2.16
12_ZA_WC_CW_16	96.02	2.12
12_ZA_WC_CW_18	107.87	1.79
12_ZA_WC_CW_20	108.41	1.83
12_ZA_WC_CW_22	117.72	1.84
12_ZA_WC_CW_23	120.54	1.83
12_ZA_WC_CW_24	96.3	1.88
12_ZA_WC_CW_37	119.56	1.86
12_ZA_WC_CW_38	110.77	1.87
12_ZA_WC_CW_39	94.59	1.85
12_ZA_WC_CW_40	94.46	1.87
12_ZA_WC_CW_41	141.09	1.88
12_ZA_WC_CW_42	140.13	1.88
12_ZA_WC_CW_44	121.96	1.82
12_ZA_WC_CW_9	121.3	1.87
12_ZA_WC_OV_13	62.09	1.82
12_ZA_WC_OV_3	63.03	1.81
12_ZA_WC_OV_4	94.48	2.21
12_ZA_WC_OV_5	94.18	2.15
12_ZA_WC_WC_10	115.45	1.86
12_ZA_WC_WC_12	115.53	1.85
12_ZA_WC_WC_14	85.9	1.84
12_ZA_WC_WC_15	85.03	1.83
12_ZA_WC_WC_2	47.67	1.99
12_ZA_WC_WC_9	57.75	1.93
13_ZA_WC_CW_1	90.6	1.81
13_ZA_WC_CW_2	91.35	1.79
13_ZA_WC_CW_28	115.16	1.86
13_ZA_WC_CW_31	117.15	1.83
13_ZA_WC_CW_45	92.89	1.87
13_ZA_WC_CW_52	91.74	1.8
13_ZA_WC_CW_61	110.34	1.82
13_ZA_WC_CW_62	109.15	1.87
13_ZA_WC_CW_63	293.75	1.83
13_ZA_WC_CW_64	85.08	1.83
13_ZA_WC_CW_65	119.5	1.83

Sample Code	ng/ul	260/280
13_ZA_WC_CW_66	78.36	2.15
13_ZA_WC_CW_67	80.09	2.21
13_ZA_WC_CW_68	101.15	2.19
13_ZA_WC_CW_69	101.85	2.24
13_ZA_WC_OV_10	82.13	2.21
13_ZA_WC_OV_14	77.71	2.22
13_ZA_WC_OV_2	77.24	1.82
13_ZA_WC_OV_22	79.06	1.73
13_ZA_WC_OV_9	113.59	1.85
13_ZA_WC_WC_19	112.95	1.87
13_ZA_WC_WC_23	91.9	1.86
13_ZA_WC_CW_2	92.1	1.86
13_ZA_WC_CW_26	127.1	1.85
13_ZA_WC_CW_29	129.36	1.84
13_ZA_WC_CW_30	99.43	1.89
13_ZA_WC_CW_31	100.25	1.84
13_ZA_WC_CW_32	78.85	1.8
13_ZA_WC_CW_36	78.67	1.79
13_ZA_WC_CW_47	107.61	1.87
13_ZA_WC_CW_5	107.85	1.81
13_ZA_WC_CW_50	117.18	1.79
13_ZA_WC_CW_52	115.48	1.87
13_ZA_WC_CW_8	117.79	1.85
13_ZA_WC_OV_1	81.86	2.21
13_ZA_WC_OV_11	82.17	2.1
13_ZA_WC_OV_12	112.13	1.85
13_ZA_WC_OV_14	124.79	1.83
13_ZA_WC_OV_16	90.98	1.82
13_ZA_WC_OV_18	88.57	1.77
13_ZA_WC_OV_21	91.85	1.86
13_ZA_WC_WC_18	92.39	1.81
13_ZA_WC_WC_22	90.93	1.76
13_ZA_WC_WC_3	89.03	1.83
13_ZA_WC_WC_7	76.97	1.77
14_ZA_WC_CW_20	76.24	1.79
14_ZA_WC_CW_25	118.99	1.87
14_ZA_WC_CW_26	109.45	1.85

Sample Code	ng/ul	260/280
14_ZA_WC_CW_29	117.4	1.83
14_ZA_WC_CW_30	113.9	1.83
14_ZA_WC_CW_70	105.41	1.83
14_ZA_WC_CW_71	106.82	1.85
14_ZA_WC_OV_11	101.64	1.8
14_ZA_WC_OV_12	69.98	1.81
14_ZA_WC_OV_13	70.28	1.84
14_ZA_WC_OV_21	129.64	1.85
14_ZA_WC_OV_7	110.11	1.85
14_ZA_WC_OV_8	116.94	1.85
14_ZA_WC_WC_11	124.02	1.83
14_ZA_WC_WC_12	97.81	2.14
14_ZA_WC_WC_13	98.29	2.2
14_ZA_WC_CW_28	86.88	1.86
14_ZA_WC_CW_33	89.52	1.85
14_ZA_WC_CW_35	124.69	1.86
14_ZA_WC_CW_43	123.77	1.86
14_ZA_WC_CW_45	124.74	1.89
14_ZA_WC_CW_46	126.99	1.86
14_ZA_WC_CW_48	121.1	1.86
14_ZA_WC_CW_49	129.8	1.87
14_ZA_WC_CW_51	88.47	1.83
14_ZA_WC_CW_56	89.28	1.83
14_ZA_WC_OV_10	73.12	1.82
14_ZA_WC_OV_17	74.52	1.82
14_ZA_WC_OV_20	97.79	2.18
14_ZA_WC_OV_6	99.67	2.19
14_ZA_WC_OV_7	99.83	1.91
14_ZA_WC_OV_8	101.14	1.91
14_ZA_WC_OV_9	54.94	1.89
14_ZA_WC_WC_11	44.99	1.89
14_ZA_WC_WC_13	121.51	1.83
14_ZA_WC_WC_16	120.66	1.86
14_ZA_WC_WC_20	109.39	1.82
15_ZA_WC_CW_10	110.44	1.86
15_ZA_WC_CW_11	103.25	1.85

Sample Code	ng/ul	260/280
15_ZA_WC_CW_12	72.5	1.89
15_ZA_WC_CW_13	72.57	1.9
15_ZA_WC_CW_14	98.98	1.85
15_ZA_WC_CW_15	98.26	1.87
15_ZA_WC_CW_16	117.78	1.89
15_ZA_WC_CW_17	116.9	1.88
15_ZA_WC_CW_18	84.45	1.83
15_ZA_WC_CW_19	84.15	1.81
15_ZA_WC_CW_21	101.08	1.84
15_ZA_WC_CW_22	165.3	1.47
15_ZA_WC_CW_23	102.86	1.82
15_ZA_WC_CW_32	114.85	1.85
15_ZA_WC_CW_33	111.04	1.85
15_ZA_WC_CW_36	76.91	1.82
15_ZA_WC_OV_4	76.82	1.82
15_ZA_WC_OV_5	101.08	1.84
15_ZA_WC_OV_6	100.04	1.85
15_ZA_WC_WC_6	117.89	1.85
15_ZA_WC_WC_7	117.82	1.86
15_ZA_WC_WC_8	88.38	1.83
15_ZA_WC_WC_9	88.88	1.82