APPLICATION OF CLONING IN THE DETECTION OF HIV-1 DRUG RESISTANT MINORITY POPULATIONS

BY

Luiza Miyanda Hatyoka

(11550021)

Dissertation submitted in fulfillment of the requirements for the award of a Master of Science degree in Microbiology at the University of Venda

SUPERVISOR: Professor Pascal O. Bessong

Department of Microbiology,

School of Mathematical and Natural Sciences,

University of Venda, South Africa.

CO-SUPERVISOR: Dr Tracy M. Masebe

Department of Microbiology,

School of Mathematical and Natural Sciences,

University of Venda, South Africa.

November, 2013

ACKNOWLEDGEMENTS

I give all the thanks and praise to the Lord God almighty for his love, grace and abundant favor during the period of my study.

To my mother Hobbia, thank you for all your support and encouragement you showed to me and for always believing that I can't do anything I put my mind to.

To my supervisor Professor Pascal Obong Bessong, thank you Sir for not giving up on me even when the project didn't seem to be getting anywhere. Thank you very much for your patience and encouragement and the time you took off and helped me with lab work, I truly appreciate it.

To my co-supervisor, Dr Tracy Madimabi Masebe, thank you very much for the support and encouragement rendered unto me.

To my colleagues, Dr Julius Nwobegahay, Ms. Nontokozo Matume, Ms. Tshifhiwa Tambani, Ms. Lufuno Mavhandu, Ms. Khumbudzo Mavhunga, Ms. Mikateko Rikhotso, Ms. Happy Tshivhula, Ms. Elisabeth Etta and Ms. Suzan Sathekge of the Department of Microbiology, thank you so much for working with me the entire time of my study and for all the encouragement you gave me.

To the funders of this project, this was made possible by the Technology Innovation Agency and the University of Venda Research and publications committee.

University of Venda
Creating Future Leaders

DECLARATION

I, Hatyoka Luiza Miyanda, hereby declare that the dissertation for Masters degree in the Microbiology Department at the University of Venda, hereby submitted by me, has not been submitted previously for the degree at this or any other University, that is my own work in design and execution, and all the references contained therein has been duly acknowledged.

Student signature _____ Date 11/02/2014

Supervisor: PASCAL BESSONG Date 11/2/2014

Co-supervisor: Masebe T.M Date 11/02/2014

DEDICATION

I dedicate this work to the lord God almighty. It is by his grace and mercy that has been sufficient for me through out the period of my study that I was able to do this research and so this is to his Glory and Honor. Amen.

To my husband Samuel Masubelele, for the support and encouragement especially after the loss of our daughter when I felt like I was losing direction, you told me to hold on that showed me positivity.

To my late daughter Dorothy Masubelele, I'm sorry that you are not here to see the fruits of this work because it became a part of you as much as it was for me.

To my family, the Hatyoka's and Masubelele's, for the love, support and simply just being there for me when I really needed a shoulder to cry on.

ABSTRACT

Human immunodeficiency virus (HIV) is the causative agent of acquired immunodeficiency syndrome (AIDS). The virus is genetically highly diversified with different genetic strains infecting different populations worldwide. Highly active antiretroviral therapy (HAART) reduces the morbidity and mortality due to AIDS. Unfortunately, the efficiency of these drugs is limited by the development of drug resistance usually caused by mutations in the protease and reverse transcriptase genes which complicates patient management.

Sequencing direct PCR products reveals the properties of major viral populations but has the potential to miss minority viral populations. This clearly means that resistant strains of minority populations may not be detected in direct sequencing of PCR products. Furthermore, if multiple drug resistance mutations are detected as mixtures in a given sample, it is not possible to determine whether the individual drug resistance mutations are present together on a single virus or exist in different viral sub-populations. Therefore, the general objective of this study was to design and evaluate a single genome sequencing protocol for the detection of HIV-1 drug resistance mutations in the protease and reverse transcriptase genes.

In this study the HIV Pol gene (for protease and reverse transcriptase genes respectively) was amplified from 8 samples of HIV-1 infected individuals (7 treatment experienced and 1 treatment naïve). Subsequent cloning was done on all 8 samples using the Topo TA cloning vector, 35 and 33 clones were obtained for protease and reverse transcriptase respectively. Sequencing of purified PCR products and viral clones was done using a BigDye terminator sequencer. A comparison of sequences obtained from direct sequencing of PCR products and cloned PCR sequences was done. Phylogenetic analysis confirmed that all sequences were of HIV-1 subtype C and amino acid determination helped revealed amino acid substitutions in comparison to the HIV-1 subtype B and C global consensus.

Using the Stanford drug resistance interpretation program, sequences obtained by direct sequencing of PCR products revealed mutations V11I (a minor resistant mutation for the protease region) and M184V (a NRTI resistant mutation), K101E and Y181C (NNRTI resistant mutations) from samples MARBB14 and MARBB73. Viral clones of samples MARBB14 and MARBB73 revealed the same mutations as those observed by direct sequencing of PCR products. In addition viral clones of sample MARBB73 also revealed reverse transcriptase mutations T215Y (a NRTI resistant mutation) and Y181V (a NNRTI resistant Mutation).

Polymorphisms such as K20R, M36I, M36V, L63P, H69K, V77I and V82I selected by protease inhibitors were also observed in the protease region from both direct sequences of PCR products and cloned PCR sequences. Using the student t-test, p values of 0.03 and 0.0002 for protease and reverse transcriptase respectively were obtained from comparisons of resistant mutations observed in sequences of PCR products and clone sequences for both protease and reverse transcriptase genes and the difference was considered significant.

The cloning technique was used as an approach to detect minority variants. Important resistant mutations were detected in clone sequences and not in sequences of PCR products. This suggests that mutations of interest could be missed in the direct sequencing of PCR products which is currently the norm in many settings.

TABLE OF CONTENT

Acknowledgement i

Declaration ii

Dedication iii

Abstract iv

Table of content vi

List of figures xi

List of Tables xiii

List of Abbreviations xiv

CHAPTER ONE: INTRODUCTION AND LITERATURE REVIEW

1.0 Introduction 1

Literature review 2

1.1 The Global distribution of HIV and epidemiology 2

1.2 The prevalence of HIV in South Africa 3

1.3 Pathogenesis of HIV 5

1.3.1 Primary infection 5

1.3.2 Chronic asymptomatic HIV-1 infection 6

1.3.3 Late stage of HIV-1 6

1.4 Transmission and prevention of HIV 8

1.4.1 Sexual 8

1.4.2 Blood and blood products 9

1.4.3 Mother-to-child 9

1.5 Laboratory diagnosis of HIV 9

1.6 Morphology and structure of HIV 10

1.7 HIV genome organization 11

1.8 Life cycle of HIV 13

1.8.1 HIV binding and entry 13

1.8.2. Replication and transcription 14

1.8.3. Assembly and release 15

1.9 HIV tropism 16

1.10 HIV classification and genetic variation 18

1.11 The HIV-1 Polymerase gene 21

1.11.1 Protease gene 21

1.11.2 Reverse Transcriptase gene 22

1.11.3 Integrase gene 23

1.12 Treatment of HIV 24

1.13 HIV drug resistance 27

1.13.1 Nucleotide Reverse Transcriptase inhibitors (NRTI) mechanism of and drug resistant mutations 27

1.13.2 Non-Nucleotide Reverse Transcriptase inhibitors (NNRTI) mechanism of action and drug resistant mutations 28

1.13.3 Protease inhibitors (PI) mechanism of action and drug resistant mutations 29

1.13.4 Fusion or Entry inhibitor mechanism of action and drug resistant mutations 29

1.13.5 Integrase inhibitors mechanism of action and drug resistant mutations 30

1.14 Minority variants of drug resistant HIV 32

1.14.1 Methods to detect minority variants 32

1.14.1.1 Point mutation assays 32

1.14.1.2 Sequencing assays 33

1.14.1.3 Phenotypic assays 33

1.14.2 Common limitations of minority variant assays 34

1.15 Polymerase Chain Reaction and Molecular cloning as tools in HIV research 35

1.15.1 Polymerase Chain Reaction (PCR) 35

1.15.1.1 Principles of PCR 35

1.15.1.2 PCR components and considerations 37

1.15.1.3 Gel electrophoresis 38

1.15.1.4 Types of PCR 38

1.15.1.4.1 Reverse transcription PCR 38

1.15.1.4.2 Nested PCR 38

1.15.1.4.3 Multiplex PCR 39

1.15.1.4.4 Quantitative PCR 39

1.15.1.5 Avoiding PCR contamination 40

1.15.2 Molecular Cloning 40

1.16 Rationale of the study 42

1.17 Objectives of the study 43

1.17.1 General objectives 43

1.17.2 Specific objectives 43

CHAPTER TWO: MATERIAL AND METHODS

2.0 Ethical clearance 44

2.1 Study population and Biospecimens 44

2.2 Total RNA isolation 44

2.3 DNA extraction from PBMCs 45

2.4 Complementary DNA (cDNA) synthesis 45

2.5 Amplification of HIV-1 pol gene 45

2.6 Agarose gel electrophoresis 47

2.7 Purification of PCR products 47

2.8 Cloning HIV DNA of the partial pol gene 47

2.8.1 Ligation of insert into vector 47

2.8.2. Transformation of Cells 47

2.9 Plasmid preparation and restriction digestion 48

2.10 Sequencing and editing of amplified products and mini preps 48

2.11 Resistant mutation analysis 49

2.12 Viral subtyping and mean genetic distance 49

2.13 Amino acid determination 49

2.14 Statistical analysis 50

CHAPTER THREE: RESULTS

3.0 Demographic and clinical features of study participants 51

3.1 Amplification of HIV pol region 53

3.2 Transformation of Cells 54

3.3 Restriction profiles of mini-preps 55

3.4 Sequencing and Drug resistance mutation analysis 56

3.4.1 Drug resistance mutation analysis of bulk PCR products 56

3.4.2 Drug resistance mutation analysis of cloned PCR products 57

3.4.3 Other mutations detected in the protease gene by bulk PCR sequencing and cloned viral sequences 59

3.4.4 Other mutations detected in the reverse transcriptase gene by bulk PCR sequencing and cloned viral sequences 60

3.5 Viral subtyping determination 62

3.6 Mean genetic distance analysis 70

3.7 Amino acid alignment 73

3.8 Statistical analysis 80

3.9 Comparisons of minority population detecting methods 81

CHAPTER FOUR: DISCUSSION, LIMITATIONS AND CONCLUSION 82

CHAPTER FIVE: REFERENCES 83

LIST OF FIGURES

Figure 1.1: Global prevalence of HIV-1 3

Figure 1.2: Prevalence of HIV-1 in South Africa 4

Figure 1.3: Disease progression in HIV-1 7

Figure 1.4: The HIV-1 virion 11

Figure 1.5: HIV-1 genome 13

Figure 1.6: HIV-1 life cycle 16

Figure 1.7: Types and groups of HIV-1 19

Figure 1.8: Target of Anti-retrovirals in the HIV-1 Life cycle 31

Figure 1.9: Steps involved in Polymerase chain reaction 36

Figure 1.10: Strategies and steps in cloning 41

Figure 3.1: A 1.5% agarose gel image of amplified products of samples obtained from HIV-1 infected individuals 53

Figure 3.2: A representative LB-kanamycin-X-gal plate of cloned PCR products into Topo TA cloning vector 54

Figure 3.3: A 1.5% agarose gel image of mini-prep restriction enzyme digestion 55

Figure 3.5a: Phylogenetic analysis of protease sequences of PCR samples 63

Figure 3.5b: Phylogenetic analysis of 35 protease cloned PCR sequences 64

Figure 3.5c: Phylogenetic analysis of protease PCR and 35 cloned PCR sequences 65

Figure 3.3d: Phylogenetic analysis of reverse transcriptase sequences of PCR samples 66

Figure 3.5e: Phylogenetic analysis of 33 reverse transcriptase cloned PCR sequences 67

Figure 3.5f: Phylogenetic analysis of reverse transcriptase PCR and 33 cloned PCR Sequences 68

Figure 3.4a: Mean nucleotide genetic distance analysis of 8 protease PCR sequences 69

Figure 3.4b: Mean nucleotide genetic distance analysis of 8 reverse transcriptase PCR sequences 69

Figure 3.6c: Mean nucleotide genetic distance analysis of 35 protease PCR Sequences 72

Figure 3.5d: Mean nucleotide genetic distance analysis of 33 reverse transcriptase cloned PCR sequences 73

Figure 3.5a: Amino acid alignment (residues 1-99) of protease PCR sequences 74

Figure 3.5b: Amino acid alignment (residues 1-99) of protease cloned PCR sequences 75

Figure 3.5c: Amino acid alignment (residues 78-285) of reverse transcriptase PCR sequences 76

Figure 3.5d: Amino acid alignment (residues 78-285) of reverse transcriptase cloned sequence 80

LIST OF TABLES

Table 2.1: Primer sequences for first round and nested PCR for the HIV-1 Pol gene, orientation and HXB2 location 46

Table 2.2: Sequencing primers for the HIV-1 pol gene, orientation and HXB2 location 49

Table 3.1: Demographic data, treatment, virologic and immunologic features of study population 52

Table 3.2a: Protease and reverse transcriptase mutation analysis for PCR 57

Table 3.2b: Protease and reverse transcriptase mutation analysis of cloned PCR products 58

Table 3.2c: Other mutations observed in the protease gene of PCR products 59

Table 3.2d: Other mutations observed in the reverse transcriptase gene PCR of products 61

Table 3.3: Advantages and Disadvantages of PCR, cloning and other methods to detect minority variants 81

LIST OF ABBREVIATIONS

µl	Microliter
µM	Micro-molar
AIDS	Acquired immunodeficiency syndrome
ART	Anti-retroviral treatment
CD4	Cluster of differentiation 4
CDC	Centers for disease control
cDNA	Complementary DNA
CRFs	circulating recombinant forms
DNA	Deoxyribonucleic Acid
$dNTP_S$	Deoxynucleoside triphosphates
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
ENV	Envelope
GAG	Group associated antigens
HAART	Highly active antiretroviral therapy
HIV	Human immunodeficiency Virus
HIV-1	Human Immunodeficiency virus type 1
HIV-2	Human immunodeficiency virus type 2
HR1	Heptad repeat 1
HR2	Heptad repeat 2

IFA	Immuno-fluorensence assay
LFA	Leucocyte function associated antigen
LTR	Long terminal repeats
ml	Milliliter
mM	millli-molar
mm^3	Cubic millimeter
NAMS	NRTI-associated mutations
NARTIs	Nucleoside analog reverse-transcriptase inhibitors
NNRTI	Non-nucleoside reverse transcriptase inhibitors
NRTIs	Nucleoside reverse transcriptase inhibitors
°C	Degree Celsius
PCR	Polymerase Chain Reaction
PI	Protease inhibitor
PMTCT	Prevention of Mother to child transmission
POL	Polymerase
Rpm	Resolutions per minute
RNA	Ribonucleic acid
RT	Reverse transcriptase
SIV_{cpz}	Simian immunodeficiency virus in Chimpanzee

SIV_{smm}	simian immunodeficiency virus in sooty mangabegs
TAE	Tris-acetate-EDTA
Tris-HCl	Tris Hydrochloric acid
UNAIDS	The joint United Nations Programme on HIV/AIDS
UNICEF	United Nations Children's Fund
URFs	Unique recombinant forms
WHO	World Health organization

CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

1.0 INTRODUCTION

Human immunodeficiency virus (HIV) is the causative agent of the acquired immune deficiency syndrome (AIDS). First reports of such diseases were published in May 1981 (CDC, 1981) when they were observed in homosexual patients suffering from previously rare diseases such as *Pneumocystis carinii pneumonia* and *Kaposi sarcoma*. It then became clear that the 'new' disease affected other population groups as well, when cases were reported in injecting drug users. However, it took almost two years before the human immunodeficiency virus (HIV) was defined as the primary cause of the acquired immunodeficiency syndrome (Barré-Sinoussi, 1983; Broder, 1984; Gallo, 1984).

An estimate of 34.2 million of people were reported to be living with HIV globally in 2011, with 2.5 million newly infected people and 1.7 million estimated deaths due to HIV/AIDS illnesses (UNAIDS, 2012). Southern Africa has the highest HIV prevalence with South Africa being one of the countries with a high prevalence. In South Africa, the province of KwaZulu-Natal has the highest HIV prevalence (37.4%) followed by the province of Mpumalanga (36.7%), Free State and North-West with overall prevalence rates greater than 30.0%. Limpopo, Gauteng and the Eastern Cape recorded prevalence rates between 20.0% and 30.0%. Northern Cape and Western Cape are the only provinces that have HIV prevalence rates below 20.0% (Department of Health, South Africa, 2012).

The current treatment for HIV infection in South Africa involves the use of highly active antiretroviral therapy (HAART). This has been highly beneficial to many HIV-infected individuals since its introduction in 2004. HAART options are combinations of three drugs belonging to at least two classes of antiretrovirals, usually nucleotide analogue reverse transcriptase inhibitors plus either a protease inhibitor or non-nucleotide analogue reverse transcriptase inhibitor. However drug classes such as entry or fusion and integrase inhibitors provide treatment options for people infected with viruses already resistant to the common therapies, although they are not widely available and not typically accessible in resource-limited settings (Nwobegahay *et al.*, 2011).

One of the major drawbacks of HAART is the development of drug resistance which usually occurs as a result of the use of anitiretrovirals. For individuals who are drug naïve, antiretroviral drug resistance is acquired through the transmission of resistant viruses hence, posing a high risk of virological failure upon the initiation of antiretroviral treatment and resulting to a lower genetic barrier to resistance. Therefore, treatment should be accompanied by regular virologic monitoring that include measurement of viral load and testing of drug resistance to guide patient management (Nwobegahay *et al.*, 2011).

LITERATURE REVIEW

1.1 THE GLOBAL DISTRIBUTION OF HIV AND EPIDEMIOLOGY

An estimated 34.2 million [31.8million - 35.9 million] people in 2011 were living with HIV globally, with 2.5 million [2.2 million – 2.8 million] newly infected and 1.7 million [1.6million – 1.9 million] estimated deaths due to HIV/AIDS- related illnesses (UNAIDS, 2012).

The HIV/AIDS prevalence continues to increase in Eastern Europe and Central Asia and in other parts of Asia. In spite of this, the epidemic seems to be stabilized in most regions even though Sub-Saharan Africa still remains the most heavily affected region. In 2011, the Sub-Saharan Africa region reported an estimated 23.5 million [22.2 million – 24.7 million] people living with HIV, 1.7 million [1.6 million – 1.9 million] newly infected people and 1.2 million [1.1 million – 1.3 million] deaths due to HIV/AIDS related illnesses (UNAIDS, 2012).

South and South-East Asia is the second heavily affected region for HIV infections globally with an estimate of 4.2 million [3.1 million – 4.7 million] people infected in 2011, with 39 000 [222,000 – 340 000] newly infected people and 25 000 [140 000 – 600 000] deaths estimated due to HIV/AIDS related illnesses. The region that recorded the lowest statistics is Oceania with 53 000 [47 000 – 60 000] people living with HIV, 2 900 [2 200 – 1 800] newly infected people and 1 300 [<1000 – 1 800] (UNAIDS, 2012). Figure 1.1 shows the global distribution of HIV infections as of end of 2011.

Figure 1.1: Number of adults and children estimated to be living with HIV/AIDS in 2011. (www.slideshare.net/UNAIDS/2012-unaids-epidemiology-slides) (Accessed on 15 May 2013)

1.2 THE PREVALENCE OF HIV IN SOUTH AFRICA

South Africa is seen to have the highest number of people living with HIV in the world with an estimate of 5.6 million people living with HIV. However, recent data shows promising signs, with national HIV prevalence and incidence stabilizing or even decline. (UNAIDS 2010).

The HIV prevalence estimates vary between provinces with provinces such as Gauteng showing a slight decrease from 30.4% in 2010 to 28.7% in 2011 and KwaZulu-Natal from 39.5 % to 37.4%. (Department of Health, South Africa 2012). A slight decrease was also seen in the Northern Cape (18.4% in 2010 to 17.0% in 2011) and the Western Cape (18.5% in 2010 t0 18.2% in 2011) and the Eastern Cape (29.9% in 2010 to 29.3% in 2011). However, an increase in HIV prevalence was observed in Free State (30.6% in 2010 to 32.5% in 2011), and the North-West (29.6% in 2010 to 30.2% in 2011). A steady increase is seen in Limpopo (21.4% in 2009 to 22.1% in 2011) and Mpumalanga (35.1% in 2010 to 36.7% in 2011) (Department of Health,

South Africa 2012). Figure 1.2 shows the South African provincial distribution of HIV infections as of end of 2011.

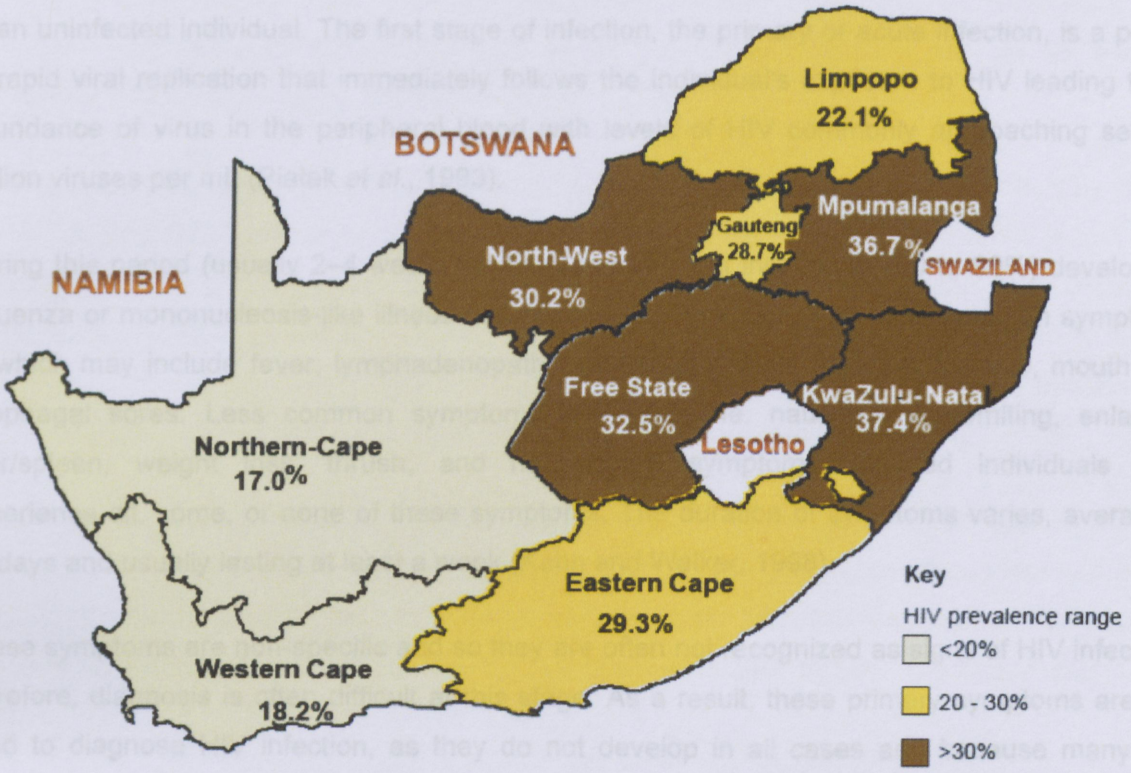


Figure 1.2: Percentage distribution of HIV-1 prevalence in South Africa. (Department of Health, South Africa, 2012) (Accessed, 10th June 2013).

1.3 PATHOGENESIS OF HIV

1.3.1 Primary infection of HIV

The initial infection with HIV generally occurs after transfer of body fluids from an infected person to an uninfected individual. The first stage of infection, the primary or acute infection, is a period of rapid viral replication that immediately follows the individual's exposure to HIV leading to an abundance of virus in the peripheral blood with levels of HIV commonly approaching several million viruses per mL (Piatak *et al.*, 1993).

During this period (usually 2–4 weeks post-exposure), most individuals (80 to 90%) develop an influenza or mononucleosis-like illness called acute HIV infection. The most common symptoms of which may include fever, lymphadenopathy, pharyngitis, rash, myalgia, malaise, mouth and esophagal sores. Less common symptoms are headache, nausea and vomiting, enlarged liver/spleen, weight loss, thrush, and neurological symptoms. Infected individuals may experience all, some, or none of these symptoms. The duration of symptoms varies, averaging 28 days and usually lasting at least a week (Kahn and Walker, 1998).

These symptoms are non-specific and so they are often not recognized as signs of HIV infection, therefore, diagnosis is often difficult at this stage. As a result, these primary symptoms are not used to diagnose HIV infection, as they do not develop in all cases and because many are caused by other more common infections. However, recognizing the syndrome can be important because the patient is much more infectious during this period (Daar *et al.*, 2001).

1.3.2 Chronic asymptomatic HIV-1 infection

A strong immune defense reduces the number of viral particles in the blood stream, marking the start of the infection's clinical latency stage. Clinical latency can vary between two weeks and 20 years. During this phase of infection, HIV is active within lymphoid organs, where large amounts of virus become trapped in the follicular dendritic cells network (Burton *et al.*, 2002) and the surrounding tissues that are rich in $CD4^+$ T cells may also become infected, with viral particles accumulating both in infected cells and as free viruses (Clapham and McKnight, 2001).

1.3.3 Late stage of HIV

When $CD4^+$ T cell numbers decline below a critical level of 200 cells per µL, cell-mediated immunity is lost, and infections with a variety of opportunistic microbes appear. The first symptoms often include moderate and unexplained weight loss, chronic respiratory tract infections (such as sinusitis, bronchitis, otitis media, pharyngitis, prostatitis, skin rashes, and oral ulcerations) (Kahn and Walker, 1989).

Common opportunistic infections and tumors, most of which are normally controlled by robust $CD4^+$ T cell-mediated immunity then start to affect the patient. Typically, resistance is lost early on to oral Candida species and to *Mycobacterium tuberculosis*, which leads to an increased susceptibility to oral candidiasis (thrush) and tuberculosis. Later, reactivation of latent herpes viruses may cause worsening recurrence of herpes simplex eruptions, shingles, Epstein-Barr virus-induced B-cell lymphomas, or Kaposi's sarcoma (Kahn and Walker, 1989).

Pneumonia caused by the fungus *Pneumocystis jirovecii* is common and often fatal. In the final stages of AIDS, infection with cytomegalovirus (another herpes virus) or *Mycobacterium avium* complex is more prominent. Not all patients with AIDS get all these infections or tumors, and there are other tumors and infections that are less prominent but still significant (Kahn and Walker, 1989).

In summary of the pathogenesis of HIV, figure 1.3 shows a typical graph of the relationship between HIV viral load and the $CD4^+$ cells count perceived to be seen on the average course of an individual not receiving treatment.

DISEASE PROGRESSION IN HIV-1

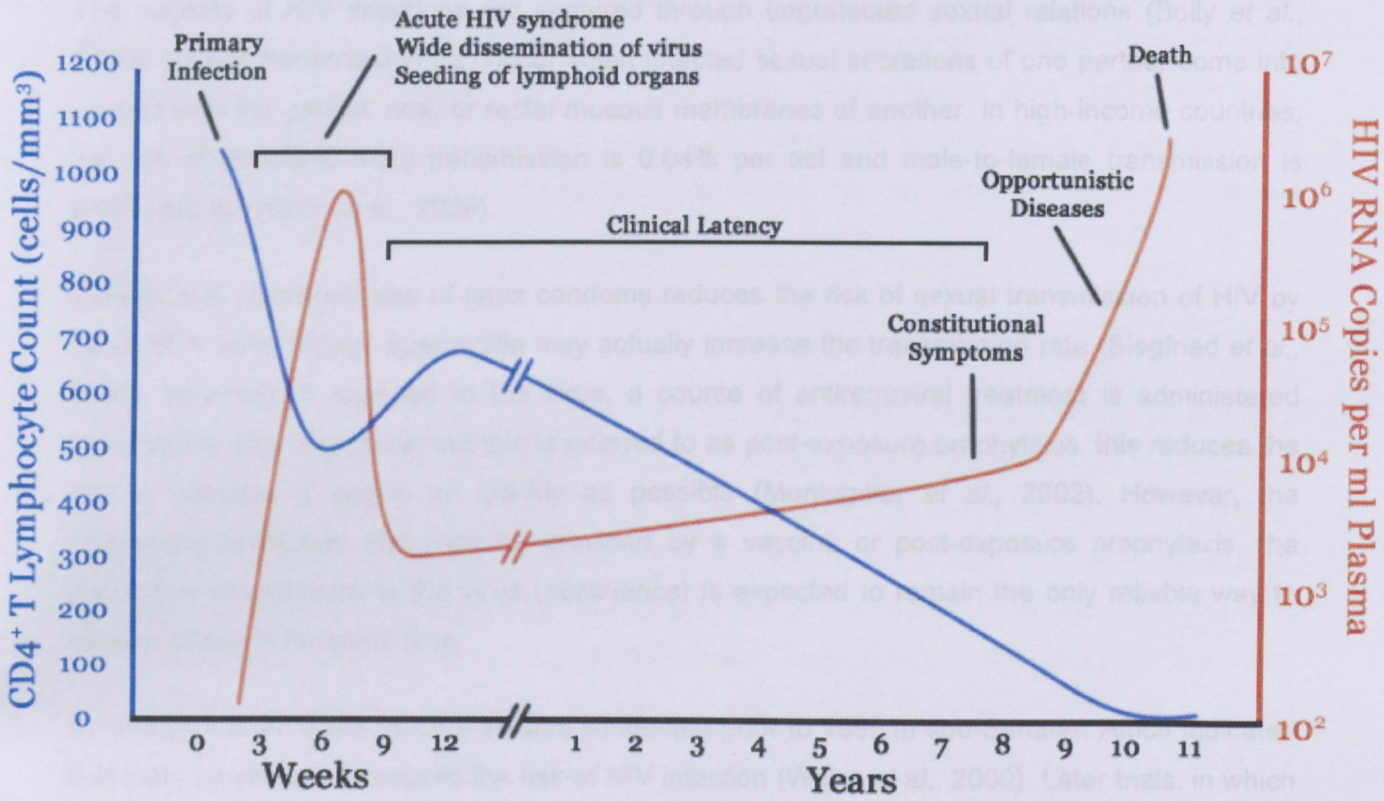


Figure 1.3: A typical graph showing disease progression in a HIV treatment naïve individual. The blue line represents CD4+ T-cell count (cells/mm^3) and the red line represents HIV viral copies per mL of plasma. (http://www.healthhype.com/cd4-count-dropping-viral-load-stable-in-hiv-infection-graph.html) (Accessed, 25th May 2013).

1.4 TRANSMISSION AND PREVENTION OF HIV

Three main transmission routes for HIV have been identified.

1.4.1 Sexual route

The majority of HIV infections are acquired through unprotected sexual relations (Boily *et al.*, 2009). Sexual transmission can occur when infected sexual secretions of one partner come into contact with the genital, oral, or rectal mucous membranes of another. In high-income countries, the risk of female-to-male transmission is 0.04% per act and male-to-female transmission is 0.08% per act (Boily *et al.*, 2009).

Correct and consistent use of latex condoms reduces the risk of sexual transmission of HIV by about 85% even though spermicide may actually increase the transmission rate (Siegfried *et al.*, 2005). However, if exposed to the virus, a course of antiretroviral treatment is administered immediately after exposure and this is referred to as post-exposure prophylaxis, this reduces the risk of infection if begun as quickly as possible (Montagnier *et al.*, 2002). However, the incomplete protection that may be provided by a vaccine or post-exposure prophylaxis, the avoidance of exposure to the virus (abstinence) is expected to remain the only reliable way to escape infection for some time.

An analysis of 27 observational studies conducted prior to 1999 in sub-Saharan Africa indicated that male circumcision reduces the risk of HIV infection (Weiss *et al.*, 2000). Later trials, in which uncircumcised men were randomly assigned to be medically circumcised in sterile conditions and given counseling and other men were not circumcised, have been conducted in South Africa, Kenya, (Bailey *et al.*,2007) and Uganda (Gray, 2007) showing reductions in female-to-male sexual HIV transmission of 60%, 53%, and 51% respectively.

As a result, a panel of experts convened by WHO and the UNAIDS stated that male circumcision is an efficacious intervention for HIV prevention but should be carried out by well-trained medical professionals and under conditions of informed consent i.e parents consent for infant boys (WHO, 2007: WHO, 2009). WHO and the CDC indicate that male circumcision may not reduce the transmission of HIV from men to women and that there is lack of sufficient data for the transmission rate of men who engage in anal sex with a female partner (WHO, 2007: CDC, 2008). Hence, WHO and UNAIDS recommend that circumcision only provide partial protection from HIV and should not replace other recommended methods of prevention (WHO, 2007).

1.4.2 Blood and blood products

In general if infected blood comes into contact with any open wound, HIV may be transmitted. This transmission route can account for infections in intravenous drug users, hemophiliacs and recipients of blood transfusions (though most transfusions are usually tested for HIV) and blood products (Lifson, 1988). Health care workers such as nurses, laboratory workers, and doctors have also been infected, although this occurs more rarely. Since transmission of HIV by blood became known, medical personnels are required to protect themselves from contact with blood by the use of universal precautions as wearing protective gloves and garments. People who give and receive tattoos, piercings and scarifications procedures can also be at risk of infection (Lifson, 1988; CDC, 2013).

1.4.3 Mother-to-child

The transmission of the virus from the mother to the child can occur in utero (during pregnancy), intrapartum (at childbirth) or via breast feeding. In the absence of treatment, the transmission rate up to birth between the mother and child is around 25% and 35% (Coovadia, 2004). However, where combination antiretroviral drug treatment and cesarian section are available, this risk can be reduced to as low as one percent (Coovadia, 2004). Postnatal mother-to-child transmission may be largely prevented by complete avoidance of breast feeding; however, this has significant associated morbidity. Exclusive breast feeding and the provision of extended antiretroviral prophylaxis to the infant are also efficacious in avoiding transmission (Hu *et al.*, 2005).

1.5 LABORATORY DIAGNOSIS OF HIV

HIV-1 testing consists of initial screening with an enzyme-linked immunosorbent assay (ELISA), to detect antibodies to HIV-1. Specimens with a nonreactive result from the initial ELISA are considered HIV-negative. Specimens with a reactive ELISA result are retested in duplicate (Celum *et al.*, 1991; Butto *et al.*, 2010). If the result of either duplicate test is reactive, the specimen is reported as repeatedly reactive and undergoes confirmatory testing with a more specific supplemental test (e.g. western blot or, less commonly, an immunoflourensence assay (IFA). Only specimens that are repeatedly reactive by ELISA and positive by IFA or reactive by Western blot are considered HIV-positive and indicative of HIV infection. Specimens that are repeatedly ELISA-reactive and occasionally provide an indeterminate Western blot result which

may be either an incomplete antibody response to HIV or nonspecific reactions in an uninfected person (Celum *et al.*, 1991; Butto *et al.*, 2010).

Although IFA can be used to confirm infection in these ambiguous cases, this assay is not widely used. Generally, a second specimen should be collected a month later and tested for persons with indeterminate Western blot results. Nucleic acid testing (e.g viral RNA or proviral DNA amplification method) can also help diagnosis in certain situations. In addition, a few tested specimens might provide inconclusive results because of a low quantity specimen. In these situations, a second specimen is collected and tested for HIV infection (Celum *et al.*, 1991; Butto *et al.*, 2010).

New born babies get antibodies from their mothers, therefore, for babies born to HIV infected mothers ELISA is not an appropriate test to use in the diagnosis of HIV. Polymerase chain reaction test (PCR) is used alternatively to diagnose because it is not an antibody/antigen based test and this shows a sensitivity of 100 %, but occasionally (in 2- 5 % of cases) led to false positive results (Hecht, 2002; Butto *et al.*, 2010).

1.6 MORPHOLOGY AND STRUCTURE OF HIV

HIV-1 particles have a diameter of 100 nm and are surrounded by a lipoprotein membrane. Each of the viral particles contains 72 glycoprotein complexes, which are integrated into this membrane, composed of trimers of an external glycoprotein gp120 and transmembrane spanning protein gp41. The bonding between gp120 and gp41 is unstable and therefore leading to the possibility of gp120 being shed spontaneously within the local environment (Hoffmann *et al.*, 2007).

Detection of gp120 in the serum and lymphatic tissue of HIV-infected patients is possible, of which during the process of budding, the virus may also incorporate different host proteins from the membrane of the host into its lipoprotein layer, as in human leukocyte antigen (HLA) class I and II proteins, or adhesion proteins such as Inter-Cellular Adhesion Molecule 1 (ICAM-1) that may facilitate adhesion to other target cells (Hoffmann *et al.*, 2007).

P17 which is the matrix protein is anchored to the inside of the viral lipoprotein membrane. The p24 core antigen contains two copies of HIV-1 RNA. The HIV-1 RNA is part of a protein-nucleic acid complex, which is composed of the nucleoprotein p7 and the reverse transcriptase p66

(RT). All enzymatic equipment that is necessary for replication; a reverse transcriptase (RT), an integrase p32 and a protease p11 are contained in the viral particle (Hoffmann *et al.*, 2007). A typical structure of the HIV virion is represented in figure 1.4.

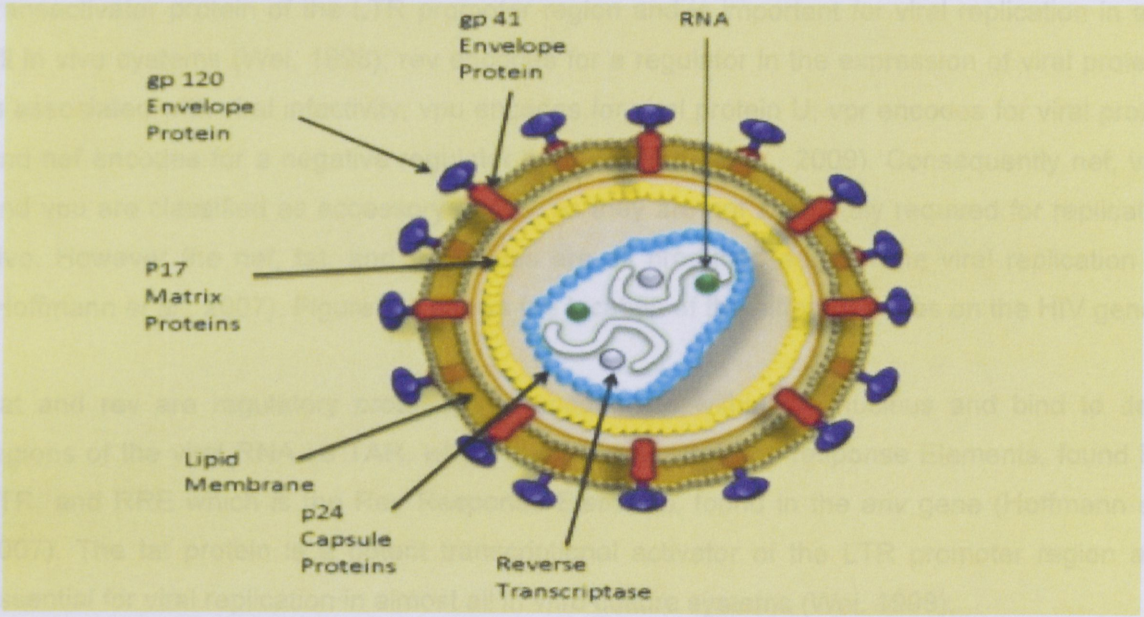


Figure 1.4: The structure of an HIV virion and its composition. (www. micro.magnet.fsu.edu) (Accessed on the 6th April 2013)

1.7 HIV GENOME ORGANIZATION

Retroviruses are characterized by three main genes: Gag (group associated antigens), Pol (polymerase) and Env (envelope) (Wong-Staal,1991).The structural scheme of a retroviral genome is: 5-LTR-gag-pol-env-LTR-3 as shown in figure 1.5. The LTR (long terminal repeats) region represents the two end parts of the viral genome, that are connected to cellular DNA of the host cell after integration and are not known to encode for any viral protein (Hoffmann *et al.*, 2007). Gag codes for gag poly-proteins processed during maturation into matrix protein p17

(MA), capsid protein p24 (CA), spacer peptide 1 p2 (SP1), nucleocaspid protein P7 (NC), spacer peptide 2 (SP2) and p6 (Watts *et al.*, 2009). Viral enzymes protease (p11), reverse transcriptase (p66/55; alpha and beta subunits) and integrase (p32) are coded by the Pol gene. Env codes for gp160, the precursor to gp120 and gp41 proteins embedded in the viral envelope allowing the virus to attach to and fuse with target cells (Watts *et al.*, 2009).

HIV also contains six other genes (vif, vpu, vpr, tat, rev and nef) in its 9 KB RNA. Tat encodes for transactivator protein of the LTR promoter region and is important for viral replication in almost all in vivo systems (Wei, 1998); rev encodes for a regulator in the expression of viral protein; vif is associated with viral infectivity; vpu encodes for viral protein U; vpr encodes for viral protein R and nef encodes for a negative regulator protein (Watts *et al.*, 2009). Consequently nef, vif, vpr and vpu are classified as accessory genes, as they are not absolutely required for replication in vivo. However the nef, tat, and rev genes are all produced early in the viral replication cycle (Hoffmann *et al.*, 2007). Figure 1.5 shows the location of the different genes on the HIV genome.

Tat and rev are regulatory proteins that accumulate within the nucleus and bind to defined regions of the viral RNA i.e TAR, which is the Transactivation-response Elements, found in the LTR; and RRE which is the Rev Response Elements, found in the *env* gene (Hoffmann *et al.*, 2007). The tat protein is a potent transcriptional activator of the LTR promoter region and is essential for viral replication in almost all in vitro culture systems (Wei, 1998).

Tat and rev are responsible for stimulating the transcription of HIV-1 DNA into RNA, promote RNA elongation and enhance the transportation of HIV RNA from the nucleus to the cytoplasm and are also essential for translation. Rev is also a nuclear export factor, is important for switching from the early expression to the structural proteins that are synthesized later during transcription (Hoffmann *et al.*, 2007).

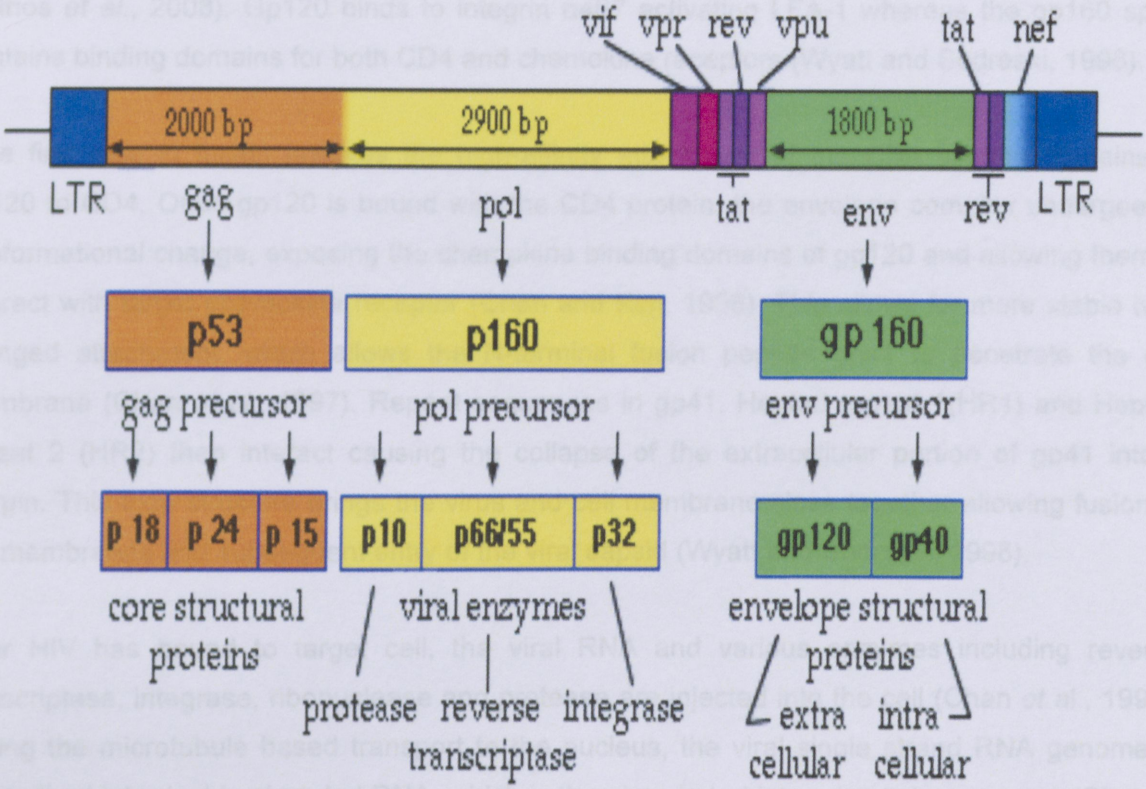


Figure 1.5: Schematic representation of HIV genes. (www.yale.edu/bio243/HIV/genome.html) (Accessed on 6th April 2013).

1.8 THE LIFE CYCLE OF HIV

1.8.1 HIV binding and entry

HIV enters the macrophages and CD4+ T-cells by the adsorption of the viral surface glycoproteins to receptors on the target cell followed by fusion of the viral envelope with the cell membrane and the release of the HIV capsid in the cell (Chan *et al.*, 1998).

The entry process begins through interaction of the trimeric envelope complex (gp160 spikes) and both CD4 and a chemokine co-receptor (either CCR5 or CXCR4) on the cell surface (Wyatt and Sodroski, 1998). Leucocyte function associated antigen (LFA) is the central integrin involved

in the establishment of virological synapse, which facilitate efficient cell-to-cell spreading of HIV-1 (Arthos *et al.*, 2008). Gp120 binds to integrin α4β7 activating LFA-1 whereas the gp160 spike contains binding domains for both CD4 and chemokine receptors (Wyatt and Sodroski, 1998).

The first step in fusion involves the high-affinity attachment of the CD4 binding domains of gp120 to CD4. Once gp120 is bound with the CD4 protein, the envelope complex undergoes a conformational change, exposing the chemokine binding domains of gp120 and allowing them to interact with target chemokine receptor (Chan and Kim, 1998). This allows for more stable two-pronged attachment, which allows the N-terminal fusion peptide gp41 to penetrate the cell membrane (Chan *et al.*, 1997). Repeat sequences in gp41, Heptad repeat 1(HR1) and Heptad repeat 2 (HR2) then interact causing the collapse of the extracellular portion of gp41 into a hairpin. This loop structure brings the virus and cell membrane close together allowing fusion of the membranes and subsequent entry of the viral capsid (Wyatt and Sodroski, 1998).

After HIV has bound to target cell, the viral RNA and various enzymes including reverse transcriptase, integrase, ribonuclease and protease are injected into the cell (Chan *et al.*, 1998). During the microtubule based transport to the nucleus, the viral single strand RNA genome is transcribed into double stranded DNA, which is then integrated into a host chromosome (Chan *et al.*, 1998).

1.8.2 Replication and transcription

Shortly after the viral capsid enters the cell, reverse transcriptase liberates the single-stranded (+) RNA genome from the attached viral protein and copies it into a complementary DNA (cDNA) molecule (Zheng *et al.*, 2005). The process of reverse transcription is extremely error-prone and the resulting mutations may cause drug resistance or allow the virus to evade the body's immune system. The reverse transcriptase also has ribonuclease activity that degrades viral RNA during synthesis of viral DNA, as well as DNA-dependant DNA polymerase that creates a sense DNA from the antisense cDNA (Doc Kaiser, 2008). Together the cDNA and its complement form a double stranded viral DNA that is then transported into the cell's nucleus. The integration of the viral DNA into the host genome is carried out by another viral enzyme called integrase (Zheng *et al.*, 2005).

During viral replication, the integrated DNA provirus is transcribed into mRNA which is then spliced into smaller pieces (exons). The exons are exported by the regulatory protein tat and rev. As newly produced rev protein accumulates in the nucleus, it binds to viral mRNA and allows unspliced RNAs to leave the nucleus, where they are otherwise retained until spliced (Pollard and Malim, 1998). At this stage, the structural protein gag and env are produced from full-length mRNA. The full-length RNA is actually the virus genome, it binds to the gag protein and is packaged into new virus particles (Pollard and Malim, 1998).

1.8.3 Assembly and release

The final step of the viral cycle is the assembly of new HIV-1 virions which begins at the plasma membrane of the host cell (Gelderblom, 1997). The env protein (gp160) goes through the endoplasmic reticulum and is transported to the golgi complex where it is cleaved by HIV protease and processed into gp41 and gp120 (Gelderblom, 1997).

Glycoproteins 41 and 120 are transported to the plasma membrane of the host cell where gp41 anchors the gp120 to the membrane of the infected cell. The gag (p55) and gag-pol (160) polyproteins also associate with the inner surface of the plasma membrane along with HIV genomic DNA as the forming virion begins to bud from host cell (Gelderblom, 1997).

During maturation, HIV protease cleaves the polyproteins into individual functional HIV provirions and enzyme. The various structural components then assemble to produce a mature HIV virion and the mature virus is then able to infect another cell (Gelderblom, 1997). Figure 1.6 summarizes the life cycle of HIV.

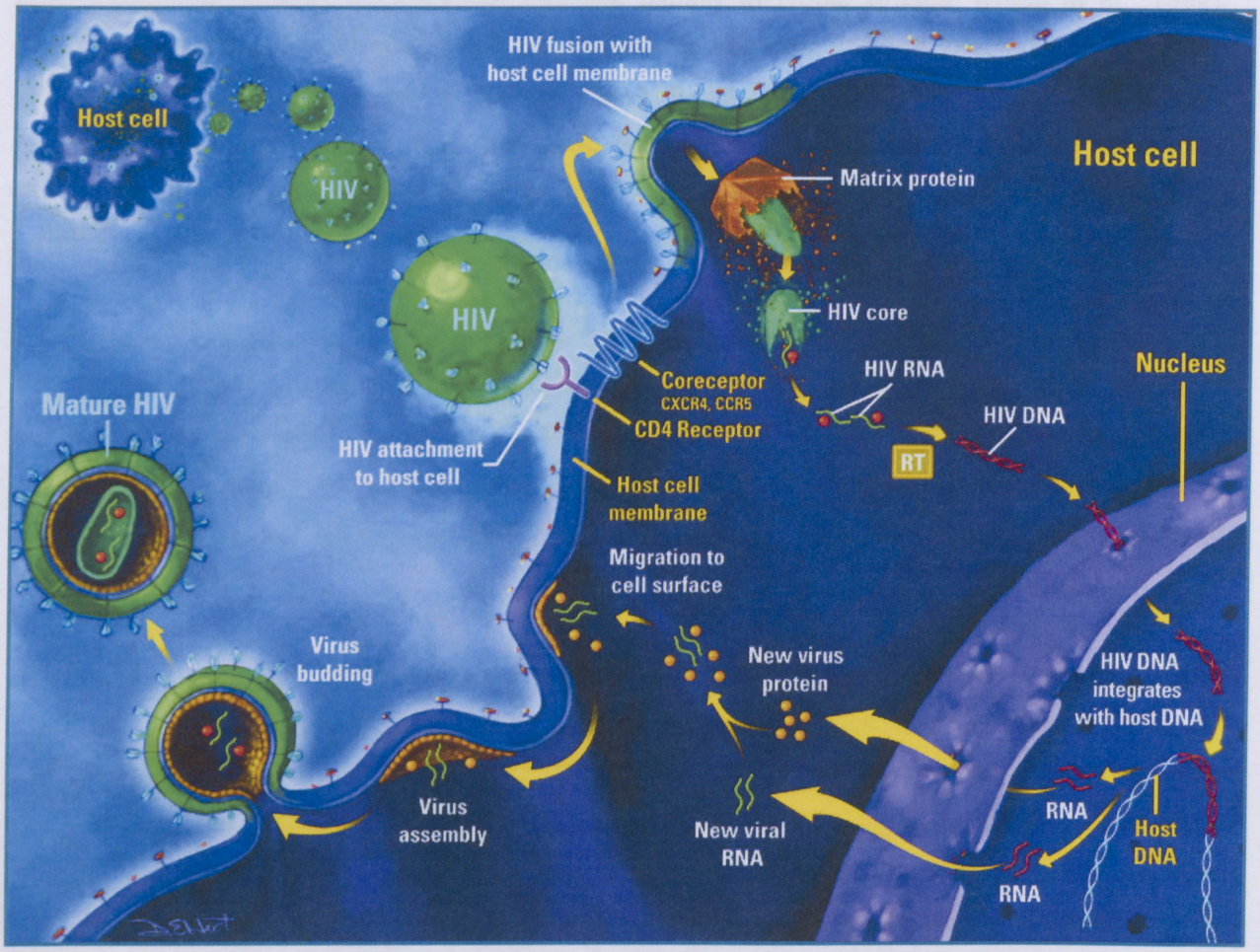


Figure 1.6: A summary of the HIV life cycle adapted from the University of Washington (www.romsla8.blogspot.com) (Accessed, 15th June 2013).

1.9 HIV TROPISM

A quintessential property of HIV and other primate lentiviruses is the sequential use of $CD4^+$ cells and a second cellular receptor during entry into susceptible cells. The main cellular targets for HIV-1 are CD4+ T-helper subset of lymphocytes and CD4+ cells of the macrophage lineage (Collman *et al.*, 1989).

All HIV-1 strains can productively infect activated Peripheral Blood Mononuclear Cells (PBMC), some can also replicate in cultures of monocytes derived macrophages (MDM) and have been classified as macrophage or M-Tropic (Collman *et al.*, 1989). Viruses that are isolated shortly

after a person has been infected with HIV are usually M-Tropic viruses. (Roos *et al.*, 1992). Other isolates that are recovered late in the disease course from persons infected with HIV-1 can establish infections in continuous human CD4+ T-cell lines but not MDM and have been designated T-cell line or TCL-Tropic Viruses. Subsequent studies have reported that primary HIV-1 viruses are able to infect both MDM and T-cell lines. Such strains were classified as dual-tropic viruses (Collman *et al.*, 1989).

Chemokine receptors (α and β) function as co-receptors for virus fusion and entry into susceptible CD4+ cells (Choe *et al.*, 1996). Hence viruses are grouped on the basis of chemokine receptor usage in fusion assays in which HIV-1 120gp and CD4 plus co-receptor proteins are expressed in separate cells or in entry assays in which small molecules inhibitors specific for co-receptor CCR5 or CXCR4, block single cycle or spreading viral infections (Baba *et al.*, 1999).

HIV-1 viruses using the CXCR4 receptor are usually T-cell lines tropic viruses. In contrast HIV-1 stains exclusively using the CCR5 (R5 viruses) are predominantly M-tropic.(Simmons *et al.*, 1996). M-tropic, CCR5 viruses are the predominant HIV-1 phenotypes detected in recently infected individual and remain the dominant viruses throughout the asymptomatic phase of HIV-1 infections whereas in some infected persons, CXCR4 (X4) virus strains emerge during the later symptomatic stage of infections (Zhu *et al.*,1993). CCR5 are exclusively expressed on the memory CD4+ T lymphocytes and at significantly higher levels on memory cells residing in non-lymphoid tissues such as the gastrointestinal tract and lungs (Douek *et al.*, 2003).

CXCR4 viruses are expressed on naïve and memory CD4+ T cells. In peripherial blood, where most of the circulating lymphocytes are naïve, 80% - 95% of CD4+ T cells express CXCR4, while only 5% -10% produce detectable CCR5 (Grivel and Margolis, 1999).

On the other hand, HIV-1 infected macrophages contribute to the viral persistence and distribution to the central nervous system (CNS), lungs, lymph nodes, bone marrow and gastrointestinal tract. Furthermore, these macrophages appear to be the main cell type involved in mucosal transmission (Grivel and Margolis, 1999). HIV-1 replication in monocytes or macrophages lineage microglia within the CNS may persist in patients receiving HAART because of ineffective transfer of drugs across the blood brain barrier. This viral reservoir could

contribute to the rebound of plasma viremia invariably observed after termination of drug administration (Grivel and Margolis, 1999).

Another variant of the CCR5 reported is the (CCR5 Δ32) mutation in humans is particularly abundant in populations of European origin. A 32bp deletion in the region encoding the second extracellular loop of the receptor results in a frame shift and early termination of the polypeptide chain. The resulting protein lacks the last three transmembrane segments and is not processed properly to the plasma membrane (Liu *et al.*, 1996).

The mutant is not functional as a chemokine receptor, nor does it act as a co-receptor for HIV. Homozygotes for the Δ32 mutation are highly protected against HIV infection (Liu *et al.*, 1996). The HIV resistance of Δ32 homozygotes has contributed greatly to the demonstration of the central role of CCR5 as a co-receptor for M-tropic strains of HIV. However, this protection is not complete, since a small number of homozygotes were found to be seropositive (Biti *et al.*, 1997; O'Brien *et al.*, Theodorou *et al.*, 1997). The strain responsible for the infection of one of these individuals was identified as a T-tropic strain utilizing CXCR4. However heterozygotes for the Δ32 mutation were reported to have on average delayed progression to clinical stages of AIDS (Michael *et al.*, 1998).

1.10 HIV CLASSIFICATION AND GENETIC VARIATION

HIV belongs to the family Retroviridae, which is subdivided into three groups; Spumavirus, Oncovirus and Lentivirus (Montagnier, 2002). HIV belongs to the genus lentivirus under the subfamily orthovirinae because its virions are spherical and is a vertebrate virus (RNA tumor virus group) (Barre-sinoussi, 1996).

There are two types of HIV: HIV type 1 (HIV-1) and HIV type 2 (HIV-2). HIV-1 is similar to a chimpanzee subspecies simian immunodeficiency virus (SIV_{cpz}) discovered in Central Africa while HIV-2 resembles the simian immunodeficiency virus among sooty mangabegs (SIV_{smm}) discovered in West Africa. These viruses (HIV-1 and HIV-2) of human are probably a result of cross-species transmission from their animal homologs (Zoonotic transmission) (Barre-sinoussi, 1996). SIV_{cpz} and SIV_{smm} are primate lentiviruses with a genomic structure similar to HIV-1 and HIV-2 respectively and both are non-pathogenic in their natural hosts (Barre-sinoussi, 1996). The relatedness of HIV and SIV is represented in figure 1.7.

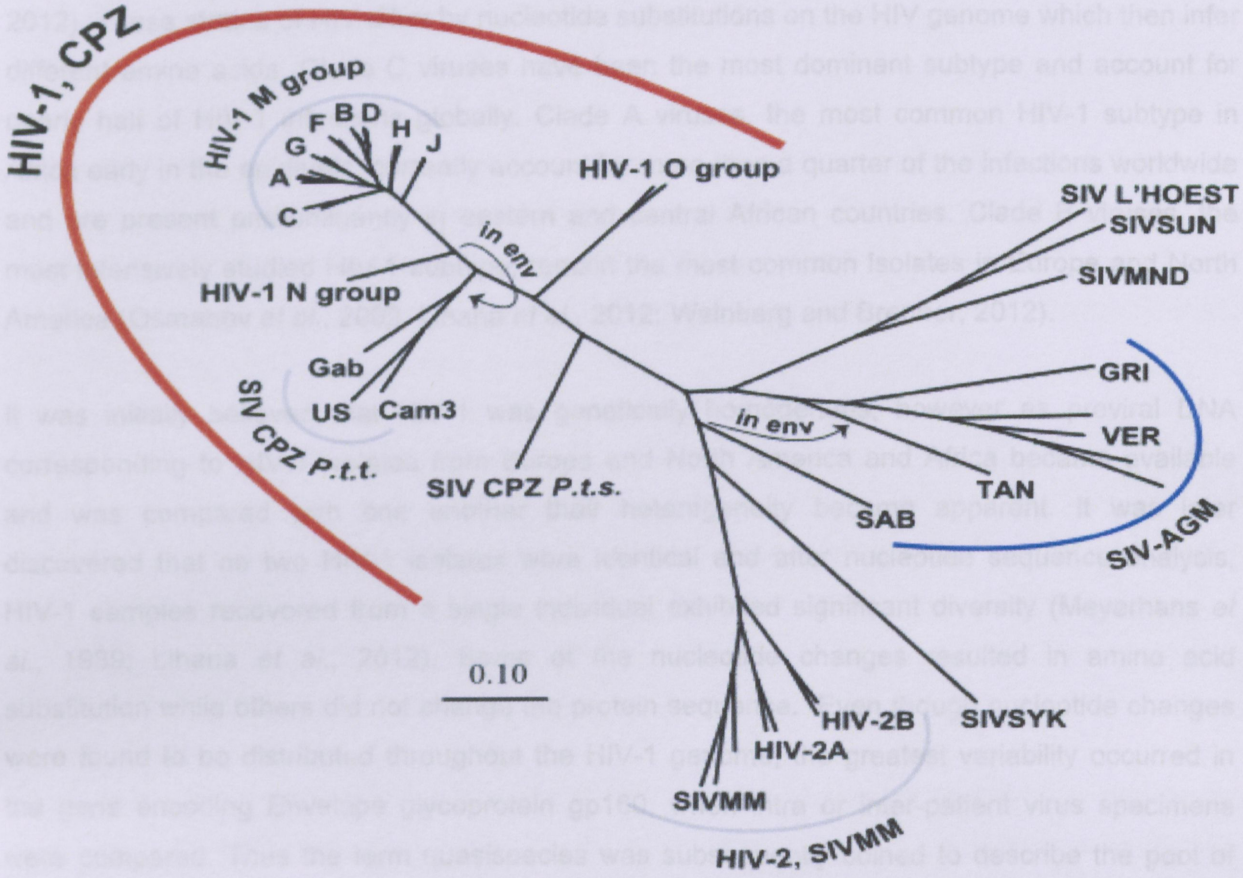


Figure 1.7: A phylogenetic tree of SIV and HIV showing HIV-1 groups and subgroups. (www.hiv.lanl.gov/content/sequence/HIV/COMPENDIUM/1999/intro.pdf) (Accessed on, 27th May 2013). The recently designated Group P is not represented in this tree.

HIV-1 is defined and classified into four groups: M group (Main), O group (Outlier), N group (non-M, non-O) (Carr *et al.*, 1998; Lihana *et al.*, 2012) and Group P, a new strain closed related to gorilla simian immunodeficiency virus discovered in a Cameroonian woman in 2009 (Plantier, 2009; Vallari *et al.*, 2011). Occasionally an individual may be infected with two genetically different viruses creating a hybrid. Most of these hybrids do not survive for long but those that infect more than one person are known as circulating recombinant forms (CRFs) (Carr *et al.*, 1998; Lihana *et al.*, 2012).

Nine subgroups and 15 circulating recombinant forms (CRFs) create genetic forms of the HIV-1 group M lineage that accounts for 95% of global virus isolates. The subgroups are A, B, C, D, F, G, H, J and K. All of the HIV-1 group M subgroups can be found in Africa, with a rapid spread of HIV-1 to southern Africa during the mid 1990s (Osmanov *et al.*, 2002; Wainberg and Brenner, 2012). These strains of HIV differ by nucleotide substitutions on the HIV genome which then infer different amino acids. Clade C viruses have been the most dominant subtype and account for nearly half of HIV-1 infections globally. Clade A viruses, the most common HIV-1 subtype in Africa early in the epidemic, currently account for more than a quarter of the infections worldwide and are present predominantly in eastern and central African countries. Clade B viruses, the most intensively studied HIV-1 subtype, remain the most common isolates in Europe and North America (Osmanov *et al.*, 2002; Lihana *et al.*, 2012; Wainberg and Brenner, 2012).

It was initially believed that HIV-1 was genetically homogenous, however as proviral DNA corresponding to HIV-1 isolates from Europe and North America and Africa became available and was compared with one another their heterogeneity became apparent. It was later discovered that no two HIV-1 isolates were identical and after nucleotide sequence analysis, HIV-1 samples recovered from a single individual exhibited significant diversity (Meyerhans *et al.*, 1989; Lihana *et al.*, 2012). Some of the nucleotide changes resulted in amino acid substitution while others did not change the protein sequence. Even though nucleotide changes were found to be distributed throughout the HIV-1 genome, the greatest variability occurred in the gene encoding Envelope glycoprotein gp160, when intra or inter-patient virus specimens were compared. Thus the term quasispecies was subsequently coined to describe the pool of diverse and changing populations of virus present in an individual infected with HIV-1 (Meyerhans *et al.*, 1989; Lihana *et al.*, 2012).

HIV isolates show extensive genetic variability and several factors such as: (a) error-prone viral DNA synthesis during reverse transcription (3×10^{-5} mutations/nucleotide/replication cycle) (b) high recombination frequencies accompanying reverse transcription, (c) the high levels of progeny virus production in vivo (10^9 particles/day; 150 to 300 replication cycles/ year), and (d) large numbers of infected individuals, contribute to this diversity (Mansky *et al.*, 1995).

High rates of genetic recombination are generated as a consequence of a single cell being infected by genetically distinct viral strains and arise as a result of template switching between two different viral RNAs during the reverse transcription process (Wainberg and Brenner, 2012).

By 1995, it became apparent that a significant fraction of global diversity of HIV-1 group M viruses include inter-clade CRFs which arose by the intermixing of viruses co-circulating in a particular geographical location. Currently, HIV-1 recombinants are found in geographical areas such Africa, South America and South Asia where multiple HIV-1 subtypes coexist and they account for nearly 30% of circulating HIV-1 strains worldwide (Osmanov *et al.*, 2002; Wainberg and Brenner, 2012).

Most HIV-1 recombinants have arisen in Africa and the majority of them contain segments originally derived from clade A viruses. The two most common CRFs are CRF01 and CRF02, a clade A/ clade E recombinant originally isolated in Thailand and a clade A/clade G recombinant which is responsible for nearly one third of new HIV-1 infections in west Africa. (Osmanov *et al.*, 2002; Wainberg and Brenner, 2012).

1.11. THE HIV-1 POLYMERASE GENE

The HIV-1 polymerase (Pol) gene codes for essential viral enzymes, protease (p10), reverse transcriptase (p66/55; alpha and beta subunits) and integrase (p32).

1.11.1 PROTEASE

HIV-1 protease is essential for the life-cycle of HIV. It cleaves newly synthesized poly-proteins to create mature protein components of an infectious HIV virion. Hence, if the enzyme is ineffective, HIV virions remain un-infectious (Krausslich, 1991).

It uses two aspartic acid (Asp) residues in the active site to co-ordinate a water molecule that catalyzes the hydrolysis of a peptide bond in the target protein. It functions as a dimer; monomers are held together by a four-stranded, anti-parallel β-sheet derived from the N and C-terminal ends of each monomer. The substrate binding site is located within a cleft formed between the two monomers.

The protease dimer contains flexible flaps that overhang the binding site and may stabilize the substrate within the cleft and the active site whereas the asp residues lie in the center of the dimer. The dimeric enzyme cleaves a number of sites in both Gag and Gag-Pol. The efficiency with which Protease cleaves the individual target sites in Gag and Gag-Pol varies widely and is

influenced by two major factors: the amino acid sequence at the site of cleavage and the degree of exposure or accessibility of the cleave site (Billich *et al.*, 1988). Mutations in Gag that disrupt the ordering nature of protease-mediated processing severely disrupt virus assembly or subsequent maturation. In addition, HIV-1 mutants engineered to over-express protease exhibit rapid, premature processing of Gag and Gag-Pol poly-protein and a block in virus production. Thus the activation of protease must be tightly controlled to prevent significant processing before assembly completion (Krausslich, 1991).

1.11.2 REVERSE TRANSCRIPTASE

Retroviruses possess the ability to convert single stranded RNA (ssRNA) to double stranded DNA (dsDNA) during the early stages of the infection process (Baltimore, 1970; Sarafianos *et al.*, 2001). HIV-1 reverse transcriptase is the enzyme that catalyzes this activity. The newly created DNA can then be incorporated into the host genome.

The HIV-1 reverse transcriptase contains two main domains: A DNA polymerase domain that either copies an RNA templates (for minus-strand DNA synthesis) or DNA template (for second or plus-strand DNA synthesis), and a ribonuclease H (RNase H) domain (for the cleavage and degradation of the template RNA after DNA synthesis so that the newly made DNA can generate a second DNA strand.) (Sarafianos *et al*, 2001).

The structure of HIV-1 reverse transcriptase reveals that p66 and p51 subunits are folded into similar sub-domains but are arranged differently. The p66 subunit can be visualized as the right hand, with the polymerase active site within the palm and a deep template binding cleft formed by the palm, fingers and thumb sub-domains (Jacobo-Molina *et al.*, 1993). The polymerase is linked to RNaseH by the connection sub-domain. The active site located in the palm contains three critical Asp residues (Asp- 110, Asp-185 and Asp 186) in close proximity and two coordinated Mg^{2+} ions. Mutation of the Asp residues abolishes the polymerase activity of Reverse Transcriptase. The orientation of the three active site Asp residues is however similar to that observed in other DNA polymerases such as that of the klenow fragment of *Escherichia coli* DNA polymerase 1 (Arnold *et al.*, 1995).

The p51 subunit appears to play a structural role and does not form a polymerizing cleft; Asp-110, Asp-185 and Asp186 are buried within the p51 subunit. Approximately 18 bp of the primer/template duplex lie in the nucleic acid binding sites (Arnold *et al.*, 1995).

Reverse transcription of the retroviral RNA genome to a dsDNA copy proceeds via a series of steps. (a) Minus-strand DNA synthesis is initiated from the 3'-OH of the tRNA bound to the PBS (Primer binding sites) then DNA synthesis proceeds to the 5' end of the genome. (b) RNaseH digests the RNA portion of the newly formed RNA-DNA hybrid, freeing the resulting short, single-stranded DNA (ssDNA) fragment known as the minus strand strong-stop DNA. (c) The minus-strand strong stop DNA is transferred to the 3'end of the genome where it hybridizes by virtue of a short region of homology (the repeat or R region) present at both 5' and 3' ends of the RNA genome. (d) Minus-strand synthesis, accompanied by RNaseH mediated degradation of the RNA in the resulting RNA-DNA hybrid, continues to the PBS at the 5' end of the genome, displacing the tRNA primer. (e) Fragments of RNA that were not removed by RNaseH serve as primers for plus-strand synthesis (Charneau *et al.*, 1992). (f) After plus-strand synthesis copies a portion of the tRNA primer, RNaseH displaces the tRNA, this then exposes the PBS at the 3' end of the plus-strand DNA, allowing the plus-strand DNA to hybridize with the homologous region at the 3' end of the minus-strand DNA (second strand transfer). (g) The plus and minus- strand synthesis proceed to completion. Plus-strand synthesis terminates at the end of the minus strand of HIV-1 at a sequence known as the central termination signal (Charnean *et al.*, 1994).

1.11.3 INTEGRASE

Integrase is a 32 kDa protein produced from the C-terminal portion of the Pol gene product (Savarino, 2006). It enables viral genetic material to be integrated into the DNA of the infected cell. This is a distinguishing replication factor of HIV and other retroviruses replication because of their ability to insert a DNA copy of the viral genome into the host cell chromosome after being reverse transcribed. Therefore the integrated viral DNA then serves as the template for the synthesis of viral RNA and maintained as part of the host cell genome for the lifetime of the infected cell.

Integrase is composed of three structurally and functionally distinct domains; an N- terminal domain, Zinc binding domain (a catalytic core domain) and a relatively non-conserved C-terminal domain. The N-terminal domain is composed of four helices with zinc coordinated by amino acid

His-12, His-12, Cys-40 and Cys- 43. The structure of the N-terminal domain is reminiscent of helical DNA-binding proteins that contain a *helix-turn helix motif;* in the HIV-1 structure this motif contributes to dimer formation (Wang *et al.*, 2001).

The catalytic core domain on integrase is composed of a five stranded β-sheet flanked by helices. The structure bears a striking resemblance to other poly-nucleotidyl tranferases, including RNaseH and the Bacteriophage MuA Transposase (Dyda *et al.*,1994). Three highly conserved residues are found in the analogous positions at the catalytic center of other poly-nucleotidyl transferases and in HIV-1 Integrase these residues are Asp-64, Asp-116 and Glu-152 and mutations at these positions blocks HIV-1 Integrase function both *In vivo* and *In vitro* (Dyda *et al.*, 1994). The C-terminal domain takes on a five-stranded β-barrel folding topology reminiscent of the Src Homology 3 (SH3) domain. It is clear that the core domain of integrase plays the central role in catalysis and the functions of the N- and C-terminal domains are less apparent. The core domain can therefore alone catalyze the disintegration reaction. However, mutations in the N-terminal domain and Zinc-binding domain disrupt Integrase tetramerization 3' processing and strand transfer (Engelman and Craigie, 1992).

The process of viral integration occurs in three main steps: (a) For 3' processing, integrase removes 2 or 3 nucleotides from the initially blunt 3' terminal of both strands of full length linear viral DNA, generating a pre-integration substrate with 3' recessed ends. (b) For DNA strand transfer, in the nucleus integrase catalyzes a staged cleavage in the cellular target DNA and the 3' recessed ends of viral DNA are joined to the 5' overhanging terminal of the cleaved cellular DNA. (c) For gap repair, the cellular repair machinery perhaps in conjunction with integrase fills the gaps thereby completing the integration step (Brown *et al.*, 1989).

1.12 TREATMENT OF HIV

There is currently no cure for HIV/AIDS (Robb, 2008). A lot of antiretroviral drugs have been approved for the treatment of HIV infection and these drugs fall into 6 different drug classes - nucleoside RT inhibitors (NRTI), non-nucleoside RT inhibitors (NNRTI), protease inhibitors (PI), entry inhibitors (EI), fusion inhibitors (FI) and integrase inhibitors (IN) (Shafer *et al.,* 2001).

Current treatment for HIV infection consists of highly active antiretroviral therapy (Montagnier *et al.,* 2002). This has been highly beneficial to many HIV-infected individuals since its introduction in 1996, when the protease inhibitor-based HAART initially became available (Palella *et al.,*

1998). Current HAART options as opposed to the previously used monotherapy, are combinations (or "cocktails") consisting of at least three drugs belonging to at least two types, or "classes," of antiretroviral agents. Typically, these classes are two nucleoside analogue reverse transcriptase inhibitors (NARTIs or NRTIs) plus either a protease inhibitor or non-nucleoside analogue reverse transcriptase inhibitors (NNRTI). South Africa is currently using protease and reverse transcriptase inhibitors and the treatment regimen is grouped into two categories; the first being for adults and adolescents and the other for infants and children (UNAIDS, 2012). The adults and adolescents category is for patients with a CD4 count of ≤ 200mm^3 or the WHO clinical stage 4 of CD4 count ≤ 350mm^3 and co-infected with TB or multiple drug resistant or extremely drug resistant tuberculosis (UNAIDS, 2011; Department of Health, South Africa, 2013). The first line regimen for this category is a combination of Tenofovir, Lamivudine or Emtricitabine and Efavirenz or Nevirapine whereas the second line is a combination of zidovudine, Lamivudine and Liponavir/Ritonavir (UNAIDS, 2012; Department of Health, South Africa, 2013).

The infants and children category is divided into two groups; one for infants/children of 12 months to 5 years with a CD4 ≤ 750mm^3 or WHO pediatric clinical stage 3 or 4 with a combination of Abacavir, Lamivudine and Liponavir/Ritonavir whereas the other is for children of over 5 years with CD4 count of 350mm^3 or WHO clinical stage 3 or 4 with a combination of Abacavir, Lamivudine and Efavirenz (UNAIDS, 2012; Department of Health, South Africa, 2013).

HAART however, neither cures the patient nor does it uniformly remove all symptoms. High levels of HIV, often HAART resistant, return if treatment is stopped (Dybul et al., 2002). Moreover, it can take more than a lifetime for HIV infection to be cleared using HAART. Despite this, many HIV-infected individuals have experienced remarkable improvements in their general health and quality of life, which has led to a large reduction in HIV-associated morbidity and mortality worldwide (Chene *et al.,* 2003).

One study suggests that the average life expectancy of an HIV infected individual is 52 years from the time of infection if treatment is started when the CD4 count is 350/μL (UNAIDS, 2010). In the absence of HAART, progression from HIV infection to AIDS has been observed to occur at a median of between nine to ten years and the median survival time after developing AIDS is only 9.2 months (Morgan *et al.,* 2002). However, HAART sometimes achieves far less than optimal results, in some circumstances being effective in less than fifty percent of patients. This is due to a variety of reasons such as medication intolerance/side effects, prior ineffective antiretroviral therapy and infection with a drug-resistant strain of HIV. However, non-adherence

and non-persistence with antiretroviral therapy is the major reason most individuals fail to benefit from HAART (Becker *et al.,* 2002).

The reasons for non-adherence and non-persistence with HAART are varied and overlapping. Major psychosocial issues, such as poor access to medical care, inadequate social supports, psychiatric disorder and drug abuse contribute to non-adherence. The complexity of these HAART regimens, whether due to dosing frequency or other issues along with side effects that create intentional non-adherence also contribute to this problem (Nieuwkerk *et al.,* 2001). Some of the side effects of anti retrovirals include lipodystrophy and insulin resistance for Stavudine, an increase in cardiovascular risks for Didanosine and Abacavir, etc (Saitoh *et al.,* 2005).

The timing for starting HIV treatment is still debated. There is no question that treatment should be started before the patient's CD4 count falls below 200, and most national guidelines state that the treatment should be started once the CD4 count falls below 350 but there is some evidence from cohort studies that treatment should be started before the CD4 count falls below 350. In those countries where CD4 counts are not available, patients with WHO stage III or IV disease should be offered treatment (Wang *et al.,* 2004).

Anti-retroviral drugs are expensive, and the majority of the world's infected individuals do not have access to medications and treatments for HIV and AIDS (Ferrantelli *et al.,* 2004). Research to improve current treatments includes decreasing side effects of current drugs, further simplifying drug regimens to improve adherence, and determining the best sequence of regimens to manage drug resistance. Unfortunately, only a vaccine is thought to be able to halt the pandemic. This is because a vaccine may cost less, thus being affordable for developing countries, and as it would not require daily administration (Ferrantell *et al.,* 2004). However, after years of research, HIV still remains a difficult target for a vaccine (Ferrantelli *et al.,* 2004).

1.13 HIV DRUG RESISTANCE

HIV-1 drug resistance is caused by mutations in the HIV genome that affect the molecular targets of antiretroviral drugs (Shafer *et al.*, 2001). Many antiretroviral drugs have been approved for the treatment of HIV-1 infection and these drugs fall into 6 different drug classes - nucleoside RT inhibitors (NRTI), non-nucleoside RT inhibitors (NNRTI), protease inhibitors (PI), fusion inhibitors (FI) and integrase inhibitors (IN) (Shafer *et al.*, 2001). In previously untreated individuals with drug-susceptible HIV-1 strains, combinations of three or more drugs from two different drug classes can lead to prolonged virus suppression and immunological reconstitution. (Shafer *et al.*, 2001).

1.13.1 NUCLEOTIDE REVERSE TRANSCRIPTASE INHIBITORS MECHANISM OF ACTION AND DRUG RESISTANT MUTATIONS

The nucleoside RT inhibitors are analogues of normal nucleotides. They are pro-drugs that are triphosphorylated by host cellular enzymes. The triphosphorylated NRTIs then compete with natural deoxynucleoside triphosphates (dNTPs) for incorporation into the newly synthesized DNA chains where they cause chain termination. There are two biochemical mechanisms of NRTI drug resistance. The first mechanism is mediated by mutations that allow the RT enzyme to discriminate against NRTI during synthesis, thereby preventing their addition to the primer DNA chain (Huang *et al.*,1998). The second mechanism is mediated by mutations in RT that increase the rate of hydrolytic removal of the chain terminating NRTI and enable continued DNA synthesis (Arion *et al.*, 1998; Arion *et al.*, 2000).

Although most of the mutations associated with NRTI resistance are not at the active site of the enzyme, they do lead to conformational changes that affect the active site aspartate residues (Ren *et al.*, 1998). Different mutations lead to 2 different mechanisms for resistance: decreased substrate binding and increased phosphorolysis (removal of the chain-terminating substrate that has already been incorporated into the growing proviral DNA chain). Both mechanisms lead to an overall net decrease in termination of the elongating chain of HIV DNA by NRTI (Arion *et al.*, 1998; Arion *et al.*, 2000). Three patterns of multi-NRTI resistance mutations have now been identified. The first being the Q151M complex which confers reduced susceptibility to all nucleoside reverse transcriptase inhibitors (Shirasaka *et al.*, 1995; Kavlick *et al.*, 1998; Mbisa *et al.*, 2011).

The second is the 69 insertion complex, consisting of a mutation at codon 69 (typically T69S) followed by an insertion of amino acids (S-S, S-A, S-G, or others) (Larder *et al.*, 1999). The 69 insertion is often accompanied by mutations at other sites. Some other amino acid changes from the wild-type Threonine (T) at codon 69 without the insertion may also be associated with broad NRTI resistance (Winters *et al.*, 1998).

The third pattern of multi-NRTI resistance involves NRTI-associated mutations (NAMs). These include the reverse-transcriptase mutations M41L, D67N, K70R, L210W, T215Y/F, and K219Q/E, which were initially recognized after zidovudine therapy (Boucher *et al.*, 1992). Although NAMs are selected for only by thymidine NRTIs (zidovudine and Tenofovir), they are associated to varying degrees of reduced susceptibility to all NRTIs. The NAMs cause resistance by improving excision of the chain terminator by phosphorolysis (Lennerstrand *et al.*, 2001) rather than the common mechanism for other reverse-transcriptase and protease mutations, which is by decreasing binding of the inhibitor to the target. Other mutations, such as the 69 insertion and the RT K65R mutation, also appear to cause resistance by the excision mechanism. (Lennerstrand *et al.*, 2001).

1.13.2 NON-NUCLEOTIDE RT INHIBITORS MECHANISM OF ACTION AND DRUG RESISTANT MUTATIONS

The non-nucleoside RT inhibitors bind to a hydrophobic pocket in RT close to the active site. These compounds inhibit HIV-1 replication allosterically by displacing the catalytic aspartate residues relative to the polymerase binding site (Esnouf *et al.*, 1995). The mutations responsible for NNRTI resistance are in the hydrophobic pocket which binds the inhibitors. A single mutation in this pocket may result in high level resistance to one or more NNRTIs. Resistance usually emerges rapidly when NNRTIs are administered as monotherapy or in the presence of incomplete virus suppression suggesting that resistance is caused by the selection of a pre-existing population of mutant viruses within an individual (Jackson *et al.*, 2000).

1.13.3 PROTEASE INHIBITOR MECHANISM OF ACTION AND DRUG RESISTANT MUTATIONS

Protease inhibitors bind to the active site of the protease and block the cleavage of viral poly-proteins by protease into structural and enzymatic components which then leads to the production of non–infectious viral particles. Mutations associated with PI resistance are found at more than 20 different residues of the enzyme. Resistance is mediated by structural changes that reduce binding affinity between the inhibitor and the mutant protease molecule. The effects of non-active site mutations are less obvious and appear to involve other mechanisms: alterations in enzyme catalysis, effects on dimer stability, alterations in inhibitor binding kinetics, or active site re-shaping through long-range structural perturbations (Miller, 2001).

1.13.4 FUSION AND ENTRY INHIBITORS (FI) MECHANISM OF ACTION AND DRUG RESISTANT MUTATIONS

HIV attaches specifically to $CD4^+$ cells on the host cell membrane through glycoprotein (gp) 120. When attachment to the host cell occurs, gp41 embeds itself in the host cell membrane. The gp41 is comprised of two adjoining subunits, the Heptad Repeat 1 and the Heptad Repeat 2. The embedding of the gp41 involves the HR1 subunit of gp41 sliding over the HR2 subunit to draw the HIV and host cell membranes closer together. The gp41 fusion peptide now undergoes a further conformational change that brings the HIV and host cell membranes in contact with one another. Fusion "pores" are formed that facilitate entry of the HIV nucleo-capsid (protein capsid *1* in the HIV genome) into the host cell (Zdanowicz, 2006).

Fusion inhibitors act on the transmembrane protein gp41 and prevent the fusion of the viral and cell membranes of susceptible cells. These inhibitors prevent HIV replication by their interference in the early stage of the HIV life cycle because fusion of HIV with the host cell membrane is an essential step in viral entry into the cell (Zdanowicz, 2006). Mutations associated with fusion inhibitors are in the HR1 region of the gp41 envelope gene and they prevent the activity of the fusion inhibitors. Fusion inhibitor Enfuvirtide is a synthetic peptide that binds directly to the HIV gp41 and prevents it from undergoing the conformational change that leads fusion of the HIV and host cell membrane. A single amino acid substitution in the gp41 can reduce the efficiency of Enfuvirtide. On the other hand, maraviroc (an Entry inhibitor) mutations are usually clustered in the C2-V3 region of the gp21 enzyme (Westby et al., 2007).

1.13.5 INTEGRASE INHIBITORS MECHANISM OF ACTION AND DRUG RESISTANT MUTATIONS

Integrase inhibitors block the action of integrase hence preventing the integration of synthesized viral DNA into the host cell DNA. Raltegravir is the FDA approved integrase inhibitor and even though very little is known about resistance development regarding this drug class, primary data shows two mutational pathways associated with mutations on the integrase gene positions 148 or 155 (Gahn and Sued, 2007). A summary of the antiretrovirals targets in the HIV-1 cycle is represented in Figure 1.11.

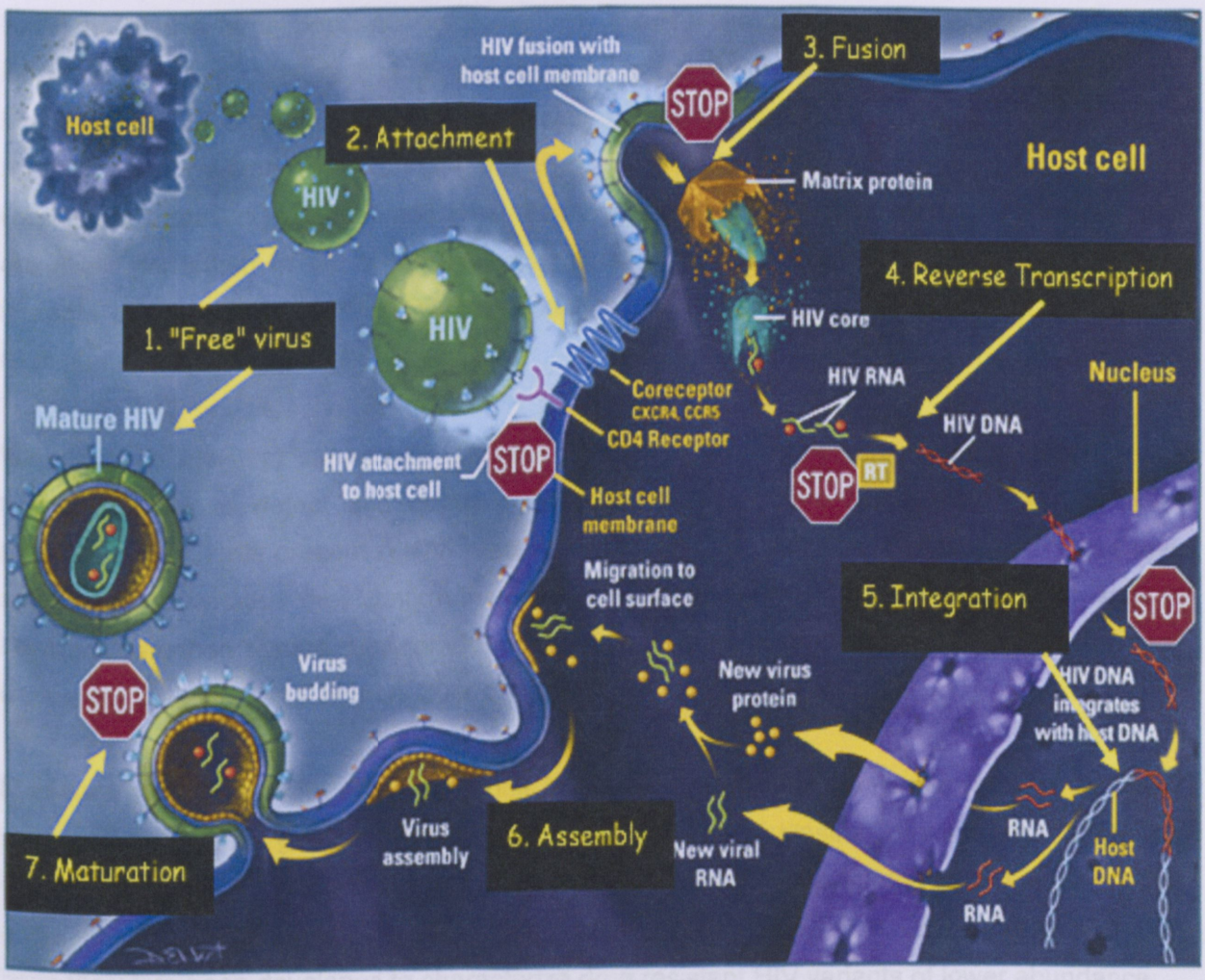


Figure 1.8: The different inhibitors at their various sites of action. The fusion and the entry inhibitors inhibit the process of fusion and entry of the viral particle. Reverse transcriptase inhibitors inhibit the process of reverse transcription. The integrase inhibitors inhibit the integration process of the viral particle into host genome and the protease inhibitors inhibiting protein synthesis of the virus and its assembly. (www.HIVwebstudy.org) (Accessed, 2nd April 2013).

1.14 MINORITY VARIANTS OF DRUG RESISTANT HIV

HIV-1 harbors a large population of genetically diverse viral variants. The genetic diversity of HIV-1 results from rapid, high level virus turnover of approximately 10^{11} virions per day and 10^{8} infected cells per day and from nucleotide mis-incorporation during replication of HIV-1 genome by error prone reverse transcriptase (Jordan *et al.*, 2010). With this combination of high rates of replication, mutation, and recombination, every possible mutation and many double mutations are likely generated in each untreated individual on a daily basis. Therefore, drug-resistant mutants are thus present in all infected patients before the initiation of therapy (Gianella and Richman, 2010).

Minority drug resistant variants of HIV are driven by the development of more sensitive and precise assays to detect and quantify minority variants in the large, genetically complex population of variants present in HIV-infected individuals. The practical implications of these new assays are not whether they confirm the preexistence of drug resistant variants. The question is however, whether results generated by these assays can inform treatment strategies and decisions (Gianella and Richman, 2010).

Mutations associated with drug resistance in clinical practice are generally detected by direct sequencing of the pol gene from the population of HIV RNA in plasma (Hirsch *et al.,* 2008). The results of standard resistance assays are limited by their inability to detect minority viral populations at levels below 20%–30% (Leitner *et al.*, 1999: Halvas *et al.*, 2006). Several assays have therefore been developed to characterize drug resistant HIV variants of lower abundance.

1.14.1 METHODS TO DETECT MINORITY VARIANTS

1.14.1.1 Point mutation assays

Point mutation assays such as mutangenically-separation PCR (MS-PCR) and line probe assay (LiPA) (Servais *et al.*, 2001) are generally highly sensitive and specific for detecting a selected minority drug resistance mutation. Although some of them depend entirely on differential hybridization to alternative viral variants, additional ligation step and/or polymerase chain reaction amplification can improve their specificity and sensitivity. These assays are relatively inexpensive and moderately labor-intensive (Gianella and Richman, 2010). The results are

generally easy to interpret. This methodology is especially suitable for epidemiological studies in which the dynamics and persistence of the most common resistance mutations are analyzed, for example, following the use of nevirapine for the prevention of mother-to-child transmission.

These assays cannot completely address the number and complexity of drug resistant mutations, for example, following treatment failure. The major limitations of point mutation assays are their ability to detect only a single point mutation at a time and the reduced ability to detect alternative polymorphisms in the codon of interest or relevant mutations in nearby codons. Moreover, point mutation assays may be prone to false-positive results at very low percent minority readouts; therefore, a careful determination of a biologically relevant lower level of detection is important, and these are likely to be different for different resistance mutations (and may be as high as 2%) (Johnson *et al.*, 2008).

1.14.1.2 Sequencing assays

In contrast to point mutation assays, sequencing methods permit analyses of the entire sequence and with sequencing of molecular clones, the genetic linkage of each detected mutation. They are less susceptible to primer and probe polymorphisms and do not need individual cutoff assessment for each mutation to avoid false-positive results.

For the detection of minority variants, sequencing techniques tend to be more labor-intensive and costly than point mutation assays. Sequencing assays such as Single Genome Sequencing or UltraDeep Pyrosequencing represent the gold standard by which point mutation assays of minority variants in clinical specimens must be validated.

1.14.1.3 Phenotypic assays

Phenotypic assays assess drug susceptibility by determining the effect of different inhibitors on the replication of viral isolates or recombinant vectors carrying patient-derived viral domains. In routine practice, their ability to detect minor drug-resistant variants is limited (Petrophoulos *et al.*, 2000). Highly sensitive phenotyping assays were developed to detect relevant, low-frequency drug resistant virus variants in clinical specimens (Nissley *et al.*, 2005; Louvel *et al.*, 2008). These methods provide phenotypic selection of drug-resistant variants, which can be further

characterized by DNA sequencing, leading to the opportunity to discover new resistance mutations, however, they do not provide quantification of minority variants.

Chemokine receptor utilization provides a unique situation in which the detection of any variants utilizing X4 can predict treatment failure with CCR5 inhibitors such as maraviroc (Su *et al.*, 2009). A phenotypic assay that can detect the presence of X4-utilizing variants down to 0.3% of the population with 100% sensitivity has proven as sensitive and specific as any approach for the prediction of the utility of this class of drugs (Whitcomb *et al.*, 2007; Su *et al.*, 2009).

1.14.2 COMMON LIMITATIONS OF MINORITY VARIANT ASSAYS

First, the specimen assayed must have a population size sufficient to permit the representative presence of minority variants. An amplification assay that uses the nucleic acid extract derived from 10 mL of plasma with 5000 copies/mL will not provide informative data about variants that represent a small percentage of the viral population (Gianella and Richman, 2010). Second, polymorphisms selectively associated with drug resistance may skew the sensitivity of primers and probes. For example, the T215Y reverse transcriptase mutation conferring zidovudine resistance is often associated with variants at codons 210 and 214.

Third, primers and probes designed to amplify commonly utilized laboratory strains may not perform as well with patient derived variants and may perform even worse with the variety of different subtypes that circulate globally. Finally, population- sequencing technologies that do not interrogate clones, as do single-genome sequencing and ultradeep pyrosequencing, cannot determine whether mutations are present on different variants in the population or are colinear within a given sequence (Gianella and Richman, 2010).

1.15 POLYMERASE CHAIN REACTION AND MOLECULAR CLONING AS TOOLS IN HIV RESEARCH

1.15.1 Polymerase chain reaction (PCR)

The polymerase chain reaction is a technique that has rapidly become one of the most widely used techniques in molecular biology because it is quick, inexpensive and simple. The technique amplifies specific DNA fragments from minute quantities of source DNA material, even when that source DNA is of relatively poor quality (Erlich and Arnhein, 1992).

1.15.1.1 Principle of PCR

PCR involves three main steps (a) Denaturation: DNA fragments are heated at high temperatures which separate the DNA double stranded helix into single strands. These strands are then accessible to primers. (b) Annealing: The reaction mixture is cooled down. The primers anneal to the complementary regions in the DNA template strand of the original sequence and new synthesized sequences and double strands are formed again between the primers and the complementary sequences (Erlich and Arnhein, 1992).

DNA polymerase known as Taq polymerase isolated from the hot-spring bacterium *Thermus aquaticus* helps extend the primer sequences when it binds to the separated DNA helix. The enzyme can withstand the high temperatures needed for DNA strand separation. (c) Extension: The DNA polymerase synthesizes a complementary strand. The enzyme reads the opposing strand sequence and extends the primers by adding nucleotides in the order in which they can pair. (Saiki *et al*,. 1988). A presentation of the PCR process is shown in figure 1.12.

PCR has conditions in which its results are better achieved. These conditions are different with each and every step involved. The first step is the complete denaturation of the DNA template at the start of the PCR reaction and it is essential. This step is called the initial denaturation. Incomplete denaturation of DNA will result in the inefficient use of the template in the first amplification cycle and, consequently, poor yield of PCR product (Saiki *et al*,. 1988).The annealing temperature may be estimated as 5°C lower than the melting temperature of the primer-template DNA duplex. If non-specific PCR products are obtained in addition to the expected product, the annealing temperature can be optimized by increasing it stepwise by 1-

2°C. Usually, the extension step is performed at 72°C and a 1-min extension is sufficient to synthesize PCR fragments as long as 2 kb (Saiki *et al*,. 1988).

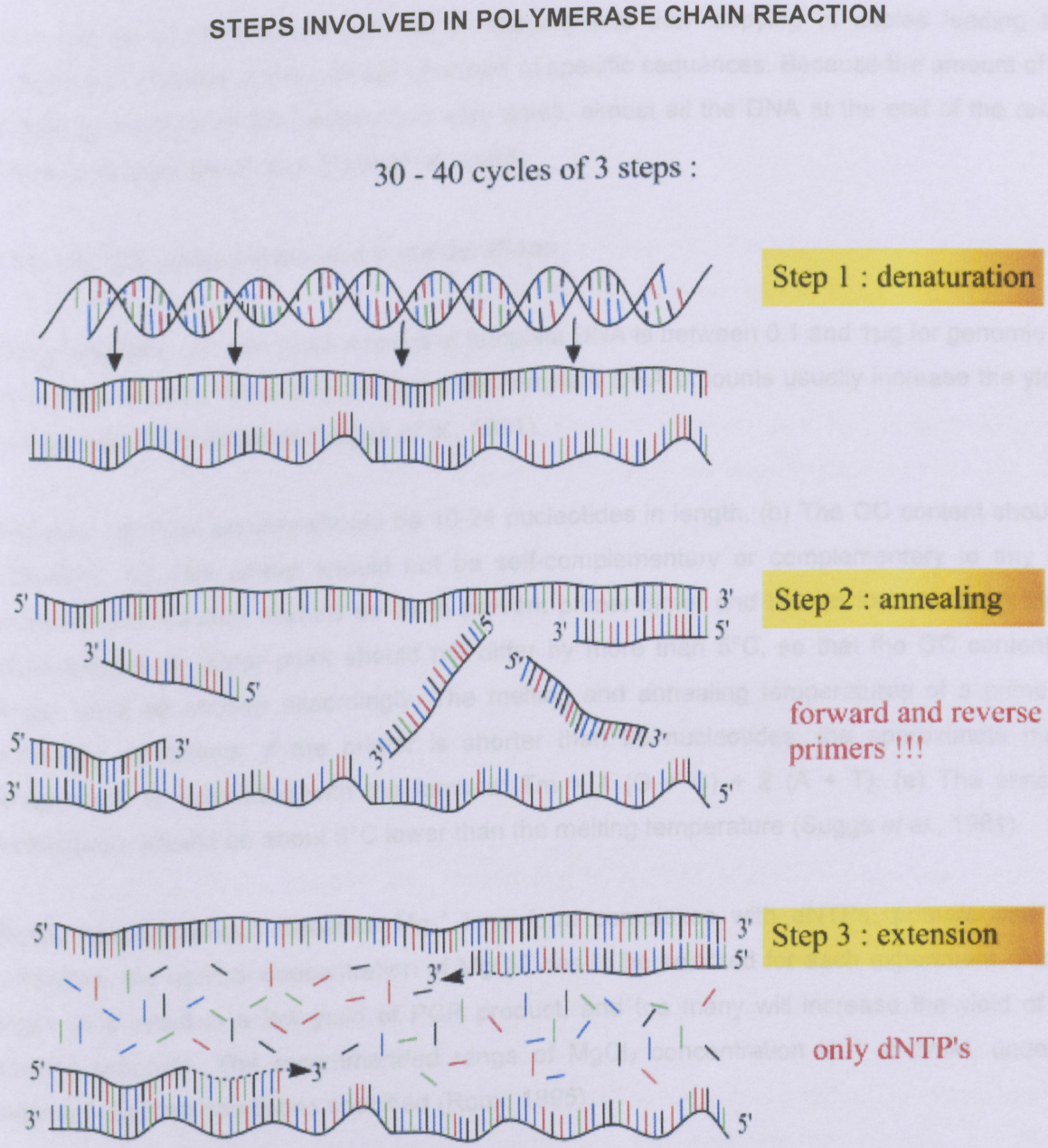


Figure 1.9: A schematic representation of the three main steps involved in PCR amplification. By Andy Vierstraete, 1999 (Accessed, 6th June 2013).

When larger DNA fragments are amplified, the amplification time is usually extended by 1 min per 1000bp. The number of PCR cycles will basically depend on the expected yield of the PCR product. After the last cycle, samples are usually incubated at 68-72°C for 5-10 min to fill-in the protruding ends of the newly synthesized PCR products (Saiki *et al*,. 1988).

This cycling of temperatures results in copying and then copying of copies leading to an exponential increase in the number of copies of specific sequences. Because the amount of DNA placed in the tube at the beginning is very small, almost all the DNA at the end of the reaction cycles is copied sequences (Saiki *et al*,. 1988).

1.15.1.2 PCR components and considerations

Template DNA: An adequate amount of template DNA is between 0.1 and 1µg for genomic DNA for a total reaction mixture of 100µl. Larger template DNA amounts usually increase the yield of non-specific PCR products (Suggs *et al*., 1981).

Primers: (a) PCR primers should be 10-24 nucleotides in length. (b) The GC content should be 40%-60%. (c) The primer should not be self-complementary or complementary to any other primer in the reaction mixture so as to prevent primer-dimer and hairpin formation. (d) Melting temperatures of primer pairs should not differ by more than 5°C, so that the GC content and length must be chosen accordingly. The melting and annealing temperatures of a primer are estimated as follows: if the primer is shorter than 25 nucleotides, the approximate melting temperature is calculated with the formula: Tm = 4 (G + C) + 2 (A + T). (e) The annealing temperature should be about 5°C lower than the melting temperature (Suggs *et al*., 1981).

$MgCl_2$ concentration: Because Mg_2^+ ions form complexes with dNTPs, primers and DNA templates, the optimal concentration of MgCl2 has to be selected for each experiment. Too few Mg2+ ions result in a low yield of PCR product, and too many will increase the yield of non-specific products. The recommended range of $MgCl_2$ concentration is 1 to 3mM, under the standard reaction conditions specified (Roux, 1995).

Taq DNA polymerase: Higher Taq DNA polymerase concentrations than needed may cause synthesis of non-specific products, at least 0.025units/µl of taq should be sufficient for a 20µl PCR reaction (Saiki *et al*,. 1988).

dNTPs: The concentration of each dNTP (dATP, dCTP, dGTP, dTTP) in the reaction mixture is usually 200μM. These concentrations must be checked as being equal, because inaccuracies will increase the degree of mis-incorporation.

Water and 10x PCR buffer are other components of PCR. The buffer promotes the taq usage. If the pH of the reaction drops by more than one full unit the reaction becomes more acidic and the enzyme fails to work and the water is used to get the reaction volume to the right amount so that the concentration of the reactants will be correct (Innis *et al.*, 1988).

1.15.1.3 Gel electrophoresis

The reaction products are resolved by gel electrophoresis. Gel electrophoresis is a technique for separating the components of a mixture of molecules (proteins, DNA or RNA) by size as a result of an electric field within a support gel. Depending on the quantity produced and the size of the amplified fragment, the reaction products can be visualized directly by staining with ethidium bromide or a silver-staining protocol, or by means of radioisotopes and autoradiography (A technique where radioactively labeled molecules are visualized through exposure to X-ray film).

1.15.1.4 .TYPES OF PCR

1.15.1.3.1 Reverse Transcription PCR (RT-PCR)

This is no different from the ordinary PCR; it is used for amplifying DNA from RNA. Reverse transcriptase enzyme reverse transcribes RNA into cDNA, which is then amplified by PCR.

1.15.1.3.2 Nested PCR

Nested sets of primers can be used to improve PCR yield and specificity of the target DNA sequence or gene region (Newton and Graham, 1994). PCR with nested primers is performed for 15 - 30 cycles with one primer set and then for an additional 15 - 30 cycles, with a second primer set, for an internal region of the first amplified DNA product. Thus, the larger fragment produced by the first round of PCR is used as the template for the second PCR. Nested PCR dramatically increases the sensitivity and specificity of DNA amplification. The specificity is mainly improved because this technique almost always eliminates any spurious non-specific amplification products (Newton and Graham, 1994). This is because after the first round of PCR any non-specific products are unlikely to be sufficiently complementary to the nested primers to be able to serve as a template for further amplification, thus the desired target sequence is preferentially

amplified. However, the increased risk of contamination is a drawback of this extreme sensitivity, and great care must be taken when performing such PCRs (Newton and Graham, 1994).

1.15.1.3.3 Multiplex PCR

Standard PCR usually uses one pair of primers to amplify a specific sequence or gene region of interest. Multiplex PCR uses multiple pairs of primers to amplify several different sequences simultaneously. For this type of PCR amplification, primers are chosen with similar annealing temperatures (Atlas and Bey, 1994).

The lengths of amplified products should be similar; large differences in the lengths of the target DNAs will favor the amplification of the shorter target over the longer one, resulting in differential yields of amplified products (Atlas and Bey, 1994). The presence of many PCR primers in a single tube could cause problems, such as the increased formation of mis-primed PCR products, "primer dimers", and the amplification discrimination of longer DNA fragments. However, multiplex PCR buffers contain Taq polymerase additive, which reduces the competition among amplicons and the discrimination of longer DNA fragments during multiplex PCR (Atlas and Bey, 1994).

1.15.1.3.4 Quantitative PCR

Quantitative PCR (qPCR) is used to measure the quantity of a PCR product (commonly in real-time). It quantitatively measures starting amounts of DNA, cDNA, or RNA. Quantitative PCR is commonly used to determine whether a DNA sequence is present in a sample and the number of its copies in the sample. This PCR has a very high degree of precision by the use of fluorescent dyes, such as Sybr Green, EvaGreen or fluorophore-containing DNA probes, such as TaqMan, to measure the amount of amplified product in real time. It is also sometimes abbreviated to RT-PCR (Real Time PCR) or QRT-PCR (Quantitative Real Time PCR) or RTQ-PCR (Real Time Quantitative PCR).

1.15.1.5 Avoiding PCR contamination

To avoid false positive PCR results that might be due to contamination or carryover of amplified DNA products a UV sterilizing PCR working station should be used. This provides a process for eliminating contamination by utilizing the microbicidal properties of shortwave 254nm UV light. The UV light should be turned on before and after PCR experiment for 30 minutes and also the working area should be decontaminated with 70% ethanol in all experiments (Niederhauser *et al*, 1994). Gloves and laboratory coats, strictly for PCR experiments should be used. Gloves are to be changed between the procedures.

The master-mix must be done in a clean room where specimens are not allowed. Nucleic acid extractions and PCR reactions are to be done in a separate room inside air cabinets. In extraction and PCR procedures, separate sets of filtered tips and pipetting devices should be used. RNase and DNase free tubes should be used. In addition both positive and negative controls must be used in each procedure (Niederhauser *et al*, 1994).

1.15.2 MOLECULAR CLONING

Cloning is a way to isolate and produce large quantities of specific genes without the need to cultivate the original organisms. It can be used to study organisms which are difficult to cultivate and growth requirements are unknown. This is achieved by transferring the desired gene to another organism which is easy to cultivate and has ideal characteristics such as, rapid growth, capable of growing in an inexpensive culture medium, not harmful or pathogenic, capable of taking up DNA, and stable in culture. The most useful hosts for cloning are microorganisms that grow well and about which a lot of genetic information is available, such as bacteria *Escherichia coli* and *Bacillus subtilis* and the yeast *Saccharomyces cerevisiae* (Kaksonen, 2005).

The basic strategy of cloning is to move the desired gene from a large, complex genome to a small, simple one. Various cloning strategies, some of which utilize PCR, have been used to allow the analysis of the nucleic acid fragments retrieved directly from the environment. This is the simplest way to obtain phylotypes from the environmental samples. The resulting idea of the community diversity will depend upon the specificity of the PCR primers and the efficiency with which the ribosomal ribonucleic acid (rRNA) genes are amplified (Kaksonen, 2005).

In so called shotgun cloning, community DNA is fragmented with restriction enzymes, the fragments are cloned, and the resulting clones are screened for the presence of rRNA genes. This is a difficult procedure, as rRNA genes will only be in a small fraction of the total clones. However this approach probably provides the most unbiased estimate of community diversity (Kaksonen, 2005).

A third alternative for obtaining rRNA gene clones is to clone complementary DNA (cDNA) obtained by reverse transcription (with or without PCR amplification). The resulting community profile will offer some reflection of the most metabolically active organisms, because cells that produce more RNA (i.e. those that are metabolically more active) will be represented to a greater extent than metabolically inactive cells (Kaksonen, 2005). Figure 1.13 shows a summary of different way to achieve cloning.

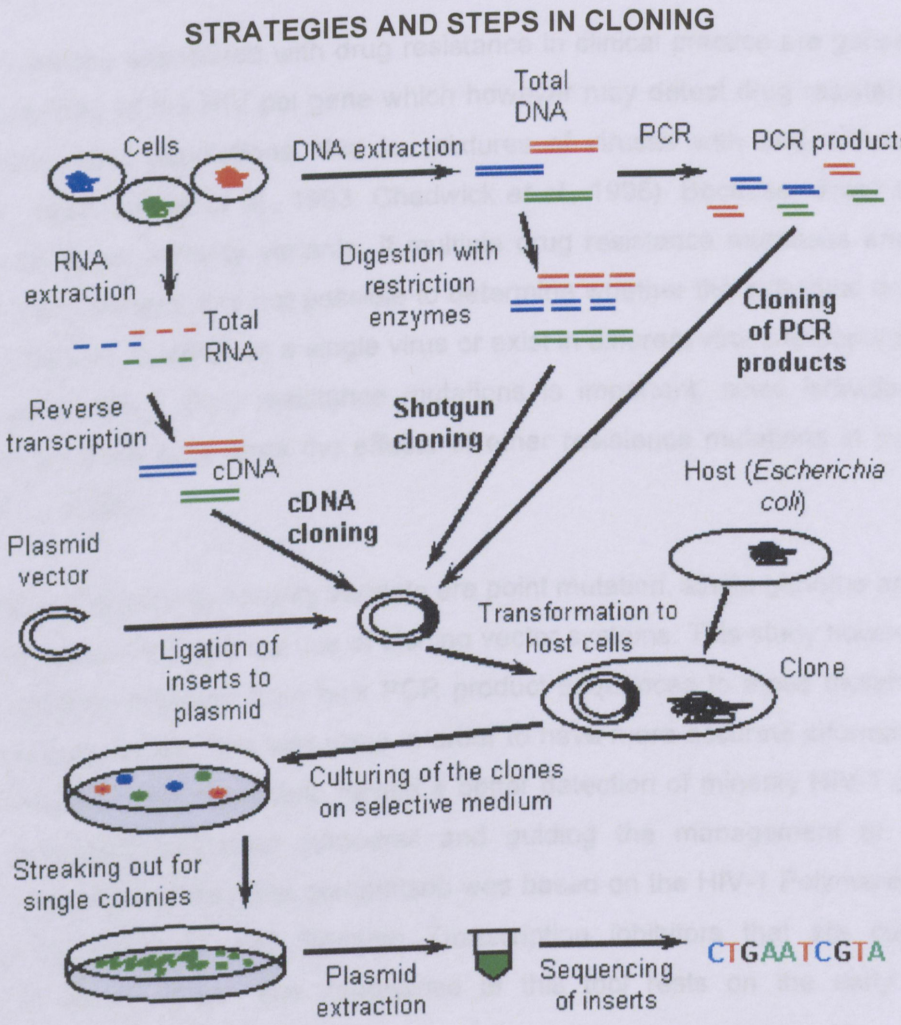


Figure 1.13: shows the different ways in which molecular cloning can be achieved. By Anna Kaksonen (Accessed, 23rd June 2013).

1.16 RATIONALE OF THE STUDY

Treatment for HIV in South Africa involves the use of first line therapy with Tenofovir, Lamivudine and Efavirenz or Nevirapine for adults and Abacavir, Lamivudine and Liponavir/Ritonavir for children. Second line therapy is usually a combination of two Reverse Transcriptase inhibitors and Protease inhibitors such as Liponavir or Ritonavir.

Generally, minority drug resistant strains exist in every individual infected with HIV, that is, individuals who are treatment-experienced and those that are treatment-naïve. Because these minority strains are usually present at very low levels, they cannot be adequately detected and quantified using direct sequencing of PCR products (Gianella and Richman, 2010) and this may the limit diagnostic of resistance testing in HIV-1 infected individuals.

In addition, mutations associated with drug resistance in clinical practice are generally detected by direct sequencing of the HIV pol gene which however may detect drug resistance mutations present in mixed viral populations, that is, mixtures of viruses with and without a mutation (Leitner *et al.*, 1993: Larder *et al.*, 1993: Chadwick *et al.*, 1996). Because, direct sequencing is not designed to detect minority variants, if multiple drug resistance mutations are detected as mixtures in a given sample, it is not possible to determine whether the individual drug resistance mutations are present together on a single virus or exist in different viral sub-populations. And so genetic linkage of HIV-1 drug resistance mutations is important, since individual resistance mutations can enhance or reverse the effects of other resistance mutations in the same virus (Eshleman *et al.*, 2003).

Current methods of detecting minority variants are point mutation, single genome amplification or single genome sequencing and the use of cloning vector systems. This study however compares the mutation patterns detected from bulk PCR product sequences to those mutations detected from viral clone sequences. This was done in order to have more accurate information on HIV-1 drug resistant strains. This could help having a better detection of minority HIV-1 drug resistant populations, predicting treatment outcomes and guiding the management of HIV infected individuals in clinical practices. This comparison was based on the HIV-1 Polymerase (pol) gene in relation to the Protease and Reverse Transcription inhibitors that are currently being administered in South Africa. The importance of this tool rests on the early detection of resistance even before virologic failure.

1.17 OBJECTIVES OF THE STUDY

1.17.1 General objective

The general objective was to design and evaluate a single genome sequencing protocol for the detection of HIV-1 drug resistance mutations in the protease and reverse transcriptase genes.

1.17.2 Specific objectives

The specific objectives were:

1. To set up a protocol and amplify protease and reverse transcriptase genes from drug naïve and drug experienced patients.

2. To compare the frequency and type of resistant mutations detected in PCR products and cloned PCR products.

CHAPTER 2: MATERIALS AND METHODS

2.0 ETHICAL CLEARANCE

This research is a part of an ongoing project (Project No: SMNS/12/MBY/06) on HIV-1 drug resistance in Limpopo province sponsored by the Department of Health, for which ethical clearance has been obtained. Permission to collect samples from government health establishments was obtained from the Department of Health in Polokwane and permission to collect samples from the University of Venda HIV/AIDS unit was obtained from the University of Venda ethical clearance committee. A signed informed consent was obtained from the study participants before sample collection.

2.1 STUDY POPULATION AND BIOSPECIMENS

Archived plasma samples and DNA extracted from Peripheral Blood Mononuclear Cells of patients recruited from the Bela-Bela HIV/AIDS Prevention Group in 2011 and 2012; and blood samples obtained from the University of Venda HIV/AIDS unit were used in this study. The total number of archived samples was 86 for the Bela-Bela HIV/AIDS and Prevention group and 10 for the University of Venda HIV/AIDS unit.

2.2 TOTAL RNA ISOLATION

Total RNA was isolated from plasma using an in-house extraction protocol. The extraction was done by pre-spinning 5ml of plasma at 5300 rpm for 10minutes at 4°C to remove cellular debris. The supernatant was dispensed into a fresh, sterile, RNase/DNase-free 1.5mL eppendorf tubes then spun at 16000 rpm at 4°C for an hour. After that, the supernatant was removed and discarded, and the pellet was resuspended in 50µl of RNase/DNase-free 5mM Tris-HCl pH 8.0 (Sigma-Aldrich (Pty) Ltd, Sydney, Australia) after which, 10µl of Proteinase K (Sigma-Aldrich (Pty) Ltd, Sydney, Australia) was added and the mixture was left to incubate at 55°C for 30 minutes. Upon incubation, 6M of Guanidinium isothyocyanate (Sigma-Aldrich (Pty) Ltd, Sydney, Australia) and 20mg/ml of glycogen (Invitrogen, California, USA) were added and then left to incubate at room temperature for 5 minutes. When the incubation was done, 270µl of 100% Isopropanol (Sigma-Aldrich (Pty) Ltd, Sydney, Australia) was added and spun at room temperature for 20 minutes at 21000 rpm. The supernatant was removed, and the pellet was washed with 500µl of 70% ethanol (Sigma-Aldrich (Pty) Ltd, Sydney, Australia) by spinning for 5

minutes at 21000rpm at 24°C. Finally the supernatant was discarded and the pellet was left to air dry for 2-3 minutes and then resuspended in 40µl 5mM RNase-free Tris-HCl pH 8.0.

2.3 DNA EXTRACTION

DNA extraction from PBMCs was done using the QIAquick DNA extraction kit (Qiagen, Hilden, Germany) according to manufacturer's instruction.

2.4 COMPLEMENTARY DNA (cDNA) SYNTHESIS

Complementary DNA was prepared for use in a subsequent amplification step. Briefly, 5µl of 10mM dNTPs (Invitrogen, California, USA) and 2µl of 2mM oligo dT (Invitrogen, California, USA) were added to 8µl of RNA then denatured at 65°C for 10 minutes. The tube was chilled on ice and then 8µl of a cDNA cocktail was added. The cDNA cocktail was made up of 2µl of 5x FSB buffer, 0.5µl of 0.1 M DTT, 0.5µl of RNase free water, 0.5µl of 40U/µl RNase-out and 0.5µl of 200U/µl superscript III (Invitrogen, California, USA). The mixture was incubated at 50°C for an hour, 55°C for another hour for further elongation, 70°C for 15minutes to inactivate the enzyme and then 4°C to hold. After this, 2µl of 2U/µl RNaseH (Invitrogen California, USA) was added and incubated at 37°C for 20 minutes and the cDNA was stored at -80°C.

2.5 AMPLIFICATION OF HIV-1 POL REGION

The amplification was carried out using the established protocol as described by Bessong et al., 2006. The protocol is set out to generate approximately 1.4kb product (HXB2 location 2253-3713) comprising the complete Protease gene (PR) and ~ 900bp of Reverse Transcriptase (RT) gene. The template used was obtained from cDNA prepared from RNA or DNA extracted from PBMCs. The amplification of PCR products was carried out in 0.2ml thin walled tubes (Axygen, Inc. California USA) in an automatic thermocycler (G-Storm GS1). First round PCR amplification assay was done by using two sets of forward and reverse primers as described in table 3.1. Briefly, 5µl of template was used for amplification in a 45µl master mix containing the following: 5µl of 10X reaction buffer + 17.5mM $Mgcl_2$ (Roche Diagnostics, Mannehiem, Germany), 1µl of 10mM dNTPs (Invitrogen, California, USA) 2µl of 10µM of sense and anti-sense primers (Inqaba Biotechnical Industries, Pretoria, South Africa) for first round, 0.5µl of 5U/µl expand long template

enzyme (Roche Diagnostics, Mannehiem, Germany) and 34.5µl of RNase/DNase Free sterile water (QIAgen, Hilden, Germany).

The amplification process included the following steps: Initial denaturation of DNA at 95°C for 3 minutes, followed by 30 cycles of denaturation at 95°C for 2 minutes, annealing at 58°C for 1 minute, extension at 68°C for 2 minutes and final extension at 68°C for 10 minutes followed by holding temperature of 4°C until the programme was manually terminated.

The second round PCR was performed using 5µl of the first round product with 10x PCR buffer + 17.5mM of $Mgcl_2$, 2µl of 10mM dNTPs, 4µl of 10mM second round primers described in table 3.1, 1µl of 5U/µl expand long template enzyme (Roche Diagnostics, Mannehiem, Germany) and 74µl of RNase/DNase free sterile water (QIAgen, Hilden, Germany) in a 100µl of reaction mixture. PCR conditiors for nested PCR were the same as those employed for the first round amplification step.

Table 3.1: Primer sequence for first round and nested PCR HIV pol gene, orientation and HXB2 location

Primer	Polarity	Target gene region	HXB2 location	Sequence
Pol 1C	Forward	Outer region	2044 →2069	5'-GAA GGA CAC CAA ATG AAA GAC TGC AC-3'
RT-RV	Reverse	Outer region	3506 ←3536	3'-TAT TTC AGC TAT CAA GTC TTT TGA GTC A- 5'
Gag P1	Forward	Inner region	2112 →2132	5'-CAA GGG GAG GCC AGG GAA TTT- 3'
Pol 2R	Reverse	Inner region	3492 ←3515	3'-TGA TGG GTC ATA ATA TACT CC ATG -5'

2.6 AGAROSE GEL ELECTROPHORESIS

Detection of the amplified products was achieved by running 5µl of the second round amplified product on 1.5% agarose gel electrophoresis stained with ethidium bromide. Briefly, 3µl of a 1kb DNA ladder (Roche Diagnostics, GmbH, Mannhiem, Germany) and 5µl of the amplified product were mixed with 10x DNA loading buffer (Roche Diagnostics, GmbH, Mannhiem, Germany). The gel was set to run at 100V for 40 minutes in a 1x Tris-acetate-EDTA (TAE) buffer (QIAgen, Hilden Germany). Finally the gel was viewed under a transilluminator.

2.7 PURIFICATION OF PCR PRODUCT AND DNA QUANTIFICATION

Amplified PCR products were purified using the QIAquick PCR purification kit (Qiagen, Hilden, Germany) according to manufacturer's instruction with centrifugation steps done at 13 000 rpm in a conventional microcentrifuge. The resulting purified DNA was quantified using a Nanophotometer (Implen, GmbH, Munich, Germany).

2.8 CLONING HIV-1 DNA OF THE PARTIAL POL GENE

A ligation reaction mixture was prepared by combining (in order) the following components: 1µl salt solution (Nacl 50mM: $Mgcl_2$ 25mM), 4µl of PCR product or 4µl of Topo T Control Insert, 1µl Topo TA cloning Vector Mix (Topo cloning vector contains ampicillin and kanamycin resistant genes therefore, insert genes resistant to both antibiotics will be detected in transformed cells) . The reaction was mixed gently by repeated pipetting, then incubated at room temperature for 30 minutes and then placed on ice.

2.8.2 TRANSFORMATION OF CELLS

Topo OneShot competent cells (100µl per tube) were thawed on ice for each ligation reaction. After thawing, 1µl of the ligation reaction was added to the tube containing competent cells then mixed gently by inverting 4-6 times. The transformation reaction was then incubated on ice for 30 minutes. During the incubation period, SOC medium was pre-warmed to 42°C. After which the transformation reaction was heat-shocked at 42°C for 45 seconds, then incubated on ice for 2 minutes. Two hundred and fifty microliters of SOC medium was added to the transformation reaction. Competent cells were left to recover for an hour at 37°C with agitation at 225rpm.

During the out growth period, LB-Kanamycin plates to be used for blue and white screening were prepared, and 40µl of X-gal was spread on each plate. After incubation, 100µl of the transformation reaction was spread on the LB-kanamycin-X-gal plates and incubated inversely at 37°C overnight. For the positive (pUC19) and negative (no insert) controls, 30µl and 10µl of the transformation reactions were spread on the plates respectively. After incubation, 5 white colonies from the ligation+insert plate, one colony from the negative control plate (no insert) and one from the pUC19 positive control plate were selected colonies and used for mini-preps.

2.9 PLASMID PREPARATION AND RESTRICTION DIGESTION

In order to obtain pure DNA from the colonies, plasmid preparation was done using the QIAgen's QIAprep Mini-prep DNA extraction kit (Qiagen, Hilden, Germany) following the manufacturers' instruction. Restriction digestion was done using EcoR1 restriction digestion enzyme. The selection of EcoR1 enzyme (Promega, Wisconsin, USA) was based on the fact that the Topo TA vector has two EcoR1 restriction sites at both ends of the inserted product and the amplified product has no EcoR1 restriction sites. Briefly the digestion was done in a 20µl reaction by the addition of 2µl restriction enzyme buffer, 1µl of plasmid DNA, 0.5µl of restriction enzyme EcoR1 and 16.5µl of sterile distilled water. The mixture was incubated at 37°C for 4 hours then verified with a 1.5% agarose gel electrophoresis.

2.10 SEQUENCING AND EDITING OF AMPLIFIED PRODUCTS AND MINI-PREPS

In order to determine of the sequence of nucleotides in the amplicons and plasmids, sequencing was done. A termination reaction using BigDye terminator was set up and analyzed on the 3500xL sequencer (24Capillary) for the sequencing of purified amplified PCR products and DNA obtained from mini-preps. Nucleotide sequences were edited using the SeqMan II (DNASTAR) software (Version 8.1). The primers used for sequencing of the HIV-1 pol gene are outlined in Table 2.2.

Table 2.2: Primer sequence for sequencing the HIV pol gene, orientation and HXB2 location

Primer	Polarity	Target gene region	HXB2 location	Sequence
Pol 1F	Forward	Outer region	2044 →2069	5'- TTT TCC CAT TAG TCC TAT TGA ACC TGT A -3'
Pol 1R	Reverse	Outer region	3506 ←3536	3'-CAT GCT ACT CTG GAA TAT TGC TGG TGA TCC- 5'
Pol 2F	Forward	Inner region	2112 →2132	5'-CTG CAT TCA CCA TAC CTA GTA TAA AC- 3'
Pol 2R	Reverse	Inner region	3492 ←3515	3'-TGA TGG GTC ATA ATA TAC TCC ATG -5'

2.11 RESISTANCE MUTATIONS ANALYSIS

Drug resistance mutations were determined according to the Stanford drug resistance interpretation program (http://hivdb.stanford.edu). This was done in order to identify known/ unknown mutations present on the HIV pol region of test samples.

2.12 VIRAL SUBTYPING AND MEAN GENETIC DISTANCE ANALYSIS

Viral subtype was determined by phylogenetic analysis. Test and reference nucleotide sequences were aligned with Clustal X (Version 2.0.11). Phylogenetic trees were drawn by the neighbour joining method using Clustal X (Version 2.0.11) and viewed with TreeViewer (Version 1.6.6). The reliability of the trees was assessed by bootstrapping of 1000 replicates. Mean genetic distance among derived sequences was determined by the Kimura-2 parameter method.

2.13 AMINO ACID ALIGNMENT DETERMINATION

All amino acid sequences were obtained using the BioEdit software. Using all test sequences, consensuses were generated from the HIV database (GenBank). Global subtype B and C consensuses were generated from the HIV database programme. The entire consensuses were aligned using the BioEdit software. This helped detect changes in the amino acid sequence of the HIV pol gene.

2.14 STATISTICAL ANAYLSIS

Mutations obtained and identified from sequences of amplified PCR products as well as mutations obtained and identified from purified clones were evaluated for statistical significance by using the student t-test. P values of <0.05 were considered significant.

CHAPTER 3: RESULTS

3.0 DEMOGRAPHICS AND CLINICAL FEATURES OF THE STUDY PARTICIPANT

Of the 96 archived samples obtained from HIV-1 infected individuals attending the Bela-Bela HIV/AIDS Prevention Group and University of Venda HIV/AIDS Unit, 30 samples were selected to be used in this study. However, of the 30 samples selected 12 had very low viral loads hence only 18 samples were analyzed for the presence of mutations on the HIV-1 Pol gene. After sequencing of amplified products and their clones, reliable nucleotide sequences were obtained from 8 samples with complete sequences of 1.4 kb hiv-1 pol gene (protease and partial reverse transcriptase genes). Of the 8 samples (4 males and 4 females) used, only 1 study participant was from the University of Venda HIV/AIDS unit. One of the participants received at least tertiary level education whereas 4 received at least secondary school level education, and 2 received primary school level of education while no educational background data was available for 1 study participant of their educational background.

All the participants were probably infected through unprotected sex and the probable period of infection was between 1997 and 2010. Only one study participant was treatment naïve. Details of the demographics and clinical features of the study participants are shown in Table 3.1.

Table 3.1: Demographic data, treatment, virologic and immunologic features of HIV-1 infected individuals attending the Bela-Bela HIV/AIDS Prevention group and the University of Venda HIV/AIDS unit

ample code	Source of Amplicon	Date of sample collection	Gender	Age	Marital status	Estimated year of infection	Year of treatment initiation	Treatment when samples were collected	CD4 count (cells/mm^3)	Viral load (RNA copies/ml)
V11/01	RNA	07-04-2011	Female	31	Single	1997	2010	AZT+3TC+LPV	N/A	N/A
MARBB14	RNA	09-08-2011	Female	37	Single	1998	2008	TDF-3TC-EFV	165	6164
MARBB68	DNA	09-02-2012	Male	46	Single	2006	2012	Not on treatment	667	311
MARBB69	DNA	09-01-2012	Female	36	Single	2003	2007	AZT-3TC-EFV	231	<50
MARBB73	DNA	09-01-2012	Male	45	Married	N/A	2011	TDF-3TC-EFV	392	<40
MARBB77	DNA	09-01-2012	Male	55	Married	2005	2011	TDF-3TC-EFV	237	32
MARBB83	DNA	09-01-2012	Male	49	Single	2002	2006	AZT-3TC-EFV	N/A	N/
MARBB84	DNA	09-01-2012	Female	29	Single	2005	2005	AZT-3TC-NVP	N/A	N/A

3.1 AMPLIFICATION OF THE HIV POL REGION

Upon amplification of the Pol region, a product size of 1.4kb (encompassing protease and partial reverse transcripase) was obtained from 2 cDNA samples and proviral DNA extracted from PBMCs of 6 samples. To obtain pure DNA, the amplicons were purified. An agarose gel with representative purified PCR samples is presented in figure 3.1.

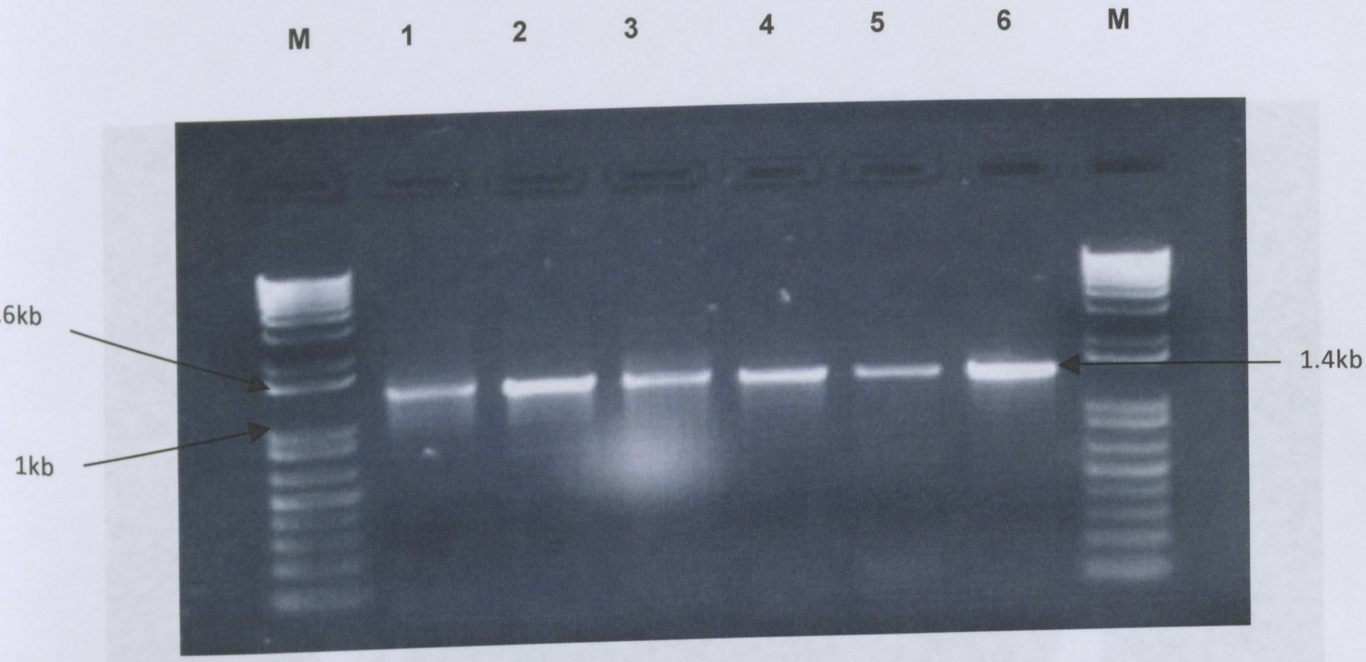


Figure 3.1: A 1.5% agarose gel of amplified products of samples obtained from HIV-1 infected individuals in Bela-Bela HIV/AIDS Prevention Group and the University of Venda HIV/AIDS unit. Lane M: Molecular weight maker (1kb plus). Lanes 1-6: samples. An expected 1.4kb product of partial pol was obtained.

3.2 TRANSFORMATION OF CELLS

After cloning, blue and white colonies were observed. White and light blue colonies are expected to contain DNA inserts of the cloned DNA sample whereas blue colonies are the ones that did not pick up the sample DNA insert. A representation of a LB-kanamycin- 5-bromo-4-chloro-3-indolyl-β-D-galatopyranoside (X-gal) plate is shown in figure 3.2.

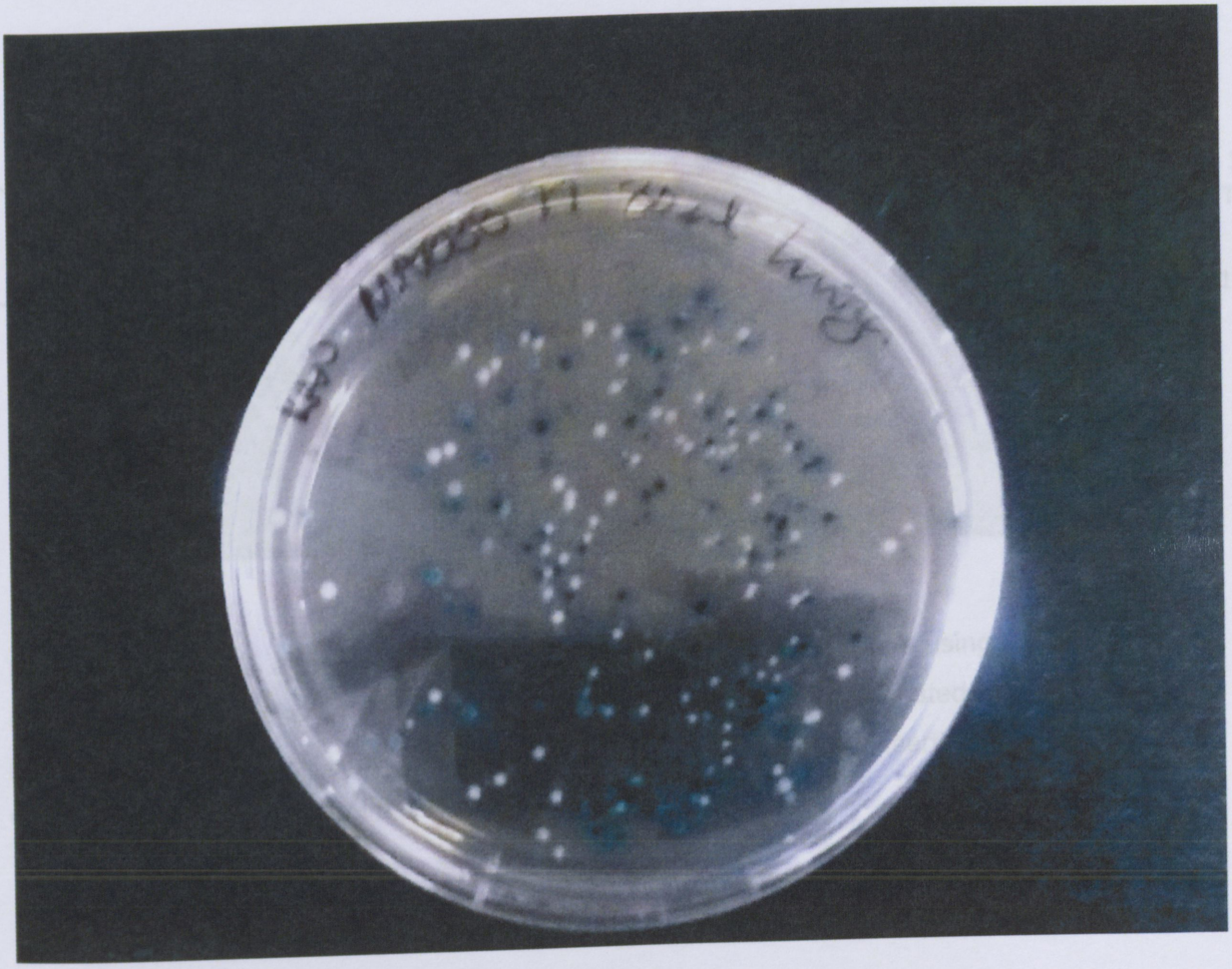


Figure 3.2: A representative LB-kanamycin-X-gal plate of a cloned PCR product into Topo TA cloning vector. Blue colonies represent vector without insert and white colonies represent vector with insert.

3.3 RESTRICTION PROFILES OF MINI PREPS

In order to verify whether cloning was indeed successful, enzyme restriction digestion on mini-preps was done using Eco R1 and two band patterns were observed; 3.9kb Topo TA vector and 1.4kb cloned DNA insert. The control lane with the Topo TA vector also showed two band patterns which were expected because the restriction sites for EcoR1 are in vector system.

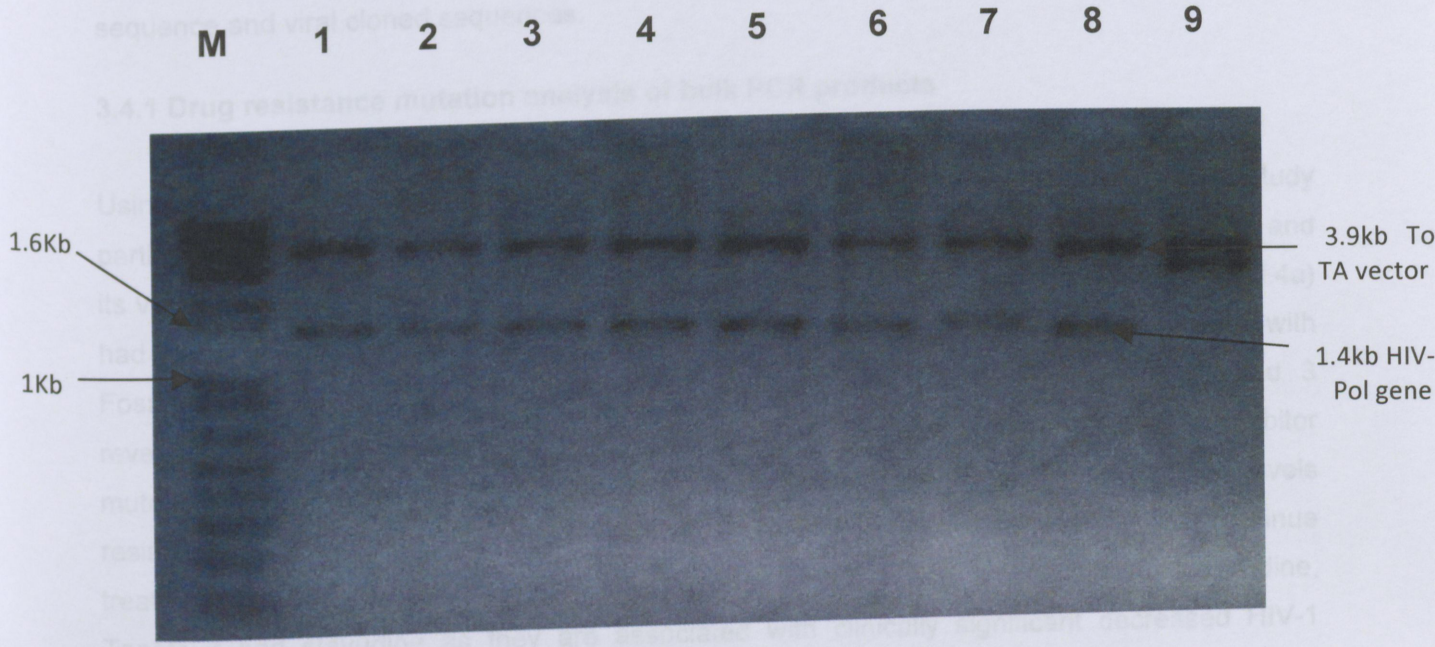


Figure 3.3: A 1.5% agarose gel of single restriction enzyme digestion using Eco R1. Lane M: Molecular weight maker (1kb plus). Lanes 1-8: Representation of digested Mini-preps with two band patterns of 3.9kb (Topo TA vector) and 1.4kb (HIV pol gene insert). Lane 9: control lane with Topo TA vector+control insert.

3.4 SEQUENCING AND DRUG RESISTANCE MUTATION ANALYSIS

Reliable nucleotide sequences were obtained for all 8 samples for the protease and the reverse transcriptase genes. Thirty five and thirty three reliable cloned PCR sequences of protease and reverse transcriptase were obtained respectively. Drug resistance mutation analysis was done on both PCR products and cloned PCR products. No mutations were observed from protease and reverse transcriptase sequences of the drug naïve sample MARBB68 for the PCR product sequence and viral cloned sequences.

3.4.1 Drug resistance mutation analysis of bulk PCR products

Using the Stanford drug resistance interpretation program, sample MARBB68 (a study participant who is drug naïve) did not reveal any mutations from the direct PCR sequence and its viral clones. However, drug resistance analysis revealed that sample MARBB14 (table 3.4a) had 1 selected protease inhibitor minor resistant mutation V11L which is weakly associated with Fosamprenavir/Ritonanvir and Darunavir/Ritonavir. Sample MARBB73 (table 3.4a) had 3 reverse transcriptase mutations; M184V, a transmitted nucleoside reverse transcriptase inhibitor mutation which causes high level resistance to Lamivudine and Emtricitabine and low levels resistance to Didonasine and Abacavir. However, M184V is not a contraindication to continue treatment with Lamivudine or Emtricitabine because they increase susceptibility to Zidovudine, Tenofovir and stavudine as they are associated with clinically significant decreased HIV-1 replication.

MARBB73 also revealed K101E alongside with Y181C both of which are non-nucleoside reverse transcriptase inhibitors resistance mutations. K101E causes intermediate resistance (about 5-10 fold decreased susceptitibilty) to Nervirapine and a low level resistance (about 2-5 fold decreased susceptibility) to Efavirenz, Etravirine and Rilpivirine. Y181C causes high level resistance to Nevirapine, ~2-fold decreased susceptibility to Efavirenz and ~5-fold drecreased susceptibility to Etravirine and Rilpivirine.

Table 3.2a: Protease and Reverse Transcriptase Mutation analysis of PCR products

Sample ID	Protease		Reverse Transcriptase		Sample type
	Major Resistant Mutations	Minor Resistant Mutations	NRTI Resistant Mutations	NNRTI Resistant Mutations	
V11-01	-	-	-	-	Drug experienced
MARBB14	-	V11I	-	-	Drug experienced
MARBB68	-	-	-	-	Drug naïve
MARBB69	-	-	-	-	Drug experienced
MARBB73	-	-	M184V	K101E, Y181C	Drug experienced
MARBB77	-	-	-	-	Drug experienced
MARBB83	-	-	-	-	Drug experienced
MARBB84	-	-	-	-	Drug experienced

3.4.2 Drug resistance mutation analysis of cloned PCR products

Drug resistance analysis for cloned PCR products (table 3.4b) was also done and it revealed that viral clones of sample MARBB14 and 1 viral clone of sample MARBB69 had V11L protease inhibitor minor resistant mutation. Viral clones of sample MARBB73 revealed mutations M184V, K101E, Y181C, Y181V (a non-nucleoside reverse transcriptase inhibitor mutation that causes high level resistance to Nevirapine and has a 10-15-fold decreased susceptibility to Etravirine and Rilpivirine) and T215Y (a nucleoside reverse transcriptase inhibitor mutation that causes Zidovudine and Stavudine resistance and reduced susceptibility to Abacavir, Didanosine and Tenofovir particularly if in combination with M41L and L120W).

Sample MARBB84 viral clones also revealed mutations not observed by direct PCR sequencing such as D30N which is a unique amino acid substitution that is a major protease inhibitor mutation that causes high level resistance to Nelfinavir. Gonzalez *et al,* 2004 found that the mutation D30N has a more significant impact on HIV-1 subtype C than HIV-1 subtype B. Another mutation was G48E which is highly unusual at this position and its effect on protease inhibitors is uncertain (www.hivdb.stanford.edu). K103N, a non-nucleoside reverse transcriptase inhibitor mutation, was revealed in all MARBB84 viral clones. This mutation causes a high level resistance to Nevirapine and Efavirenz and no effect on Etravirine and Rilpivirine (www.hivdb.stanford.edu).

Table 3.2b: Protease and Reverse Transcriptase Mutation analysis of cloned PCR products

Sample ID	Protease		Reveres Transcriptase	
	Major Resistant Mutations	Minor Resistant Mutations	NRTI Resistant Mutations	NNRTI Resistant Mutations
V11-01-1	-	-	-	-
V11-01-3	-	-	-	-
V11-01-5	-	-	-	-
V11-01-6	-	-	-	-
V11-01-7	-	-	-	-
MARBB14-4	-	V11I	-	-
MARBB14-5	-	V11I	-	-
MARBB68-1	-	-	*	*
MARBB68-3	-	-	-	-
MARBB68-4	-	-	-	-
MARBB68-5	-	-	-	-
MARBB68-7	-	-	*	*
MARBB69-1	-	V11I	-	-
MARBB69-3	-	-	-	-
MARBB69-4	-	-	-	-
MARBB69-6	-	-	-	-
MARBB69-7	-	-	-	-
MARBB73-2	-	-	M184V, T215Y	K101E, Y181V
MARBB73-3	-	-	M184V, T215Y	K101E, Y181V
MARBB73-4	-	-	M184V	K101E, Y181C
MARBB73-5	-	-	M184V	K101E, Y181C
MARBB73-7	-	-	M184V, T215Y	K101E, Y181C
MARBB77-1	-	-	-	-
MARBB77-2	-	-	-	-
MARBB77-4	-	-	-	-
MARBB77-6	-	-	-	-
MARBB77-7	-	-	-	-
MARBB83-2	-	-	-	-
MARBB83-4	-	-	-	-
MARBB83-5	-	-	-	-
MARBB84-2	-	-	-	-
MARBB84-3	-	-	-	K103N
MARBB84-4	D30N	G48E	-	K103N
MARBB84-5	-	-	-	K103N
MARBB84-6	-	-	-	K103N

* Indicate that no RT viral clones were obtained

3.4.3 Other mutation detected in the protease gene by bulk PCR sequencing and cloned viral sequences

Nine other protease mutations (table 3.4c) such as K20R (highly polymorphic and commonly selected by Nelfinavir, Ritonavir, Atazanavir and Lopinavir) M36I (a consensus amino acid in subtype C selected by Ritinavir) M36V (a polymorphism associated with Atazanvir and Ritonavir; L63P, a polymorphism selected by Ritonavir) H69K (a polymorphism selected by Tipranavir or Ritonavir) V77I (a polymorphism selected by Nelfinavir; V82I, polymorphism selected by Atanazavir or Ritonavir) L89M (a polymorphism selected by Tipranavir or Ritonavir) and I93L (a polymorphism weakly associated with subtype B protease inhibitor treatments) were revealed in direct PCR sequencing and one other I62V a polymorphism selected by Ritonavir was revealed in all sequenced viral clones. The majority of these mutations are uncharacterized and as such, their effect on protease inhibitors is unknown. This could be because the HIV RNA polymerase has a high error rates and lacks a proofreading mechanism. Hence not all mutations present in the HIV genome are drug related mutations.

Table 3.2c: Other mutations observed in the Protease gene of PCR and cloned sequences

Mutation	Frequency (n=8)	Frequency (n=35)	HIV-1 subtype	Comment
L10P	-	3%	C	Unknown
T12S	88%	89%	C	Unknown
K14R	25%	7%	C	Unknown
I15V	88%	86%	C	Unknown
Q18K	13%	3%	C	Unknown
L19I	88%	54%	C	Unknown
L19T	13%	14%	C	Unknown
K20R	50%	66%	C	Highly polymorphic ,selected commonly by Nelfinavir, Ritonavir, Atazanavir, Lopinavir
E35D	13%	23%	C	Unknown
E35K	-	3%	C	Unknown
M36I	75%	71%	C	Polymorphism weakly associated with PI Resistance in subtype B viruses; Consensus amino acid in subtype C selected by Ritonavir
M36V	13%	11%	C	Polymorphism associated with Atazanavir+/- Ritonavir; Ritonavir
N37R	-	3%	C	Unknown
N37S	13%	11%	C	Unknown
P39S	25%	11%	C	Unknown

R41K	100%	94%	C	Unknown
K45R	-	6%	C	Unknown
Q58R	-	7%	C	Unknown
I62V	-	3%	C	Unknown
L63P	75%	51%	C	Polymorphism selected by Ritonavir
L63T	25%	23%	C	Unknown
L63A	25%	9%	C	Unknown
E65G	-	35	C	Unknown
H69K	88%	94%	C	Polymorphism selected by Tipranavir/Ritonavir
V75I	13%	14%	C	Unknown
V77I	50%	37%	C	Polymorphism selected by Nelfinavir
V82I	13%	14%	C	Polymorphism selected by Atazanavir+/-Ritonavir
R87G	-	6%	C	Unknown
L89M	88%	91%	C	Polymorphism selected by Tipranavir/Ritonavir
I93L	100%	100%	C	Polymorphism, weakly associated with subtype B PI treatment
T96S	13%	3%	C	Unknown

In comparison to other mutations observed from viral clones, mutations revealed in table 3.4c were also revealed in all viral clones. However, other mutations such as I62V a polymorphism selected by Ritonavir and uncharacterized mutations such as L10P, K45R, Q58R, L63S and R87G were revealed in protease viral clones.

3.4.4 Other mutations detected in the reverse transcriptase gene by bulk PCR sequencing and cloned viral sequences

Thirty five reverse transcriptase uncharacterized mutations were revealed from direct PCR sequences as shown in table 3.4d. The same mutations were revealed in all reverse transcriptase viral clones. Other mutations such as K82R, K102R, V111E, Y115K, V118A, G153R, I167V, I195V, I202V, Q207N, K220E, Y232H, L282P and R284K were revealed in reverse transcriptase viral clones sequences and are all uncharacterized.

Table 3.2d: Other mutations observed in the Reverse Transcriptase gene of PCR products

Mutation	Frequency (n=8)	Frequency (n=33)	HIV-1 subtype	Comment
K102Q	12.5%		C	Unknown
K102R	-	12%	C	Unknown
V111E	-	3%	C	Unknown
Y115X	-	3%	C	Unknown
V118A	-	9%	C	Unknown
D121H	12.5%	6%	C	Unknown
D121N	12.5%	-	C	Unknown
D121Y	12.5%	15%	C	Unknown
K122E	100%	100%	C	Unknown
D123G	37.5%	48%	C	Unknown
D123S	12.5%	-	C	Unknown
D123N	25%	27%	C	Unknown
I135V	12.5%	9%	C	Unknown
I135T	12.5%	15%	C	Unknown
T139A	12.5%	15%	C	Unknown
G152R	-	3%	C	Unknown
S162C	12.5%	6%	C	Unknown
S156P	-	3%	C	Unknown
T165I	12.5%	15%	C	Unknown
K166R	12.5%	30%	C	Unknown
I167V	-	3%	C	Unknown
K173T	12.5%	15%	C	Unknown
K173A	75%	70%	C	Unknown
Q174R	25%	30%	C	Unknown
D177E	87.5%	91%	C	Unknown
D177G	12.5%	9%	C	Unknown
I178V	12.5%	9%	C	Unknown
T200A	100%	100%	C	Unknown
I202A	12.5%	-	C	Unknown
I202V	-	12%	C	Unknown
I202Q	-	3%	C	Unknown
E203D	12.5%	9%	C	Unknown
E204K	25%	24%	C	Unknown
Q207E	75%	79%	C	Unknown
Q207A	12.5%	-	C	Unknown
Q207X	12.5%	-	C	Unknown
Q207N	-	15%	C	Unknown
R211K	75%	61%	C	Unknown
F214L	12.5%	9%	C	Unknown
K220E	-	3%	C	Unknown
L228R	12.%	6%	C	Unknown
Y232H	-	3%	C	Unknown
V245Q	100%	100%	C	Unknown
E248T	25%	30%	C	Unknown
D250E	12.5%	9%	C	Unknown

S251N	12.5%	9%	C	Unknown
T253S	12.5%	9%	C	Unknown
G262R	-	3%	C	Unknown
A272P	-	84%	C	Unknown
I274V	25%	6%	C	Unknown
K277R	50%	76%	C	Unknown
Q278N	12.5%	6%	C	Unknown
L282P	-	3%	C	Unknown
L283I	12.5%	9%	C	Unknown
R284K	-	12%	C	Unknown

3.5 VIRAL SUBTYPE DETERMINATION

Viral subtyping for Protease sequences and Reverse Transcriptase sequences was determined by drawing phylogenetic trees to analyze sample relatedness. Direct PCR sequences and viral cloned sequences showed intermingling with diversity variations and were HIV-1 subtype C. There was no contamination because bulk PCR sequences did not cluster on one tree branch for both protease and reverse transcriptase genes.

Separate trees were drawn for protease and reverse transcriptase genes for sequences obtained from direct PCR sequencing (figures 3.5a and 3.5b) and viral cloned sequences (figures 3.5c and 3.5d). Figures 3.5e and 3.5f are phylogenetic trees for protease and reverse transcriptase gene bulk PCR sequences in comparison to viral cloned sequences.

Figure 3.5a: Phylogenetic analysis of Protease nucleotide sequences of 8 samples from HIV-1 infected individuals attending the Bela-Bela HIV/AIDS Prevention Group and University of Venda HIV/AIDS unit. Using the neighbor-joining method, the tree was drawn with the test sequences (shown in bold letters) clustering and intermingling with HIV-1 subtype C reference sequences obtained from the HIV GenBank.

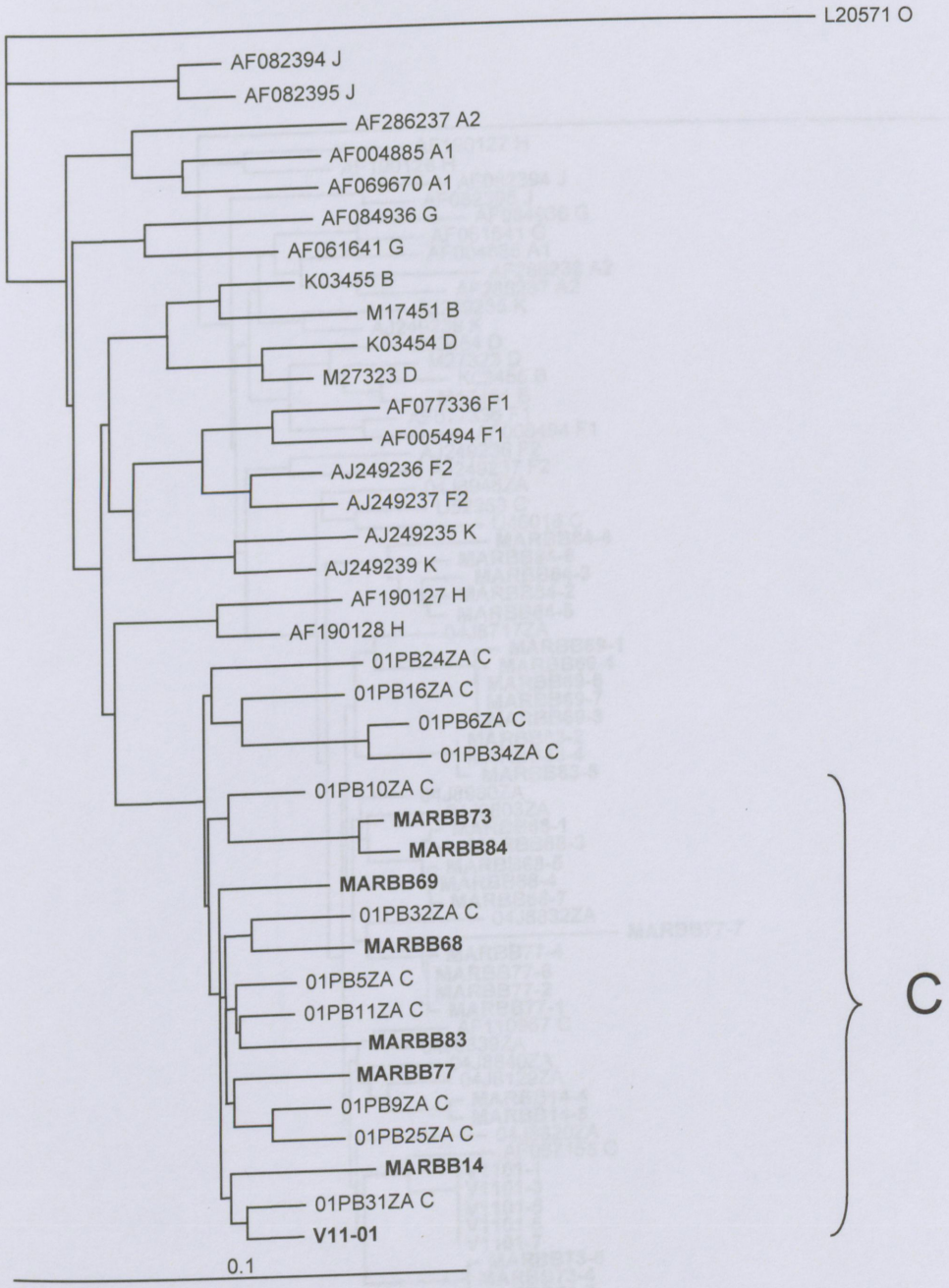


Figure 3.5b: Phylogenetic analysis of reverse transcriptase nucleotide PCR sequences from HIV-1 infected individuals attending the Bela-Bela HIV/AIDS Prevention Group and University of Venda HIV/AIDS unit. Using the neighbor-joining method, the tree was drawn with the test sequences (shown in bold letters) clustering and intermingling with HIV-1 subtype C reference sequences obtained from the HIV GenBank.

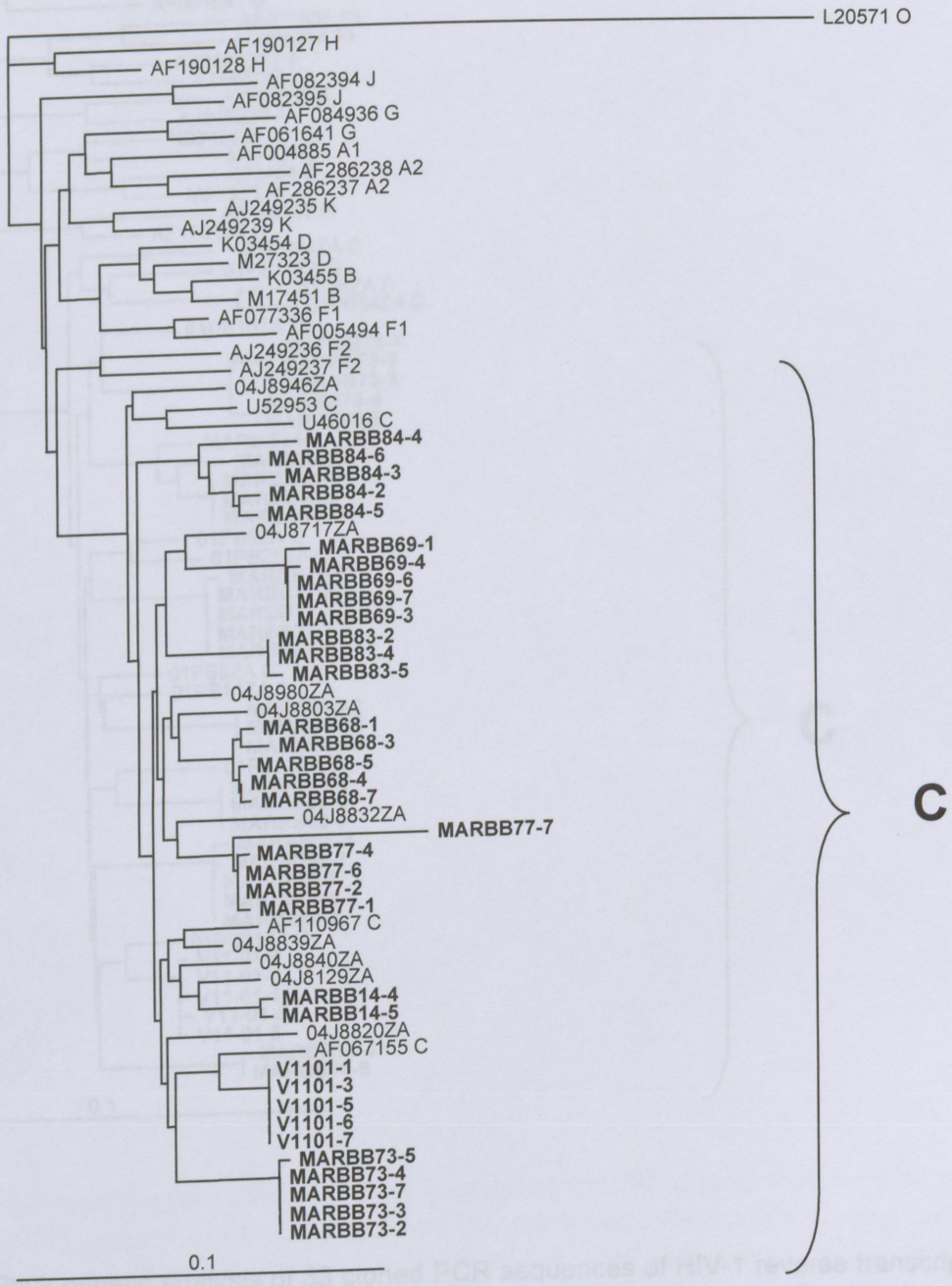


Figure 3.5c: Phylogenetic analysis of 35 cloned PCR sequences of HIV-1 Protease nucleotide sequences from 8 PCR samples from HIV-1 infected individuals attending the Bela-Bela HIV/AIDS Prevention Group and University of Venda HIV/AIDS unit. Using the neighbor-joining method, the tree was drawn with the test sequences (shown in bold letters) clustering and intermingling with HIV-1 subtype C reference sequences obtained from the HIV GenBank. Sequences are identical because they are cloned sequences and show little variations in the viral population of the protease gene.

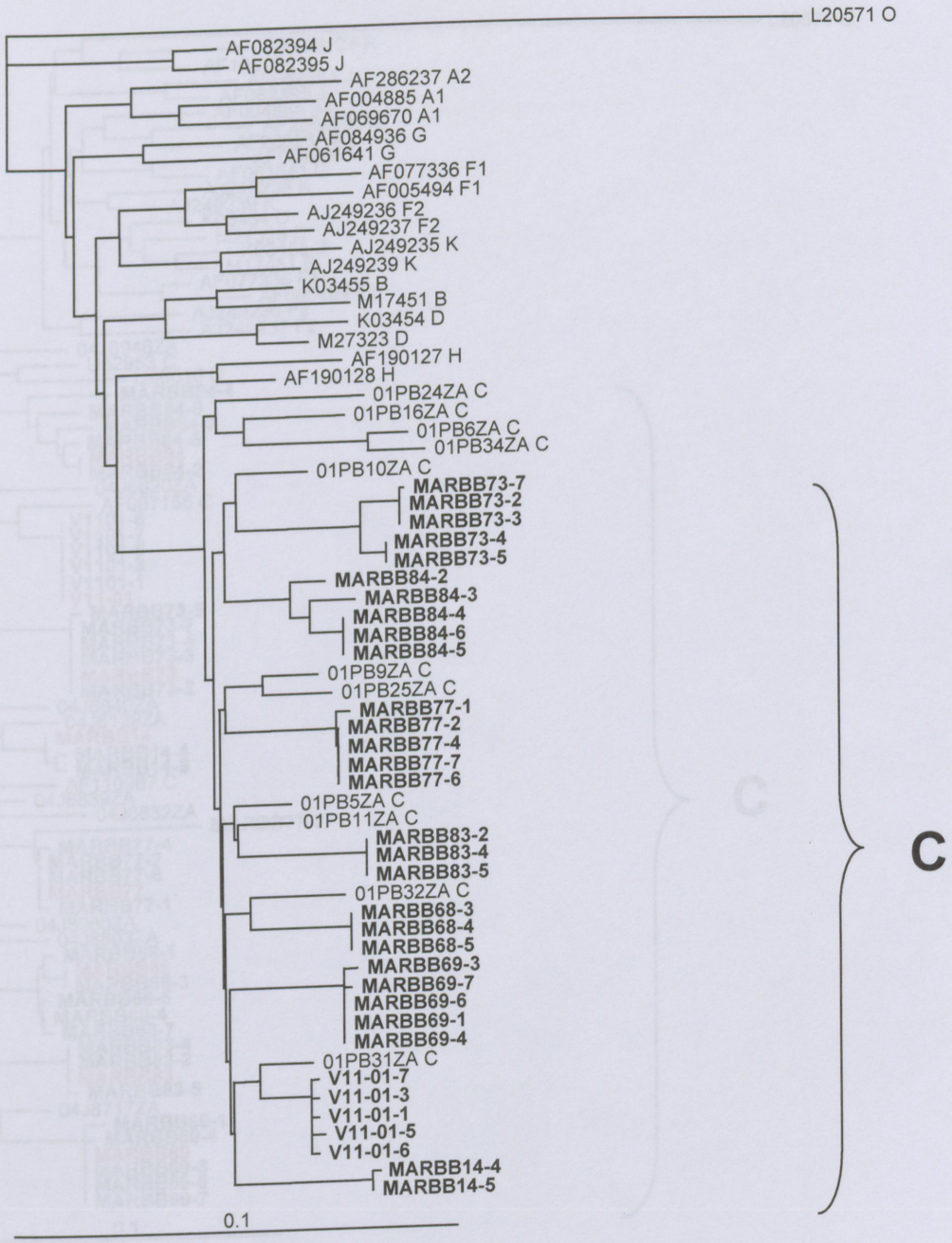
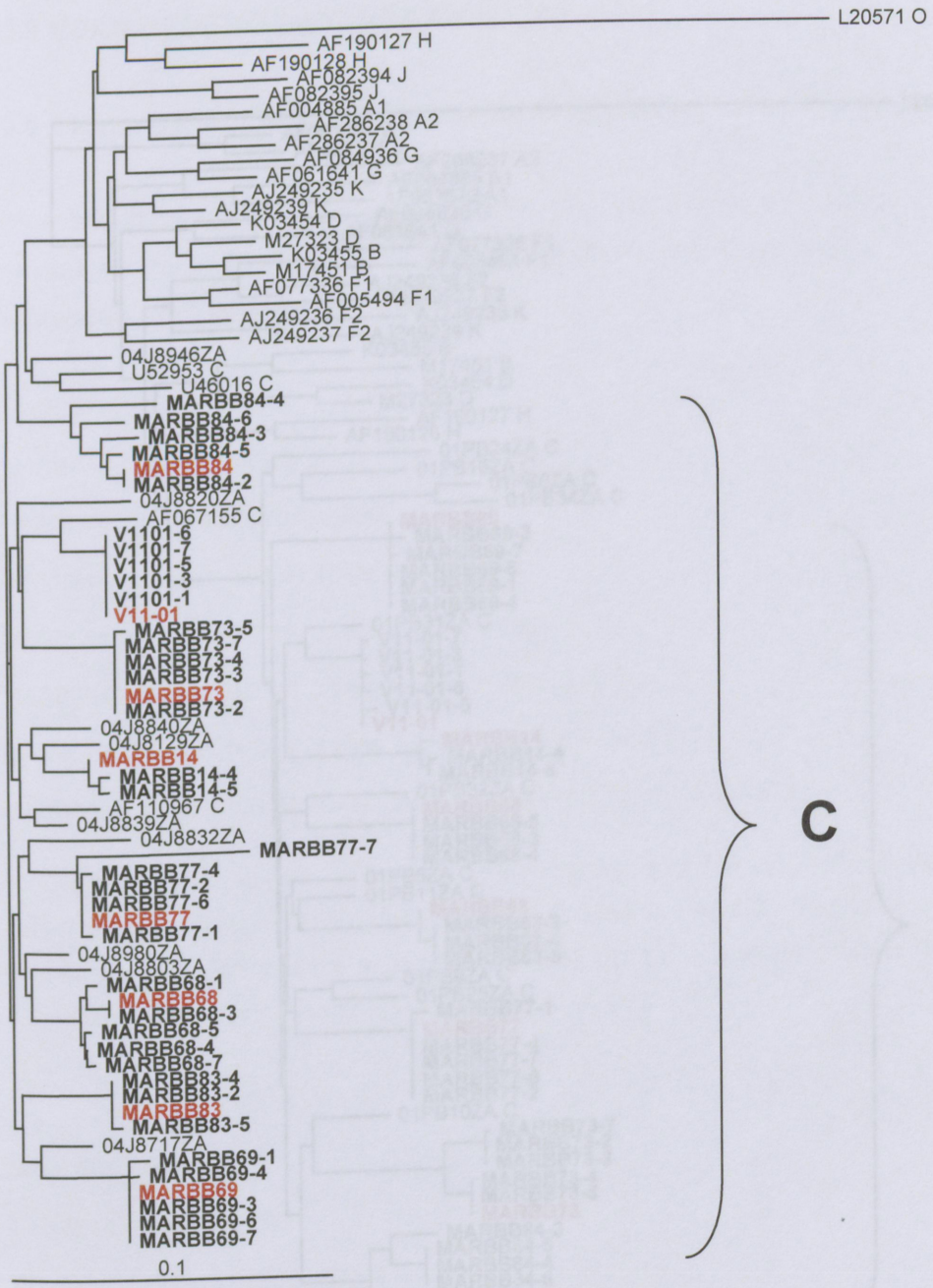


Figure 3.5d: Phylogenetic analysis of 33 cloned PCR sequences of HIV-1 reverse transcriptase nucleotide sequences from 8 PCR samples from HIV-1 infected individuals attending the Bela-Bela HIV/AIDS Prevention Group and University of Venda HIV/AIDS unit. Using the neighbor-joining method, the tree was drawn with the test sequences (shown in bold letters) clustering and intermingling with HIV-1 subtype C reference sequences obtained from the HIV GenBank. Sequences are identical because there are cloned sequences and show little variations in the viral population of the protease gene.



-Figure 3.5e: Phylogenetic analysis of 8 PCR sequences and 35 cloned PCR sequences for protease nucleotide sequences of individuals from HIV-1 infected individuals attending the Bela-Bela HIV/AIDS Prevention Group and University of Venda HIV/AIDS unit. Using the neighbor-joining method, the tree was drawn with the PCR sequences and cloned PCR sequences (shown in bold letters; PCR and cloned PCR sequences are denoted by Red and Black colors respectively) clustering and intermingling with HIV-1 subtype C reference sequences obtained from the HIV GenBank. Close relatedness was seen between PCR sequences and their cloned PCR sequences. Some sequences of the same samples were identical because they are clones and show little variation in the viral population.

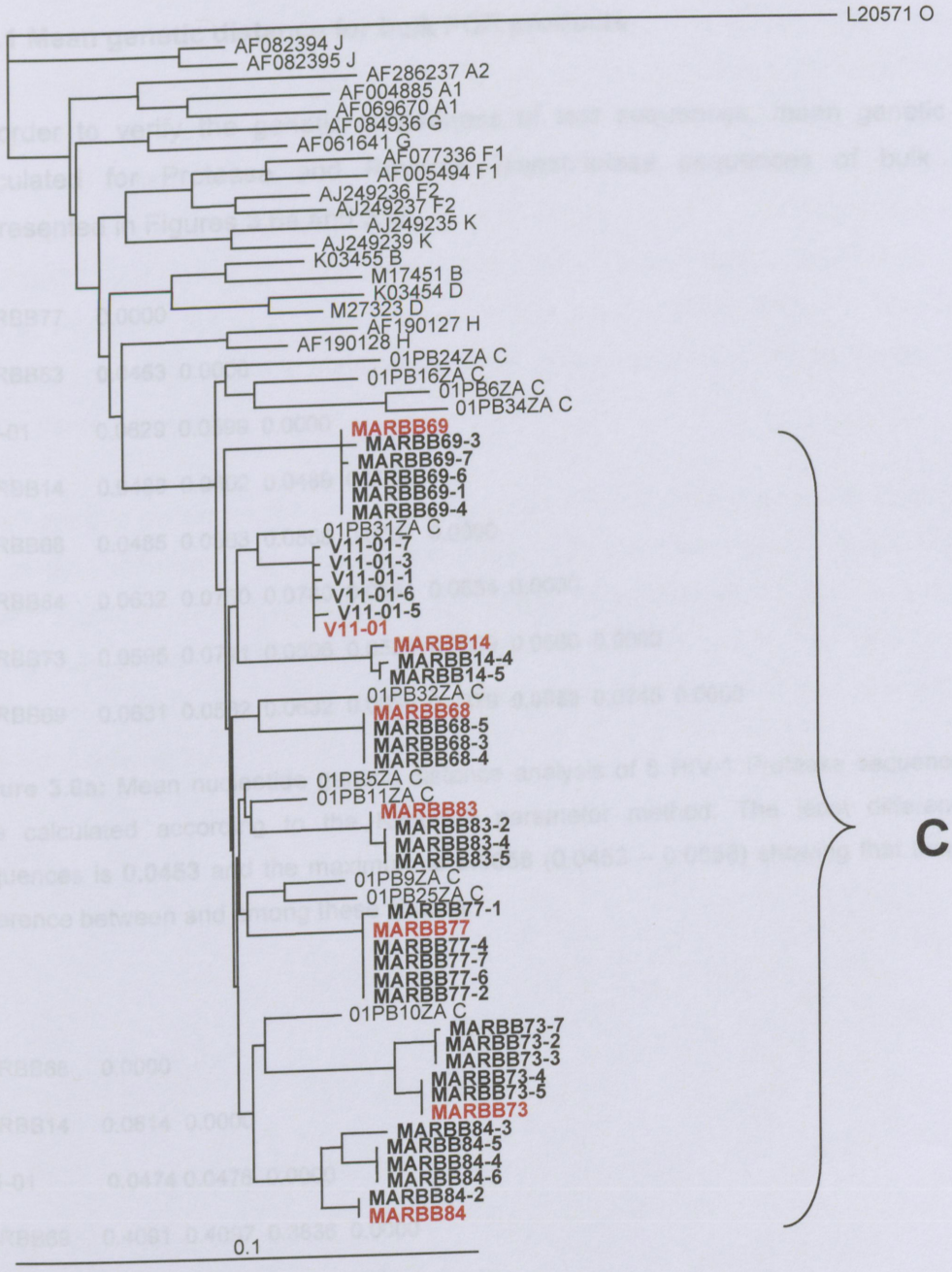


Figure 3.5f: Phylogenetic analysis of 8 PCR and 33 cloned PCR reverse transcriptase nucleotide sequences of individuals from HIV-1 infected individuals at the Bela-Bela HIV/AIDS prevention group and University of Venda HIV/AIDS Unit. Using the neighbor-joining method, the tree was drawn with the test sequences (shown in bold letters; PCR and cloned PCR sequences are denoted by Red and Black colors respectively) clustering and intermingling with HIV-1 subtype C reference sequences obtained from the HIV GenBank. Close relatedness was seen between PCR sequences and their cloned PCR sequences and some samples were identical because there are clones and show little variation in the viral population.

3.6 MEAN GENETIC DISTANCE ANALYSIS

3.6.1 Mean genetic distance for bulk PCR products

In order to verify the genetic relatedness of test sequences, mean genetic distance were calculated for Protease and Reverse Transcriptase sequences of bulk PCR products represented in Figures 3.6a and 3.6b.

MARBB77	0.0000						
MARBB83	**0.0453**	0.0000					
V11-01	0.0629	0.0599	0.0000				
MARBB14	0.0488	0.0602	0.0489	0.0000			
MARBB68	0.0485	0.0563	0.0558	0.0561	0.0000		
MARBB84	0.0632	0.0750	0.0780	0.0561	0.0634	0.0000	
MARBB73	0.0595	0.0711	0.0596	0.0524	0.0669	0.0560	0.0000
MARBB69	0.0631	0.0562	0.0632	0.0636	0.0379	**0.0858**	0.0745	0.0000

Figure 3.6a: Mean nucleotide genetic distance analysis of 8 HIV-1 Protease sequences. The distance was calculated according to the Kimura-2 parameter method. The least difference between the sequences is **0.0453** and the maximum is **0.0858** (0.0453 – 0.0858) showing that there is little genetic difference between and among these viruses.

MARBB68	0.0000						
MARBB14	0.0614	0.0000					
V11-01	**0.0474**	0.0476	0.0000				
MARBB69	0.4091	0.4097	0.3836	0.0000			
MARBB73	0.0738	0.0723	0.0565	0.4192	0.0000		
MARBB77	0.0561	0.0650	0.0475	0.3953	0.0759	0.0000	
MARBB83	0.0629	0.0650	0.0526	0.4108	0.0706	0.0612	0.0000
MARBB84	0.0561	0.0668	0.0477	**0.4275**	0.0617	0.0598	0.0581	0.0000

Figure 3.6b: Mean nucleotide genetic distance analysis of 8 HIV-1 Reverse Transcriptase sequences. The distance was calculated according to the Kimura-2 parameter method. The least difference between the sequences is **0.0474** and the maximum is **0.4275** (0.0474 – 0.4275) showing that there is little genetic difference between and among these viruses.

3.6.2 Mean genetic distance for cloned PCR products

Mean genetic distance for protease and reverse transcriptase cloned PCR sequences were also calculated. This was done in order to verify the genetic relatedness of cloned PCR sequences in comparison to PCR product sequences. For the protease gene (figure 3.6c), viral cloned sequences of sample V11-01 showed no genetic differences. The same was observed in cloned PCR sequences of samples MARBB68 viral clones 5 and 7, MARBB69 viral clones 6 and 7, MARBB73 viral clones 2, 3 and 5, MARBB77 viral clones 2 and 6 and MARBB83 viral clones 2 and 4. However, for viral cloned sequences of sample MARBB14 and MARBB84, their viral clones were not identical and revealed slight genetic variations among cloned PCR sequences of the same samples.

Mean genetic distances for viral cloned reverse transcriptase sequences (figure 3.6d) revealed no genetic difference for viral cloned sequences from samples MARBB68, MARBB73 viral clones 2 and 3, MARBB73 viral clones 4 and 5, MARBB84 viral clones 4, 5 and 6, MARBB84 viral clones 3 and 5, MARBB77 viral clones 2, 4, 6 and 7 and V11-01 viral clones 1, 3 and 7. Viral cloned sequences of sample MARBB14 were not identical and revealed slight genetic variations among them.

MARBB68-1	0.0000
MARBB68-3	0.0101	0.0000
MARBB68-4	0.0034	0.0135	0.0000
MARBB68-5	0.0067	0.0169	0.0034	0.0000
MARBB68-7	0.0067	0.0169	0.0034	0.0067	0.0000
MARBB69-1	0.0817	0.0853	0.0781	0.0818	0.0817	0.0000
MARBB69-3	0.0743	0.0778	0.0707	0.0744	0.0743	0.0068	0.0000
MARBB69-4	0.0781	0.0817	0.0745	0.0782	0.0781	0.0102	0.0034	0.0000
MARBB69-6	0.0743	0.0778	0.0707	0.0744	0.0743	0.0068	0.0000	0.0034	0.0000
MARBB69-7	0.0743	0.0778	0.0707	0.0744	0.0743	0.0068	0.0000	0.0034	0.0000	0.0000
MARBB73-2	0.0669	0.0705	0.0633	0.0670	0.0669	0.0820	0.0746	0.0784	0.0746	0.0746	0.0000
MARBB73-3	0.0669	0.0705	0.0633	0.0670	0.0669	0.0820	0.0746	0.0784	0.0746	0.0746	0.0000	0.0000
MARBB73-5	0.0669	0.0705	0.0633	0.0670	0.0669	0.0820	0.0746	0.0784	0.0746	0.0746	0.0000	0.0000	0.0000
MARBB73-6	0.0705	0.0741	0.0669	0.0706	0.0705	0.0857	0.0782	0.0821	0.0782	0.0782	0.0034	0.0034	0.0034	0.0000
MARBB73-7	0.0669	0.0705	0.0633	0.0670	0.0669	0.0820	0.0746	0.0784	0.0746	0.0746	0.0000	0.0000	0.0000	0.0034	0.0000
MARBB84-2	0.0634	0.0669	0.0598	0.0635	0.0561	0.0935	0.0859	0.0898	0.0859	0.0859	0.0561	0.0561	0.0561	0.0596	0.0561	0.0000
MARBB84-3	0.0668	0.0631	0.0632	0.0669	0.0596	0.1046	0.0970	0.1009	0.0970	0.0970	0.0668	0.0668	0.0668	0.0704	0.0668	0.0170	0.0000
MARBB84-4	0.0739	0.0701	0.0703	0.0740	0.0666	0.1042	0.0966	0.1006	0.0966	0.0966	0.0816	0.0816	0.0816	0.0778	0.0816	0.0486	0.0449	0.0000
MARBB84-6	0.0597	0.0632	0.0561	0.0598	0.0525	0.0896	0.0821	0.0860	0.0821	0.0821	0.0670	0.0670	0.0670	0.0706	0.0670	0.0240	0.0344	0.0450	0.0000
MARBB84-7	0.0633	0.0668	0.0597	0.0634	0.0561	0.0934	0.0858	0.0897	0.0858	0.0858	0.0524	0.0524	0.0524	0.0560	0.0524	0.0102	0.0204	0.0486	0.0275	0.0000
MARBB14-4	0.0708	0.0744	0.0672	0.0709	0.0708	0.0711	0.0711	0.0750	0.0711	0.0711	0.0598	0.0598	0.0598	0.0634	0.0598	0.0635	0.0743	0.0818	0.0598	0.0598	0.0000
MARBB14-5	0.0706	0.0742	0.0670	0.0707	0.0706	0.0709	0.0709	0.0748	0.0709	0.0709	0.0597	0.0597	0.0597	0.0633	0.0597	0.0633	0.0741	0.0816	0.0597	0.0597	0.0068	0.0000
V1101-1	0.0703	0.0738	0.0667	0.0704	0.0703	0.0706	0.0633	0.0670	0.0633	0.0633	0.0596	0.0596	0.0596	0.0632	0.0596	0.0781	0.0890	0.0887	0.0818	0.0780	0.0562	0.0561	0.0000
V1101-3	0.0703	0.0738	0.0667	0.0704	0.0703	0.0706	0.0633	0.0670	0.0633	0.0633	0.0596	0.0596	0.0596	0.0632	0.0596	0.0781	0.0890	0.0887	0.0818	0.0780	0.0562	0.0561	0.0000	0.0000
V1101-5	0.0703	0.0738	0.0667	0.0704	0.0703	0.0706	0.0633	0.0670	0.0633	0.0633	0.0596	0.0596	0.0596	0.0632	0.0596	0.0781	0.0890	0.0887	0.0818	0.0780	0.0562	0.0561	0.0000	0.0000	0.0000
V1101-6	0.0703	0.0738	0.0667	0.0704	0.0703	0.0706	0.0633	0.0670	0.0633	0.0633	0.0596	0.0596	0.0596	0.0632	0.0596	0.0781	0.0890	0.0887	0.0818	0.0780	0.0562	0.0561	0.0000	0.0000	0.0000	0.0000
V1101-7	0.0703	0.0738	0.0667	0.0704	0.0703	0.0706	0.0633	0.0670	0.0633	0.0633	0.0596	0.0596	0.0596	0.0632	0.0596	0.0781	0.0890	0.0887	0.0818	0.0780	0.0562	0.0561	0.0000	0.0000	0.0000	0.0000	0.0000
MARBB77-1	0.0450	0.0413	0.0414	0.0379	0.0450	0.0741	0.0668	0.0705	0.0668	0.0668	0.0632	0.0632	0.0632	0.0668	0.0632	0.0669	0.0631	0.0629	0.0632	0.0669	0.0597	0.0596	0.0665	0.0665	0.0665	0.0665	0.0665	0.0000
MARBB77-2	0.0414	0.0378	0.0379	0.0344	0.0414	0.0705	0.0632	0.0669	0.0632	0.0632	0.0596	0.0596	0.0596	0.0631	0.0596	0.0633	0.0595	0.0593	0.0596	0.0632	0.0561	0.0560	0.0629	0.0629	0.0629	0.0629	0.0629	0.0034	0.0000
MARBB77-4	0.0449	0.0413	0.0414	0.0379	0.0449	0.0741	0.0668	0.0705	0.0668	0.0668	0.0631	0.0631	0.0631	0.0667	0.0631	0.0669	0.0631	0.0628	0.0632	0.0668	0.0597	0.0523	0.0665	0.0665	0.0665	0.0665	0.0665	0.0068	0.0034	0.0000
MARBB77-6	0.0414	0.0378	0.0379	0.0344	0.0414	0.0705	0.0632	0.0669	0.0632	0.0632	0.0596	0.0596	0.0596	0.0631	0.0596	0.0633	0.0595	0.0593	0.0596	0.0632	0.0561	0.0560	0.0629	0.0629	0.0629	0.0629	0.0629	0.0034	0.0000	0.0034	0.0000
MARBB77-7	0.1230	0.1188	0.1191	0.1152	0.1230	0.1523	0.1440	0.1484	0.1440	0.1440	0.1439	0.1439	0.1439	0.1480	0.1439	0.1483	0.1438	0.1432	0.1441	0.1400	0.1404	0.1399	0.1474	0.1474	0.1474	0.1474	0.1474	0.0799	0.0762	0.0799	0.0762	0.0000
MARBB83-2	0.0524	0.0559	0.0489	0.0525	0.0524	0.0636	0.0563	0.0600	0.0563	0.0563	0.0712	0.0712	0.0712	0.0749	0.0712	0.0751	0.0860	0.0782	0.0639	0.0750	0.0677	0.0676	0.0599	0.0599	0.0599	0.0599	0.0599	0.0489	0.0454	0.0489	0.0454	0.1282	0.0000
MARBB83-4	0.0524	0.0559	0.0489	0.0525	0.0524	0.0636	0.0563	0.0600	0.0563	0.0563	0.0712	0.0712	0.0712	0.0749	0.0712	0.0751	0.0860	0.0782	0.0639	0.0750	0.0677	0.0676	0.0599	0.0599	0.0599	0.0599	0.0599	0.0489	0.0454	0.0489	0.0454	0.1282	0.0000	0.0000
MARBB83-5	0.0560	0.0595	0.0524	0.0561	0.0560	0.0672	0.0599	0.0636	0.0599	0.0599	0.0749	0.0749	0.0749	0.0785	0.0749	0.0787	0.0898	0.0819	0.0675	0.0787	0.0714	0.0712	0.0635	0.0635	0.0635	0.0635	0.0635	0.0525	0.0489	0.0525	0.0489	0.1322	0.0034	0.0034	0.0000

Figure 3.6c: Mean nucleotide genetic distance analysis of 35 HIV-1 Protease cloned sequences. The distance was calculated according to the Kimura-2 parameter method. Cloned PCR sequences revealed slight genetic variations among cloned sequences of the same samples, however, no genetic differences were revealed in some cloned PCR sequences of sample V11-01.

MARBB68-3	0.0000																																
MARBB68-4	0.0000	0.0000																															
MARBB68-5	0.0000	0.0000	0.0000																														
MARBB69-1	0.0595	0.0595	0.0595	0.0000																													
MARBB69-3	0.0630	0.0630	0.0630	0.0032	0.0032	0.0000																											
MARBB69-4	0.0595	0.0595	0.0595	0.0000	0.0000	0.0032	0.0000																										
MARBB69-6	0.0595	0.0595	0.0595	0.0000	0.0000	0.0032	0.0000																										
MARBB69-7	0.0612	0.0612	0.0612	0.0016	0.0016	0.0048	0.0016	0.0000																									
MARBB73-2	0.0788	0.0788	0.0788	0.0719	0.0719	0.0754	0.0719	0.0736	0.0000																								
MARBB73-3	0.0788	0.0788	0.0788	0.0719	0.0719	0.0754	0.0719	0.0736	0.0000	0.0000																							
MARBB73-4	0.0752	0.0752	0.0752	0.0683	0.0683	0.0718	0.0683	0.0700	0.0145	0.0145	0.0000																						
MARBB73-5	0.0752	0.0752	0.0752	0.0683	0.0683	0.0718	0.0683	0.0700	0.0145	0.0145	0.0000	0.0000																					
MARBB73-7	0.0806	0.0806	0.0806	0.0736	0.0736	0.0772	0.0736	0.0754	0.0016	0.0016	0.0162	0.0162	0.0000																				
MARBB84-5	0.0612	0.0612	0.0612	0.0647	0.0647	0.0683	0.0647	0.0664	0.0701	0.0701	0.0665	0.0665	0.0684	0.0000																			
MARBB84-2	0.0562	0.0562	0.0562	0.0597	0.0597	0.0632	0.0597	0.0614	0.0685	0.0685	0.0615	0.0615	0.0668	0.0210	0.0000																		
MARBB84-3	0.0629	0.0629	0.0629	0.0665	0.0665	0.0700	0.0665	0.0682	0.0753	0.0753	0.0717	0.0717	0.0736	0.0176	0.0226	0.0000																	
MARBB84-4	0.0612	0.0612	0.0612	0.0647	0.0647	0.0683	0.0647	0.0664	0.0701	0.0701	0.0665	0.0665	0.0684	0.0000	0.0210	0.0176	0.0000																
MARBB84-6	0.0612	0.0612	0.0612	0.0647	0.0647	0.0683	0.0647	0.0664	0.0701	0.0701	0.0665	0.0665	0.0684	0.0000	0.0210	0.0176	0.0000	0.0000															
MARBB14-4	1.6876	1.6876	1.6876	1.6450	1.6450	1.6650	1.6450	1.6450	1.5571	1.5571	1.6221	1.6221	1.5508	1.6919	1.7110	1.6919	1.6919	1.6919	0.0000														
MARBB14-5	0.0563	0.0563	0.0563	0.0581	0.0581	0.0616	0.0581	0.0598	0.0756	0.0756	0.0719	0.0719	0.0774	0.0701	0.0650	0.0719	0.0701	0.0701	1.5901	0.0000													
V11-01-1	0.0492	0.0492	0.0492	0.0460	0.0460	0.0494	0.0460	0.0476	0.0615	0.0615	0.0580	0.0580	0.0633	0.0544	0.0494	0.0561	0.0544	0.0544	1.6769	0.0493	0.0000												
V11-01-7	0.0492	0.0492	0.0492	0.0459	0.0459	0.0494	0.0459	0.0476	0.0615	0.0615	0.0580	0.0580	0.0632	0.0543	0.0494	0.0561	0.0543	0.0543	1.6731	0.0493	0.0032	0.0000											
V11-01-3	0.0492	0.0492	0.0492	0.0459	0.0459	0.0494	0.0459	0.0476	0.0615	0.0615	0.0580	0.0580	0.0633	0.0544	0.0494	0.0561	0.0544	0.0544	1.6564	0.0493	0.0032	0.0032	0.0000										
V11-01-5	0.0509	0.0509	0.0509	0.0476	0.0476	0.0510	0.0476	0.0493	0.0632	0.0632	0.0597	0.0597	0.0650	0.0560	0.0511	0.0578	0.0560	0.0560	1.6564	0.0510	0.0048	0.0048	0.0048	0.0000									
V11-01-6	0.0492	0.0492	0.0492	0.0460	0.0460	0.0494	0.0460	0.0476	0.0615	0.0615	0.0580	0.0580	0.0633	0.0544	0.0494	0.0561	0.0544	0.0544	1.6769	0.0493	0.0032	0.0032	0.0032	0.0048	0.0000								
MARBB83-2	0.0647	0.0647	0.0647	0.0597	0.0597	0.0632	0.0597	0.0614	0.0757	0.0757	0.0720	0.0720	0.0774	0.0649	0.0598	0.0666	0.0649	0.0649	1.6767	0.0633	0.0562	0.0562	0.0562	0.0579	0.0562	0.0000							
MARBB83-3	0.0647	0.0647	0.0647	0.0597	0.0597	0.0632	0.0597	0.0614	0.0757	0.0757	0.0720	0.0720	0.0774	0.0649	0.0598	0.0666	0.0649	0.0649	1.6767	0.0633	0.0562	0.0562	0.0562	0.0579	0.0562	0.0000	0.0000						
MARBB83-5	0.0647	0.0647	0.0647	0.0597	0.0597	0.0632	0.0597	0.0614	0.0757	0.0757	0.0720	0.0720	0.0774	0.0649	0.0598	0.0666	0.0649	0.0649	1.6767	0.0633	0.0562	0.0562	0.0562	0.0579	0.0562	0.0000	0.0000	0.0000					
MARBB77-1	0.0596	0.0596	0.0596	0.0614	0.0614	0.0649	0.0614	0.0631	0.0826	0.0826	0.0790	0.0790	0.0844	0.0648	0.0632	0.0665	0.0648	0.0648	1.6528	0.0685	0.0528	0.0527	0.0528	0.0544	0.0528	0.0631	0.0631	0.0631	0.0000				
MARBB77-2	0.0562	0.0562	0.0562	0.0580	0.0580	0.0615	0.0580	0.0597	0.0792	0.0792	0.0755	0.0755	0.0810	0.0614	0.0598	0.0631	0.0614	0.0614	1.6565	0.0651	0.0494	0.0494	0.0494	0.0511	0.0494	0.0632	0.0632	0.0632	0.0032	0.0000			
MARBB77-4	0.0562	0.0562	0.0562	0.0580	0.0580	0.0615	0.0580	0.0597	0.0792	0.0792	0.0755	0.0755	0.0810	0.0614	0.0598	0.0631	0.0614	0.0614	1.6565	0.0651	0.0494	0.0494	0.0494	0.0511	0.0494	0.0632	0.0632	0.0632	0.0032	0.0000	0.0000		
MARBB77-7	0.0562	0.0562	0.0562	0.0580	0.0580	0.0615	0.0580	0.0597	0.0792	0.0792	0.0755	0.0755	0.0810	0.0614	0.0598	0.0631	0.0614	0.0614	1.6565	0.0651	0.0494	0.0494	0.0494	0.0511	0.0494	0.0632	0.0632	0.0632	0.0032	0.0000	0.0000	0.00[illegible]	
MARBB77-6	0.0562	0.0562	0.0562	0.0580	0.0580	0.0615	0.0580	0.0597	0.0792	0.0792	0.0755	0.0755	0.0810	0.0614	0.0598	0.0631	0.0614	0.0614	1.6565	0.0651	0.0494	0.0494	0.0494	0.0511	0.0494	0.0632	0.0632	0.0632	0.0032	0.0000	0.0000	0.0000	0.0000

Figure 3.6d: Mean nucleotide genetic distance analysis of 33 HIV-1 Reverse Transcriptase sequences. The distance was calculated according to the Kimura-2 parameter method. Cloned PCR sequences revealed slight variations among cloned PCR sequences of the same samples, however, no genetic differences were revealed in cloned PCR sequences of the same sample such as MARBB68, MARBB77, and MARBB83.

3.7 Amino acid alignment

Amino acid alignments for protease and reverse transcriptase PCR products and viral cloned sequences was done in order to detect changes in the amino acid sequences of PCR products and viral cloned sequences in comparison to the HIV-1 global consensus sequences B and C. HIV-1 global consensus B and C sequences obtained from the HIV GenBank were used for this alignment. Amino acid substitutions obtained are shown in figures 3.7a- 3.7d. Dots in the alignments indicate identity with the consensus whereas dots replaced by amino acid symbols represent the changes (differences) between the consensus sequences and the PCR product sequences and viral cloned sequences.

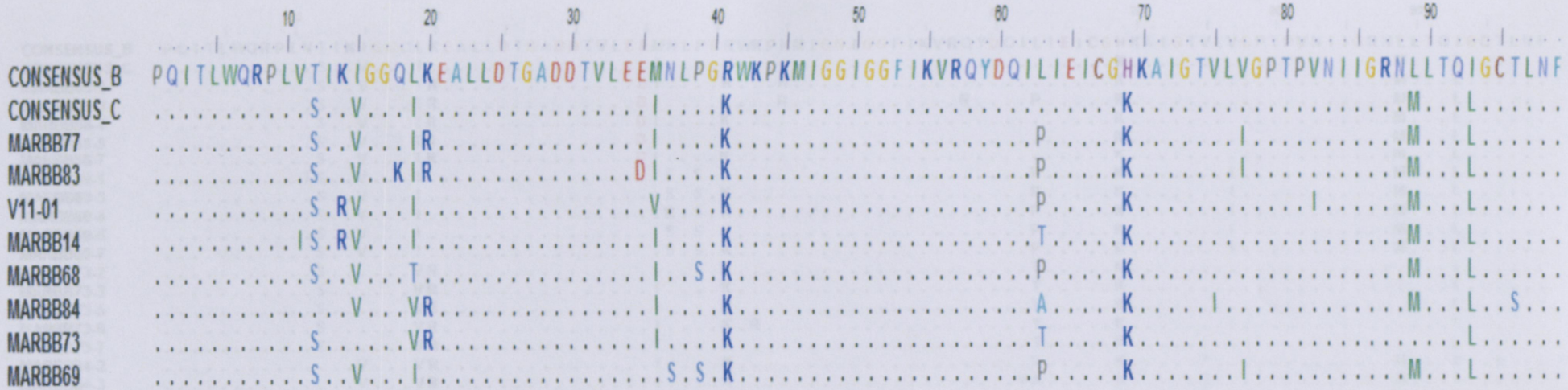


Figure 3.7a: Amino acid alignments (residues 1-99) of HIV-1 Protease gene of HIV-1 infected individuals from the Bela-Bela HIV/AIDS Prevention Group and the University of Venda HIV/AIDS Unit. The test consensus was generated form the HIV database using 8 isotopes from this study this. The global subtype B and subtype C consensus' were obtained from the HIV database programme and the entire consensus was aligned using BioEdit software. Dots in the alignments indicate identity with the consensus whereas amino acids that differ from the consensus are shown.

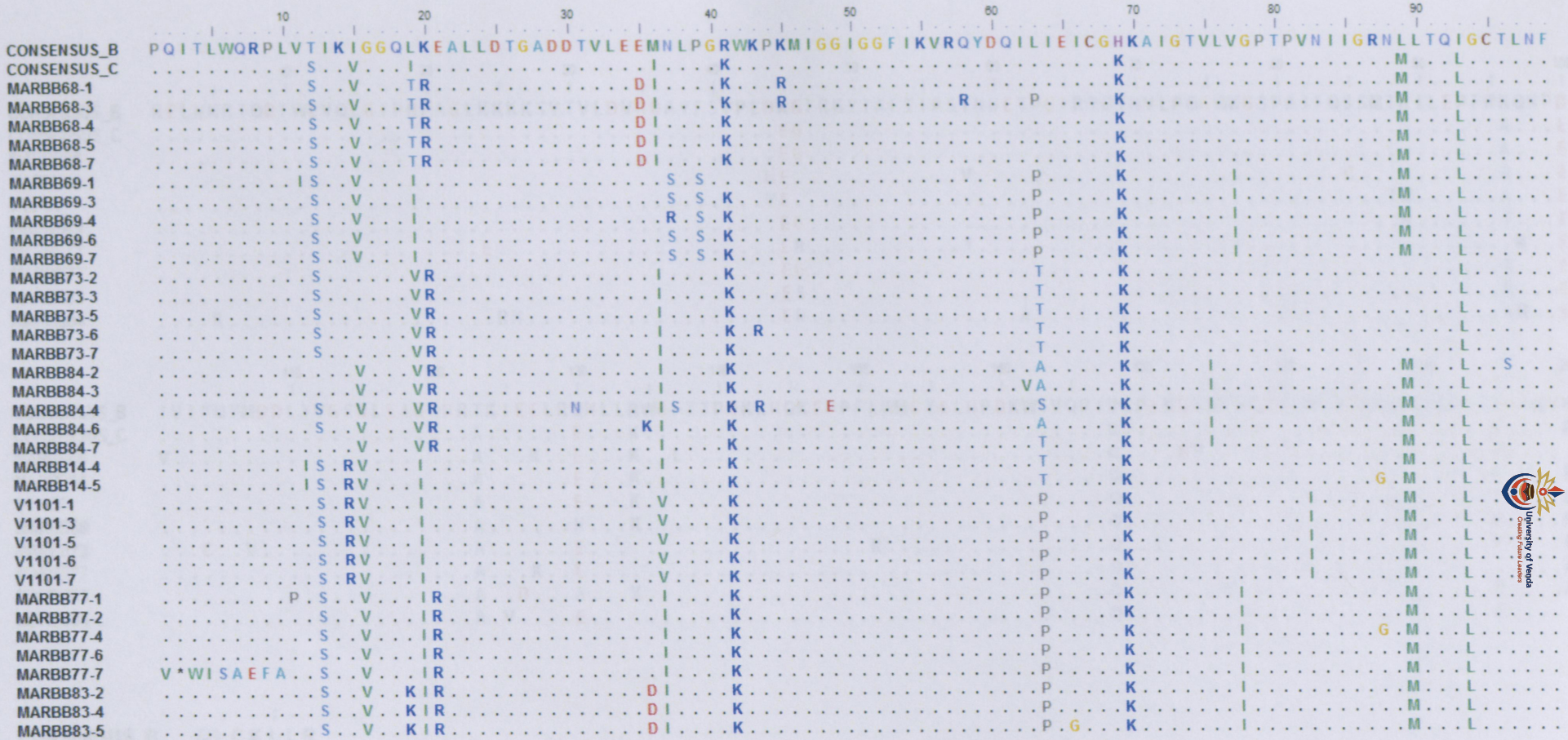


Figure 3.5b: Amino acid alignments (residues 1-99) of HIV-1 Protease cloned PCR sequences for samples from HIV-1 infected individuals from the Bela-Bela HIV/AIDS Prevention Group and the University of Venda HIV/AIDS Unit. Five clones of each PCR product were used for analysis except for samples **MARBB14** and **MARBB83** that had two and three clones respectively. Dots in the alignments indicate identity with the consensus whereas amino acids that differ from the consensus are shown and the asterisk on the viral clone **MARBB77-7** indicates an in-frame stop codon.

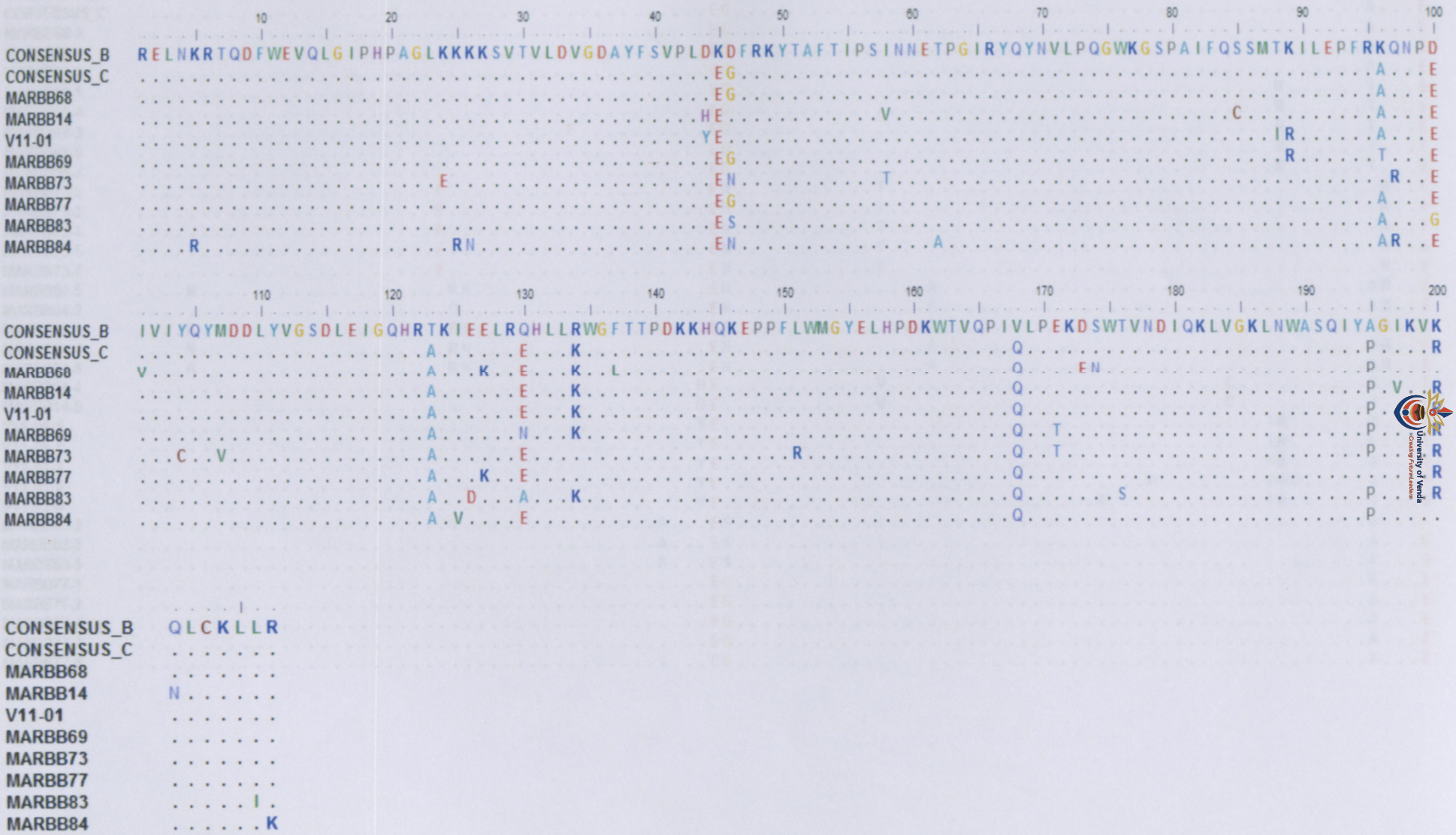
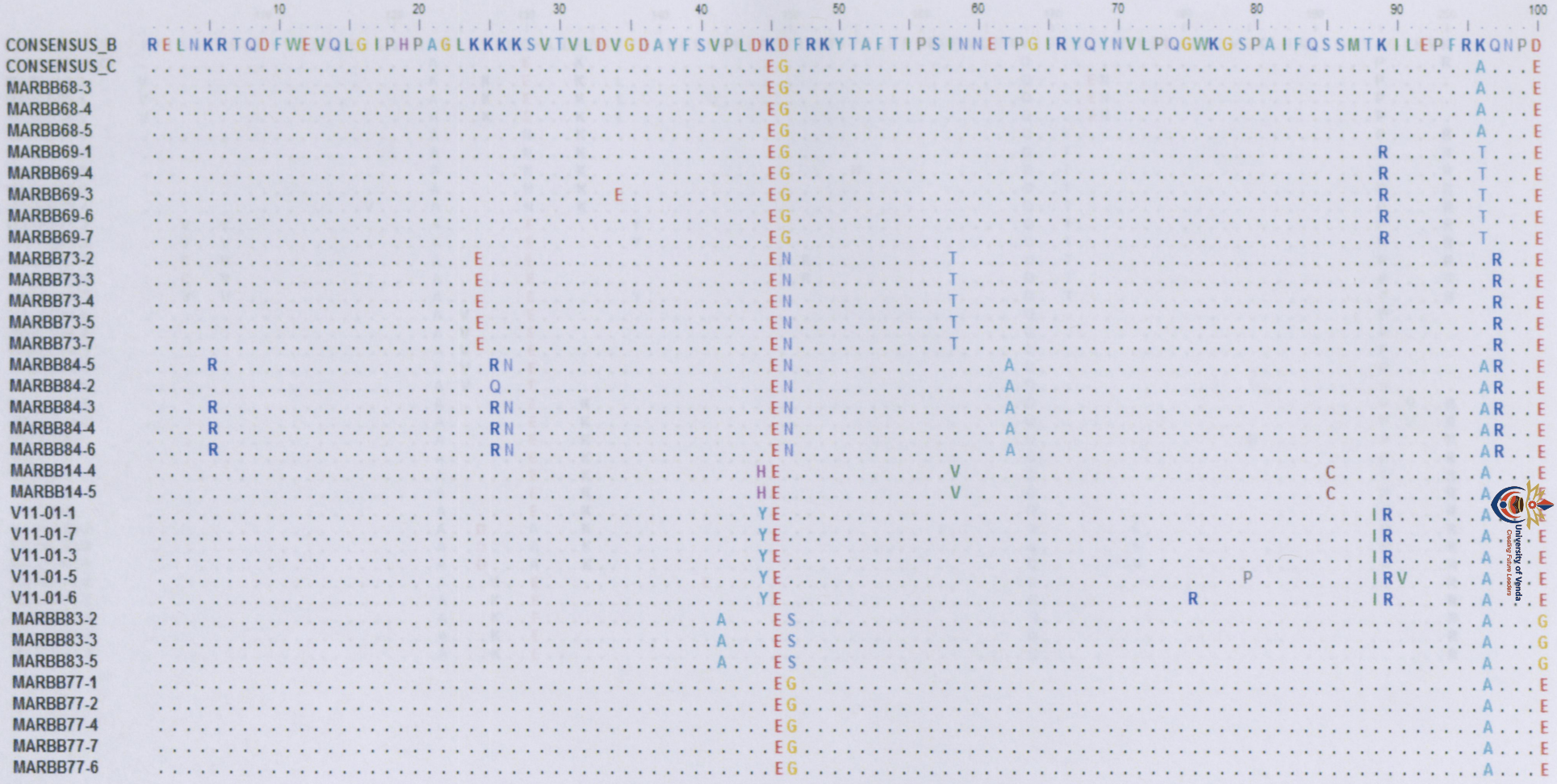
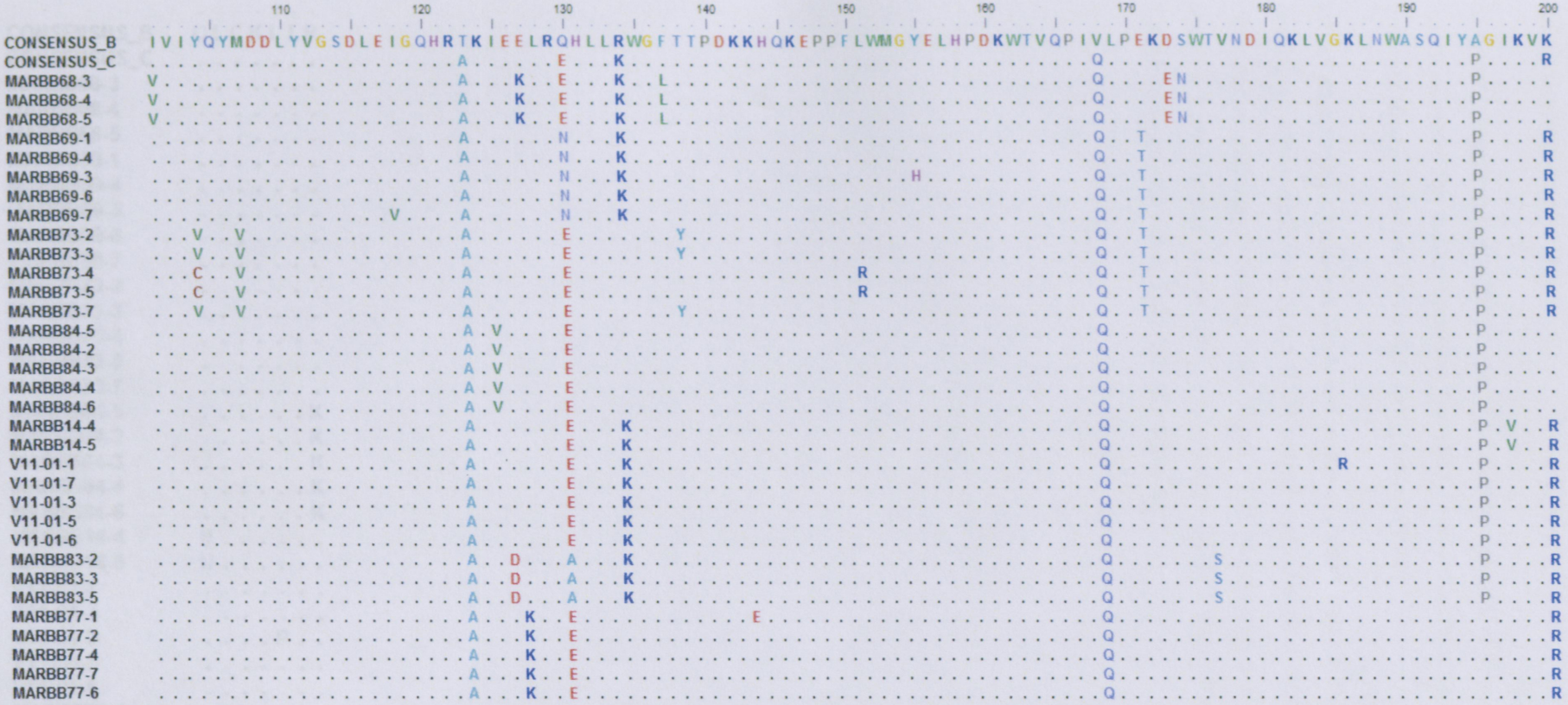


Figure 3.7c: Amino acid alignments (residues 78-285) of HIV-1 Reverse transcriptase of HIV-1 infected individuals from the Bela-Bela HIV/AIDS Prevention Group and the University of Venda HIV/AIDS Unit. The test consensus was generated form the HIV database using 8 isotopes from this study. The global subtype B and subtype C were obtained from the HIV database programme and the entire consensus was aligned using BioEdit software. Dots in the alignments indicate identity with the consensus whereas amino acids that differ from the consensus are shown.





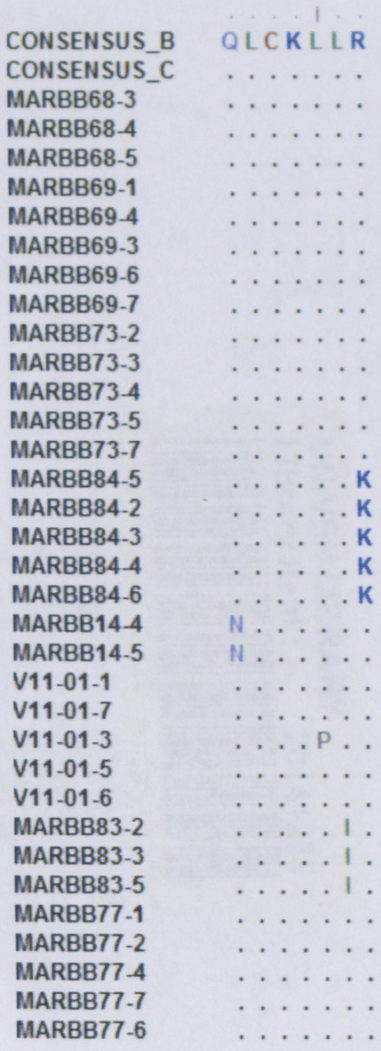


Figure 3.7d: Amino acid alignments (residues 78-285) of HIV-1 Reverse transcriptase of HIV-1 infected individuals from the Bela-Bela HIV/AIDS Prevention Group and the University of Venda HIV/AIDS Unit. Five clones of each PCR product were used for analysis expect for samples **MARBB14** which hand two clones analyzed and **MARBB83** and **MARBB68** which each had three clones analyzed. Dots in the alignments indicate identity with the consensus whereas amino acids that differ from the consensus are shown

3.8 Statistical analysis

In order to confirm the significance of observed minority mutations through cloning, a student t-test was used for the comparison between mutations observed from PCR product sequences and mutations observed from viral cloned sequences.

When comparing mutations observed in the protease gene a p-value of 0.03 was obtained from mutations observed from direct PCR sequences and viral cloned sequences. The same was done for mutations observed from direct PCR sequences and viral cloned sequences for the reverse transcriptase gene and a p-value of 0.0002 was obtained. Both values were <0.05 therefore were considered statistically significant.

3.9 Method comparisons

Table 3.6 shows a comparison of methods used in this study and other methods used for the detection of minority variants. This helped to clarify the limitations of this study and provides other options of achieving better results when detecting minority variants.

Table 3.3 Advantages and Disadvantages of PCR, cloning and other methods to detect minority variants.

	Bulk PCR sequencing	**Cloned PCR sequencing**	**Sequencing assays** (uitra-deep Pyrosequening)	**Point mutation assays** (Mutagenically-separated PCR)	**Phenotypic Assays** (P_{24})
Cost	inexpensive	Relatively inexpensive	expensive	inexpensive	expensive
Easiness of the method	Moderately labor-intensive	Labor-intensive	Labor-intensive	Moderately labor-intensive	Labor-intensive
Efficiency	Sensitivity of 100-1000 RNA copies/mL (Shafer, 2002)	100% sensitivity based on sequenced clones	100% sensitivity to all mutations in a given population	Sensitivity based on specific mutations in a population	Sensitivity of >20-50 RNA copies/mL (Shafer, 2002)
Time needed to perform assay	2-3days	3-4 days	2-3 days	2-3 days	2-3 weeks
Advantages	Detects mutations that cause drug resistance and can detect a nucleotide mixture when the least common nucleotide is present in at least 20% of the sample viral population (Lardar *et al.*, 1993)	Gives insight about the evolution of HIV-1 drug resistance (Shafer, 2002)	Detects mutations even if the phenotypic effect of the mutation is suppressed by other mutations in the sequence (Shafer, 2002)	Sensitive and specific for the detecting selected minority drug resistance mutation (Shafer, 2002)	Clinically useful by providing phenotypic selection of drug resistant variants leading to the opportunity to discover new resistant mutations (Shafer, 2002)
Disadvantages	Cannot determine whether mutations are present on different variants in the population (Gianella and Richman, 2010)	Multiple clones have to be tested individually in order to detect minority variants (Shafer, 2002)	Equipment used for sequencing is very expensive	Detects only single point mutations at a time and reduce the ability to detect alternative polymorphisms in the codons of interest (Johnson *et al.*, 2008)	In routine practice, their ability to detect minor drug-resistant variants is limited (Petrophoulos *et al.*, 2000).

CHAPTER 4: DISCUSSION, LIMITATIONS AND CONCLUSION

DISCUSSION

Human immunodeficiency virus type 1 (HIV-1) gene mutations conferring resistance to protease and reverse transcriptase inhibitors may lead to virological failure of the antiretroviral drugs being used to treat HIV-1 infections. Drug resistant mutations may be present in a minority population of the HIV-1 population and may not be detected when screening for mutations using direct PCR sequencing. Therefore, better detection of minority HIV-1 gene mutations is important because it will help on therapy decision making in clinical practices and hence the aim of this study to develop a protocol for the detection of HIV-1 drug resistant mutations.

A study by Jordan *et al.*, 2010 showed that it is important to assess intrahost virus evolutions and their relationship to disease progression, including the existence and development of low frequency drug resistance mutations and their impact on treatment outcomes. In addition not all patients respond to the current HAART regime because patients eventually develop drug resistance through the accumulation of several patterns of drug resistance mutations. Minority drug resistant variants of HIV are revealed by the development of more sensitive and precise methods to detect and quantify minority variants in the large, genetically complex population of viruses present in infected individuals. Therefore, in this study HIV-1 populations were assessed by PCR and cloning techniques to better describe HIV-1 population diversity and detect low frequency mutations.

A total of 8 samples used were of RNA extracted from plasma of 2 samples and DNA of extracted from PBMCs of 6 samples because the viral load of the samples was low leading to difficulties in viral RNA amplification and left proviral DNA as a second option. In fact, Guimaraes *et al.*, 2008 found that the resistance profile in plasma is very similar to that in PBMCs, indicating that quantification of proviral DNA is a useful and less expensive and less cumbersome tool for identifying resistance related mutations.

A DNA fragment of approximately 1.4 kb of the HIV-1 Pol (Protease and Reverse Transcriptase) gene was amplified using an optimized PCR protocol, and subsequent sequencing and phylogenetic profiling revealed that all test sequences are of HIV-1 subtype C. However, because direct PCR sequencing only depicts a population average of the viral population and not a single genome, cloning was therefore used as an alternative to single genome amplification. All 8 PCR products were cloned and a total of 35 and 33 protease and reverse

transcriptase viral clones were obtained respectively. Subsequent sequencing and phylogenetic profiling of the viral clones also revealed the relationship between the PCR sequences and viral cloned sequences by showing genetic relatedness of sequences from each PCR sample in the different branches of the tree.

Of the 8 samples, sample MARBB68 was of a treatment naïve individual and no mutation was detected from sequences obtained by direct PCR sequencing and viral cloned sequences of this sample. However, sequences obtained by direct PCR sequencing for samples of treatment experienced individuals such as MARBB14 revealed mutation V11I (a minor protease inhibitor resistant mutation) whereas MARBB73 revealed M184V (a NRTI resistant mutation) and K101E and Y181C (both NNRTI resistant mutations). Interestingly, when compared to mutations observed from sequences of cloned PCR samples, mutation V11I was revealed in viral clones of sample MARBB14 and also in one clone of sample MARBB69.

All sequences of sample MARBB73 clones revealed mutation M184V which was observed in the direct PCR sequencing of this sample. However, alongside M184V, another NRTI resistant mutation T215Y was revealed in 3 viral cloned sequences of sample MARBB73 which was not observed by direct PCR sequencing. Furthermore, NNRTI mutations K101E and Y181C observed by direct PCR sequencing were revealed in all viral cloned sequences and 3 viral cloned sequences respectively. Additionally, mutation Y181V not observed by direct PCR was revealed in 2 viral cloned sequences of sample MARBB73.

Cloned PCR sequences of samples MARBB84 all revealed mutations that were not observed by direct PCR sequencing such as K103N, a NNRTI resistant mutation revealed in all viral clones and D30N a major protease resistant mutation. D30N has more significant impact on HIV-1 subtype C than HIV-1 subtype B (Gonzalez *et al.*, 2004). Another mutation revealed was G48E which is a protease minor resistant inhibitor.

Other mutations were also observed in all direct PCR sequences and all viral cloned sequences for the protease and reverse transcriptase genes. The protease gene sequences for direct PCR and viral clones revealed polymorphisms K20R, M36I, M36V, L63P, H69K, V77I and V82I selected by protease inhibitors. However, the protease gene sequences for direct PCR and viral clones also revealed other mutations not associated with HIV-1 drug resistance. The same was observed with other mutations revealed from reverse transcriptase all direct PCR and all viral cloned sequences.

Using the student t-test, p values of >0.05 were obtained from comparisons of resistant mutations observed in PCR sequences and clone sequences for both protease and reverse transcriptase genes.

LIMITATIONS OF THE STUDY

The first limitation to this study is that, mutations were not analyzed in a broader view because of the drastically reduced sample size and therefore lead to reduced number of PCR product obtained for the mutation analysis. Secondly, the Sanger sequencing used for the sequencing of viral clones although a less expensive method rarely detects mutant strains present in lower than 20% of circulating plasma viruses or provirals (Mitsuya *et al.*, 2008). An alternative to the Sanger sequencing method would have been Utradeep Pyrosequencing, which usually reliably detects minor variants missed by the Sanger method. However, this could not be done because of limited resources.

CONCLUSION

In conclusion, the cloning technique was used as an approach to detect minority variants. Important mutations were detected in clone sequences and not in PCR sequences. This suggests that mutations of interest could be missed in the bulk PCR sequencing which is currently the norm in many settings.

CHAPTER 5: REFERENCE

Arion, D., Kaushik, N., McCormick, S., Borkow, G., Parniak, M. A. (1998). Phenotypic mechanism of HIV-1 resistance to 3_-azido-3_deoxythymidine (AZT): increased polymerization processivity and enhanced sensitivityto pyrophosphate of the mutant viral reverse transcriptase. *Biochemistry* **37**:15908–17.

Arion, D., Sluis-Cremer, N., Parniak, M. A. (2000). Mechanism by which phosphonoformic acid resistance mutations restore 3_-azido-3_-deoxythymidine (AZT) sensitivity to AZT-resistant HIV-1 reverse transcriptase. *Journal of Biological Chemistry* **275**:9251–5.

Arnold, E., Ding, L., Hughes, S. H., Hostomsky Z (1995). Structures of DNA and RNA polymerases and their interactions with nucleic acid substrates. *Current opinion in structural biology* **5**(1):27-38.

Arthos , J., Cicala, C., Martinelli, E., Macleod, K., Van Ryk, D., Wei, D *et at.* (2008). "HIV-1 envelope protein binds to and signals through integrin alpha (4) beta (7), the gut mucosal homing receptor for peripheral T cells". Nature Immunology. *In Press: 301.*

Atlas, R. M. and Bej, A. K. (1994). Polymerase Chain Reaction. In: Gerhardt, P., Murrey, R.G.E., Wood, W.A. and Krieg, N.R., (Eds.) Methods for general and molecular bacteriology. Washington, D.C.: *American Society for Microbiology*, **pp**. 418–435.

Baba, M., Nishimura, O., Kanzaki, N., Cormier E. G., Kajumo F. A., Maxwell. E *et al.* (1999). A small-molecule, nonpeptide CCR5 antagonist with highly potent and selective anti-HIV-1 activity. *Proceedings of the National Academy of Science USA* **69**(10):5698-703.

Bailey, R. C., Moses, S, Parker, C. B., Agot, K., Maclean, I., Krieger, J. N., *et al.* (2007). Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial. *Lancet* ***369*** (9562): 643–56.

Baltimore, D. (1970). Viral RNA-dependent DNA Polymerase: RNA-dependent DNA Polymerase in Virions of RNA Tumour Viruses. *Nature* **226**:1209 – 11.

Barre-Sinoussi, F., Chermann, J. C., Rey, F. (1996). HIV as the cause of AIDS. *The lancet*, **348**: 31-35.

Becker , S. L., Dezii, C. M., Burtcel, B., Kawabata, H., Hodder, S. (2002). Young HIV-infected adults are at greater risk for medication nonadherence. *Medscape General Medicine*. ***4*** (3): 21.

Bessong, P. O., Mphahlele, J., Choge I. A., Obi, L. C., Morris, L., Hammarskjold, M. L., Rekosh, D. M. (2006). Resistance mutational analysis of HIV type 1 subtype C among rural South African drug-naïve patients prior to large scale availability of antiretrovirals. *AIDS Research Human Retroviruses*. 22(12):1306-12

Billich, S., Knoop, M. T., Hansen, J., Strop, P., Sedlacek, J., Mertz, R., *et al.* (1988). Synthetic peptides as substrates and inhibitors of human immunodeficiency virus type 1 protease. *Journal of Biological Chemistry* **263**(34):17905-8.

Biti, R., French, R., Young, J., Bennetts, B., Stewar, T. G., Liang, T. (1997). HIV-1 infection in an individual homozygous for the CCR5 deletion allele. *Nature Medicine* **3**(3):252-53.

Boily, M. C., Baggaley, R. F., Wang, L., Masse, B., White, R. G., Hayes, R. J., *et al.* (2009). Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies. *Lancet Infectious Disease* ***9*** (2): 118–29.

Boucher, C. A. B., O'Sullivan, E., Mulder, J. W., Ramautarsing, C., Kellam, P., Darby, G., *et al.* (1992). Ordered appearance of zidovudine resistance mutations during treatment of 18 human immunodeficiency virus–positive subjects. *Journal of Infectious Disease* **165**:105–10.

Broder, S. and Gallo, R. C. (1984). A pathogenic retrovirus (HTLV-III) linked to AIDS. *New England Journal of Medicine* **311**:1292-7.

Brown, P. O., Bowerman, B., Varmus, H. E., Bishop J. M. (1989). Retroviral integration: structure of the initial covalent product and its precursor and role for the viral integrase protein. *Proceedings of the National Academy of Sciences* **59**(2):373–383.

Burton, G. F., Keele, B. F., Estes, J. D., Thacker, T .C., Gartner, S. (2002). Follicular dendritic cell contributions to HIV pathogenesis. *Seminars in Immunology*. **14** (4): 275–284.

Butto, S., Suligoi, B., Fanales-Belasio, E., Raimondo M. (2010). Laboratory diagnostics for HIV infection. *Annali dell'Istituto Superiore di Sanita*. **46** (1): 24-33

Carr, J. K., MO Salminen, J., Albert, E., Sanders-Buell, D., Gotte, D. L., Birx and FE McCutchan. (1998). Full genome sequences of human immunodeficiency virus type 1 subtypesG and A/G intersubtype recombinants. *Virology* **247**(1):22–31.

Celum, C. L., Coombs, R. W., Lafferty, W., Inui, T. S., Louie, P. H., Gates, C. A., *et al* (1991). Indeterminate human immunodeficiency virus type 1 western blots: seroconversion risk, specificity of supplemental tests, and an algorithm for evaluation. *Journal of Infectious Diseases*. **164** (4): 656–664.

Centers for Disease Control (1981). Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California. *Morbidity and Mortality Weekly Report* **30**: 305-8.

Centers for Disease Control and Prevention. (2001). Revised guidelines for HIV counseling, testing, and referral. *Morbidity and Mortality Weekly Report; Recommendations and Reports*. **50** (19): 1–57.

Centers for Disease Control and Prevention. (2008). Male circumcision and risk for HIV transmission and other health conditions: implications for the United States.

Centers for Disease Control and prevention. (2013). Occupational HIV transmission and prevention among Health care workers.

Chadwick, R.B., M.P. Conrad, M.D. McGin-nis, L. Johnston-Dow, S.L. Spurgeon, and M.N. Kronick. 1996. Heterozygote and mutation detection by direct automated fluorescent DNA sequencing using a mutant Taq DNA polymerase. *BioTechniques* **20**:676-683.

Chan, D. C., Fass, D., Berger, J. M., Kim, P. S. (1997). Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **89**: 263-73.

Chan, D., Kim, P. (1998). HIV entry and its inhibition. *Cell* **93** (5): 681–4.

Charneau, P., Alizon, M., Clavel, F. (1992) A second origin of DNA plus strand synthesis is required for optimal human immunodeficiency virus replication. *Journal of Virology* **66**(5)2814-20.

Charneau, P., Mirambeau, G., Roux, P., Paulous, H., Buc., Clavel, F. (1994). HIV-1 reverse transcriptase. A termination step at the center of the genome. *Journal of Molecular Biology* **241**(5):651-62.

Chene, G., Sterne, J. A., May, M., Costagliola, D., Ledergerber, B., Phillips, A. *et al* and the Antiretroviral Therapy Cohort Collaboration. (2003). Prognostic importance of initial response in HIV-1 infected patients starting potent antiretroviral therapy: analysis of prospective studies. *Lancet* ***362*** (9385): 679–686.

Choe, H., Farzan, M., Sun, Y *et al.* (1996). The beta-chemokines CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell* **85**(7):1135-48.

Clapham, P. R., McKnight, A. (2001). HIV-1 receptors and cell tropism. *British Medical Bulletin.* **58**(4): 43–59.

Collman, R., Hassan, N.F., Walker, R., Godfrey, B., Cutilli, J., Hastings, J. C., *et al.* (1989). Infection of monocyte-drived macrophages with Human immunodeficiency virus type-1 (HIV-1). Monocyte-tropic and lymphocyte-tropic strains of HIV-1 shows distinctive patterns of replication in a panel of cell types. *Journal of Experimental Medicine*. **170**(4):1149-63.

Cooper, D. A., Steigbigel, R. T., Gatell, J. M., Rockstroh, J. K., Katlama, C., Yeni, P *et al.,* (2008). Clinical trials o HIV antiretroviral therapy: Integrase inhibitors *New England Journal of Medicine* **359**(4):355-65.

Coovadia, H. (2004). Antiretroviral agents—how best to protect infants from HIV and save their mothers from AIDS. *New England Journal Medicine*. **351** (3): 289–292.

Daar, E. S., Little, S., Pitt, J. (2001). Diagnosis of primary HIV-1 infection. Los Angeles County Primary HIV Infection Recruitment Network. *Annals of Internal Medicine.* **134** (1): 25–9.

Department of Health South Africa. (2012). The 2011 National Antenatal Sentinel HIV & Syphilis Prevalence Survey in South Africa. **Pp:** 16-32.

Department of Health South Africa. (2013). The South African Antiretroviral Treatment Guidelines. **Pp:** 6-10.

Doc Kaiser's Microbiology Home Page. (2008). IV. viruses F. animal virus life cycles 3. The Life Cycle of HIV Community College of Baltimore County.

Douek, D. C., Picker, L. J., Koup, R. A. (2003). T cell dynamics in HIV-1 infection. *Annual Reviews of Immunology* **21**:265-304.

Dybul, M., Fauci, A. S., Bartlett, J. G., Kaplan, J. E., Pau, A. K.; Panel on Clinical Practices for Treatment of HIV. (2002). Guidelines for using antiretroviral agents among HIV-infected adults and adolescents. *Annals of Internal Medicine*. **137** (2): 381–433.

Dyda, F., Hickman, A. B., Jenkins, T. M., Engelman, A,. Craigie, R., Davies, D. R. (1994). Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl Transferase. *Science* 23; **266** (5193): 1981-6.

Engelman, A and Craigie, R (1992). Identification of conserved amino acid residues critical for Human immunodeficiency virus type 1 integrase function in vitro. *Journal of Virology* **66** (11): 6361-69.

Erlich, H. A. and Arnhein, N. (1992). Genetic analysis using the polymerase chain reaction. *Annual Reviews of Genetics* **26**: 479-506.

Eshleman, S. H., Jones, D., Flys, T., Petrauskene, O., Jackson, J. B. (2003). Analysis of HIV-1 Variants by cloning DNA generated with the Viroseq™ HIV-1 genotyping system. *BioTechnigues* **35**: 614-622.

Esnouf, R., Ren, J., Ross, C. Jones, Y., Stammers, D., Stuart, D. (1995) Mechanism of inhibition of HIV-1 reverse transcriptase by non- nucleoside inhibitors. *Nature Structural and molecular Biology* **2**:303-8.

Ferrantelli, F., Cafaro, A., Ensoli, B. (2004). Nonstructural HIV proteins as targets for prophylactic or therapeutic vaccines. *Current. Opinion of Biotechnology*. **15** (6): 543–56.

Gallo, R. C., Salahuddin, S. Z., Popovic, M., Shearer, G. M., Kaplan, M., Haynes, B. F., *et al.* (1984). Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* **224**: 500-3.

Gelderblom, H. R. (1997). Fine structure of HIV and SIV. In Los Alamos National Laboratory (ed.) HIV Sequence Compendium. *Los Alamos National Laboratory*. **pp**. 31–44.

Gianella, S and Richman, D. D. (2010) Minority variants of drug resistant HIV. *Journal of Infectious Diseases* **202**(5):657-666.

Gonzalez, L. M. F., Brinderio, R. M., Aguiar, R. S., Pereira, H. S., Abreu, C. M., Soares, M. A *et al.* (2004). Impact of Nelfinavir resistance mutations on in vitro phenotype, fitness and replication capacity of Human Immunodeficiency Virus type 1 with subtype B and C protease. *Antimicrobial Agents and Chemotherapy* **48**(9) 3552-3555.

Gray, R. H. (2007). Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial. *Lancet* ***369*** (9562): 657–66.

Grivel, J. C and Margolis, L. B. (1999). CCR5 and CXCR4-tropic HIV-1 are equally cytopathic for their T-cell targets in human lymphoid tissue. *Nature Medicine* **5**(3):344-6.

Guimaraes, A . P., Sa-Filho, D.J., Sucupira, M C., Janni, L. M., Diaz, R. S. (2008). Profiling resistance realted mutations in the protease region of the pol gene: single genome sequencing of HIV in plasma and peripheral blood mononuclear cells. *Aids Research and Human Retrovirueses* **24**(7): 969-71.

Halvas, E. K., Aldrovandi, G. M., Balfe, P., Beck, I.A., Boltz, V.F., Coffin, J.M., *et al.* (2006). Blinded, multicenter comparison of methods to detect a drug-resistant mutant of human immunodeficiency virus type 1 at low frequency. *Journal of Clinical Microbiology* **44**:2612–2614.

Hecht, F. M., Busch, M.P., Rawal, B., Webb, M., Rosenberg, E., Swanson, M., *et al.* (2002). Use of laboratory tests and clinical symptoms for identification of primary HIV infection. *AIDS* **16**:1119-1129.

Hirsch, M. S., Gunthard, H. F., Schapiro. J. M., Brun-Vézinet, F., Clotet, B., Hammer, S. M., *et al.* (2008) Antiretroviral drug resistance testing in adult HIV-1 infection: 2008 recommendations of an International AIDS Society–USA panel. *Clinical Infectious Diseases* **47**:266–285.

Hoffmann, C., Rockstroh, J. K., Kamps, B. S. (.2007). HIV Medicine 2007. *Flying pulishers, Hamburg.*

Hu, D. J., Subbarao, S., Vanichseni, S., Mock, P. A., Ramos, A., Nguyen, L., *et al.* (2005). Frequency of HIV-1 dual subtype infections, including intersubtype superinfections, among injection drug users in Bangkok, Thailand. *AIDS* **19** (3): 303–8.

Huang, H., Chopra, R., Verdine, G. L., Harrison, S. C. (1998). Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* **28**(5394):1669-75.

Innis, M. A., Myambo, K. B., Gelfand, D. H. and Brow, M. A. (1988). DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proceedings of the National Academy of Science USA* **85**:9436-9440.

Jackson, J. B., Becker-Pergola, G., Guay, L. A., Musoke, P., Mracna, M., Fowler, M. G., *et al.* (2000) Identification of the K103N resistance mutation in Ugandan women receiving nevirapine to prevent HIV-1 vertical transmission. *AIDS* **14**:111-5.

Jacobo-Molina, A., Ding, J., Nanni, R. Clark, G. A. D. Jr., Lu. X., Tantillo, C., *et al.* (1993). Crystal structure of Human immunodeficiency virus type 1 reverse transcriptase complexed with double stranded DNA at 3.0 a resolution show bent DNA. *Proceedings of the National Academy of Science USA* **90**(13):6320-4.

Johnson, J. A., Li, J. F., Wei, X., Lipscomb, J., Irlbeck, D., Craig, C., *et al.* (2008). Minority HIV-1 drug resistance mutations are present in antiretroviral treatment-naive populations and associate with reduced treatment efficacy. *PLoS Medicine* **5**:e158.

Joint United Nations Program on HIV/AIDS (2006). Overview of the global AIDS epidemic. *Report on the global AIDS epidemic.*

Jordan, M. R., Kearney, M., Palmer, S., Shao, W., Maldarelli, F., Coakley, E. P *et al.* (2010). Comparison of standard PCR/cloning to single genome sequencing for the analysis of HIV-1 populations. *Journal of Virological Methods* **168** pp 114-120.

Kahn, J. O. and Walker, B. D. (1998). Acute Human Immunodeficiency Virus type 1 infection. *New. England Journal of Medicine.* **331** (1): 33–39.

Kaksonen, A. (2005). Molecular approaches for microbial community analysis. *BioMine* 500329

Kavlick, M. F., Wyvill, K., Yarchoan, R., Mitsuya. (1998). Emergence of multidideoxynucleoside–resistant human immunodeficiency virus type 1 variants, viral sequence variation, and disease progression in patients receiving antiretroviral chemotherapy. *Journal of Infectious Diseases* **177**:1506–13.

Krausslich, H.G (1991). Human immunodeficiency virus Proteinase dimer as component of the viral polyprotein prevents particle assembly and viral infectivity. *Proceedings of the National Academy of Sciences USA* **88**(8):3213-7.

Larder, B. A., Bloor, S., Kemp, S. D., Hertogs, K., Desmet, R. L., Miller, V., *et al* (1999). A family of insertion mutations between codons 67 and 70 of human immunodeficiency virus type 1 reverse transcriptase confer multinucleoside analog resistance. *Antimicrobial Agents and Chemotherapy* **43**:1961–7.

Larder, B.A., A. Kohli, S.D. Kemp, M.N. Kro-nick, and R.D. Henfrey. (1993). Quantitative detection of HIV-1 drug resistance mutations by automated DNA sequencing. *Nature* **365**:671-673.

Leitner, T., E. Halapi, G. Scarlatti, P. Rossi, J. Albert, E.M. Fenyo, and M. Uhlen. (1993). Analysis of heterogeneous viral populations by direct DNA sequencing. *BioTechniques* **15**:120-127.

Lennerstrand, J., Hertogs, K., Stammers, D. K., Larder, B. A. (2001). Correlation between viral resistance to zidovudine and resistance at the reverse transcriptase level for a panel of human immunodeficiency virus type1 mutants. *Journal of Virology* **75**:7202–5.

Lifson, A. R. (1988). Do alternate modes for transmission of human immunodeficiency virus exist? *A review. Journal of the American Medical Association* **259** (9): 1353–6.

Liu, R., Paxton, W. A., Choe, S., Ceradini, D., Martin, S. R., Horuk, R., *et al.* (1996). Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiple-exposed individuals to HIV-1 Infections. *Cell* **86**(3):367-77.

Louvel, S., Battegay, M., Vernazza, P., Bregenzer, T., Klimkait, T., Hamy, F. (2008). Detection of drug resistant HIV minorities in clinical specimens and therapy failure. *HIV Medicine* **9**:133–141.

Mansky, L. M and Temin, H. M. (1995). Lower in vivo mutation rate of human immunodeficiency virus type-1 than that predicted from the fidelity of purified reverse transcriptase. *Journal of Virology* **69**(8):5087-94.

Martinez-Picado, J., DePasquale, M. P., Kartsonis, N., Hanna, G. J., Wong, J., Finzi, D., *et al.* (2000). Antiretroviral resistance during successful therapy of human immunodeficiency virus type 1 infection. *Proceedings of the National Academy of Sciences USA*. **97** (20): 10948–10953.

Mbisa, J.L., Gupta, R. K., Kabamba, D., Mulenga, V., Kalumba, M., Chintu, C., *et al.* (2011). The evaluation of HIV-1 Reverse transcriptase in route to acquisition of Q151M multi drug resistance is a complex and involves mutation in multiple domains. *Retrovirology*. **8**(1): 31

Meyerhans, A., Cheynier, R., Albert, J,. Seth, M., Kwok, S., Sninsky, J., *et al.* (1989). Temporal fluctuations in HIV Quasispecies in vivo are not reflected by sequential HIV isolations. *Cell* **58**(5) 901-10.

Michael, N. L., Nelson, J. A., Kewalramani, V. N., Chang, G., O'Brien, S. J., Mascola, J. R., *et al.* (1998). Exclusive and persistent use of the entry coreceptor CXCR4 by human immunodeficiency virus type 1 from a subject homozygous for CCR5 delta32. *Journal of Virology* 72: 6040-47.

Miller, V. (2001). Resistance to protease inhibitors. *Journal of Acquired Immune Deficiency Syndromes* **26**(1):34-50.

Mitsuya, Y., Varghese, v., Wang, C., Liu, T F., Holmes, S. P., Jayakumar, P *et al.* (2008). *Journal of Virology* **82**(21):10747-10755.

Montagnier, L. (2002). A history of HIV discovery. *Science* **298** (5599): 1727-1728.

Morgan, D., Mahe, C., Mayanja, B., Whitworth, J. A (2002). Progression to symptomatic disease in people infected with HIV-1 in rural Uganda: prospective cohort study. *British Medical Journal* **324** (7331): 193–196.

Newton, C. R. and Graham. A. (1994). PCR, part 1: Basic principles and methods. *EngBios Scientific Publishers*, Oxford **pp** 23-28.

Newton, C.R. and Graham, A. (1994). PCR. BIOS Scientific Publishers, Limited, Oxford.

Niederhauser, C., Höfelein, C., Wegmüller, *et al* (1994). Reliability of PCR Decontamination Systems. *PCR Methods and Applications* **4**,117–123.

Nieuwkerk, P., Sprangers, M., Burger, D., Hoetelmans, R. M., Hugen, P. W., Danner, S. A *et al* and the ATHENA Project. (2001). Limited Patient Adherence to Highly Active Antiretroviral Therapy for HIV-1 Infection in an Observational Cohort Study. *Archives of Internal Medince*. **161** (16): 1962–1968.

Nissley, D. V., Halvas, E. K., Hoppman, N. L., Garfinkel, D. J., Mellors, J. W., Strathern, J. N. (2005). Sensitive phenotypic detection of minor drug-resistant human immunodeficiency virus type 1 reverse transcriptase variants. *Journal of Clinical Microbiology* **43**:5696–5704.

Nwobegahay, J. M., Bessong, P. O., Masebe, T. M., Mavhandu, L. G., Iweriebor, B. C., Selabe, G. (2011). Prevalence of antiretroviral drug resistance mutations and HIV-1 subtypes among newly diagnosed drug-naïve persons visiting a voluntary testing and counseling center in northeastern South Africa. *Journal of Health, population and Nutrition* **29** (4): 303-309.

Osmanov, S., Pattou, C., walker, N., Schwardländer, B., Esparza, J., and WHO-UNAIDS Network for HIV isolation and Charaterization. (2002). Estimated global distribution and regional spread of HIV-1 genetic subytpes in the year 2000. *Acquired Immune deficiency Syndrome* **29**(2):184-90.

Palella, F. J., Delaney, K. M., Moorman, A. C., Loveless, M. O., Fuhrer, J., Satten, G. A., *et al.* (1998). Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. *New England Journal of Medicine.* **338** (13): 853–860.

Palella, F. J., Delaney, K. M., Moorman, A. C., Loveless, M. O., Fuhrer, J., Satten, G. A., *et al.* (1997). The qualitative nature of the primary immune response to HIV infection is a prognosticator of disease progression independent of the initial level of plasma viremia. *Proceedings of the National Academy of Sciences USA.* **94** (1): 254–258.

Petropoulos, C. J., Parkin, N. T., Limoli, K. L., Lie, Y. S., Wrin, T., Huang, W., et al. (2000). A novel phenotypic drug susceptibility assay for human immunodeficiency virus type1. *Antimicrobial Agents Chemotherapy* **44**: 920–928.

Piatak, M., Jr, Saag, M. S., Yang, L. C., Clark, S. J., Kappes, J. C., Luk, K. C., *et al.* (1993). High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. *Science* **259** (5102): 1749–1754.

Plantier, J.C. (2009) A new human immunodeficiency virus derived from gorillas. *Nature Medicine* **15**, 871 – 872.

Pollard, V. W. and Malim, M. H. (1998). The HIV-1 Rev protein. *Annual Review of Microbiology.* **52**: 491–532.

Ren, J., Esnouf, R. M. and Hopkins, A, L. (1998). 3-Azido-3-deoxythymidine drug resistance mutations in HIV-1 reverse transcriptase can induce long range conformational changes. *Proceedings of the National Academy of Sciences USA*; **95**:9518–23.

Robb, M. L. (2008). Failure of the Merck HIV vaccine: an uncertain step forward. *Lancet* **372** (9653): 1857–1858.

Roos, M.T., Lange, J.M., De Goed, R.E., Coutinho, R. A., Schellekens, P. T., Miedema. F. J. *et al.* (1992). Viral phenotype and immune response in primary Human immunodeficiency virus type-1 infection. *Journal of Infectious Diseases* **165**(3):427-32.

Roux, K.H. (1995). Optimization and troubleshooting in PCR. *PCR Methods and Applications* **4**:185-194.

Saiki, R.K., D.H. Gelfand, S. Stoffel, S.J. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis and H.A. Erlich. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**:487-491.

Saitoh, A., Hull, A. D., Franklin, P. and Spector, S. A. (2005). Myelomeningocele in an infant with intrauterine exposure to efavirenz. *Journal of Perinatology* **25** (8): 555–556.

Sarafianos, S. G., Das, K., Tantillo, C., Clark, A. D. Jr., Ding, J., Whitcomb, J. M., *et al.* (2001). Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA: DNA. *European Molecular Biology Organization journals* **20**(6):1449-61.

Savarino, A. (2006). A historical sketch of the discovery and development of HIV-1 integrase inhibitors. *Expert Opinion on Investigational Drugs*. **15** (12):1507-22.

Schackman, B. R, Gebo, K. A, Walensky, R. P, Losina, E., Muccio, T., Sax, P. E *et al* (2006). The lifetime cost of current HIV care in the United States. *Medical Care* **44** (11): 990–997.

Servias, J., Lambert, C., Fontaine, E., Plesseria, J. M., Robert, I., Arendt, V., *et al.* (2001). Comparison of DNA sequencing and a line probe assay for the detection of human immunodeficiency virus type 1 drug resistance mutations in patients failing highly active antiretroviral therapy. *Journal of clinical Microbiology*. **39** (2): 454-9

Shafer, R. W., Gonzales, M. J., Brun-Vezinet, F. (2001). Online comparison of HIV-1 drug resistance algorithms identifies rates and causes of discordant interpretations. *Antiviral Therapy* **6**(1):101.

Shafer, R. W. (2002). Assays for Antiretroviral Therapy resistance. *HIV insite knowledge, Base chapter.*

Shirasaka, T., Kavlick, M. F., Ueno, T. (1995). Emergence of human immunodeficiency virus type-1 variants with resistance to multiple dideoxynucleosides in patients receiving therapy with dideoxynucleosides. *Proceedings of the National Academy of Sciences USA*; **92**:2398–402.

Siegfried, N., Muller, M., Deeks, J., Volmink, J., Egger, M., Low, N., *et al.* (2005). HIV and male circumcision-a systematic review with assessment of the quality of studies. *Lancet Infect. Dis.* **5** (3): 165–73.

Simmons, G., Wilkinson, D., Reeves, J.D., Dittmar, M. T., Beddows, S., Weber, J., *et al.* (1996). Primary, syncytium-iducing human immunodeficiency virus type-1 isolates are dual-tropic and most can use either Lestr or CCR5 as coreceptors for virus entry. *Journal of Virology* **70**(12):8533-60.

Stanford HIV drug resistance database (www.hivdb.stanford.edu).

Su, Z., Gulick, R. M., Krambrink, A., Coakley, E., Hughes, M. D., Han, D., *et al.* (2009). Response to vicriviroc in treatment-experienced subjects, as determined by an enhanced-sensitivity coreceptor tropism assay: reanalysis of AIDS clinical trials group A5211. *Journal of Infectious Diseases* **200**:1724–1728.

Suggs, S.V., Hirose, T., Miyake, E.H., Kawashima, M.J., Johnson, K.I., and Wallace, R.B. *et al.* (1981). Using Purified Genes. In *ICN-UCLA Symp. Developmental Biology*, Vol. 23, Brown, D.D. Ed., *Academic Press, New York*, 683.

UNAIDS. (2010). *Report on the Global AIDS Epidemic.USA* **86**(8):2525-9.

UNAIDS (2012). Core Slides: Global Summary of the AIDS Epidemic.

Vallari, A., Holzmayer, V., Harris, B et al. (2011). Confirmation of putative HIV-1 group P Cameroon. *Journal of Virology.* **85**:1403-7.

Wainberg, M. A and Brenner. B. G. (2012). The impact of HIV genetic polymorphisms and subtype differences on the occurrences of resistance to anitiretroviral drugs. *Molecular Biology International*. 256982-10

Wang, C., Vlahov, D., Galai, N., Bareta, J., Strathdee, S. A.,. Nelson, k. E., *et al.* (2004). Mortality in HIV-seropositive versus seronegative persons in the era of highly active antiretroviral therapy. *Journal of Infectious Disease*. **190**: 1046–54.

Wang, J. Y., Ling, H., Yang, W., Craigie, R. (2001). Structure of a two- domain fragment of HIV-1 integrase implication for domain organisation in the intact protein. *European Molecular Biology Organization* **20**:7333-43.

Watts, J. M., Dang, K. K., Gorelick, R. J., Leonard, C. W., Bess, J. W Jr, Swanstrom, R., *et al.* (2009). Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature* **460**:711-716.

Wei, P., Garber, M. E., Fang, S. M., Fischer, W. H., Jones, K. A. (1998). A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. *Cell* **92**: 451-62.

Weiss, H. A., Quigley, M.A., Hayes, R.J. (2000). Male circumcision and risk of HIV infection in sub-Saharan Africa: A systematic review and meta-analysis. *AIDS* **14** (15): 2361–70.

Whitcomb, J. M., Huang, W., Fransen, S., Limoli, K., Toma, J., Wrin T,. *et al.* (2007). Development and characterization of a novel single-cycle recombinant-virus assay to determine human immunodeficiency virus type 1 coreceptor tropism. *Antimicrobial Agents Chemotherapy* **51**:566–575.

World Health Organization. (2007). Male circumcision: Global trends and determinants of prevalence, safety and acceptability.

World Health Organization. (2007). New data on male circumcision and HIV prevention: policy and programme implications.

World Health Organization. (2007). WHO and UNAIDS announce recommendations from expert consultation on male circumcision for HIV prevention.

Winters, M. A., Coolley, K. L., Girard, Y. A., Levee, D. J., Hamdan, H., Shafer R.W., *et al.* (1998). 6-Basepair insert in the reverse transcriptase gene of human immunodeficiency virus type 1 confers resistance to multiple nucleoside inhibitors. *Journal of Clinical Investigation* **102**:1769–75.

Wong-Staal, F. (1991). HIVes and their replication. In: Fundamental Virology, Ed.: Fields, B. N., Knipe, D. M *et al. Raven Press, Ltd., New York.*

Wyatt, R and Sodroski, J (1998). The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science* **280** (5371): 1884–8.

Zheng, Y. H., Lovsin, N. and Peterlin, B. M. (2005). Newly identified host factors modulate HIV replication. *Immunology Letters* **97** (2): 225–34.

Zhu, T., Mo, H., Wang, N *et al.* (1993). Genotype and phenotype characterization of HIV-1 patients with primary infection. *Science* **261**(5125):1170-81.