

MIDLANDS STATE UNIVERSITY



**MOLECULAR CHARACTERISATION OF GP41 AND GP120 V3 LOOP IN HIV-1C
INFECTED PATIENTS FAILING SALVAGE THERAPY IN BOTSWANA**

By

Nokuthula Sibusiso Ndlovu

R147337X

**A thesis submitted in partial fulfillment of the requirements for the Bachelor of
Science Honors Degree in Biosciences and Biotechnology**

**Department of Biosciences and Biotechnology
Faculty of science and Technology**

December 2018

RELEASE FORM

NAME OF AUTHOR: Nokuthula Sibusiso Ndlovu
REG NUMBER: R147337X
DEGREE PROGRAM: Bsc in Applied Biological Sciences and Biotechnology
PROJECT TITLE: **Molecular Characterisation of gp41 and gp120V3
Loop in HIV-1C Infected Patients Failing Salvage
Therapy in Botswana**
SUPERVISOR: DR F. Songwe

YEAR GRANTED: 2018

Permission is hereby granted to the Midlands State University library to produce single copies of this dissertation and to lend or sell such copies for scholarly or scientific research only. The author reserves other publication rights. Neither the dissertation nor extensive extracts from it may be printed or otherwise reproduced without the author's permission.

Signed:

Address: 8444 Nkulumane
Bulawayo
Phone 0783496753
Email Address: nokkiethula1@gmail.com
Date: 4 February 2019

APPROVAL FORM

This is to certify that that the dissertation entitled “**Molecular characterization of gp41 and gp120 V3 loop in HIV-1C infected patients failing salvage therapy in Botswana**”, submitted in partial fulfillment of the requirements for Bachelor of Science Honors Degree in Biosciences and Biotechnology at Midlands State University, is a record of the original research carried out by Nokuthula Sibusiso Ndlovu R147337X under my supervision and no part of the dissertation has been submitted for any other degree or diploma.

Any assistance received during the course of this research has been duly acknowledged. Therefore, I recommend that it be accepted as fulfilling the dissertation requirements of the Midlands State University.

Name of supervisor

Signature

Chairperson signature

ABSTRACT

Multi-drug resistant (MDR) HIV-1 infection remains a challenge in individuals with extensive antiretroviral treatment (ART) experience, in terms of high mortality and probability of onward transmission. New therapeutic options within old drug and new drug classes are therefore essential. This study sought to determine coreceptor usage and T20 resistance associated mutations in patients failing salvage therapy from clinics around Botswana, through characterizing gp120 V3 loop and gp41 of the viral envelope. Thirty-eight blood samples from deep salvage patients were included in this analysis. Gp41 and gp120 V3 regions of HIV-1 envelope were amplified. Drug resistance mutations were analyzed according to the IAS-USA 2017 reference mutation lists. Coreceptor usage was determined using PSSM and Geno2Pheno using a false positive rate (FPR) of 10%. Among 38 participants, 34(89%) GP41 sequences and 25(68%) gp120 V3 loop sequences were obtained. Major T20 resistance mutation G36S and minor mutations E137K, and S138A were observed at a 2.9% frequency each, occurring in different sequences. Polymorphisms N42S(70.6%), L54M(61.8%), A67T(32.4%), I69V(97.1%), V72I(61.8%), E119Q(88.2%), N125S(44.1%), S129D(32.4%), S129N(29.4%), L130T(67.7%), H132Y(97.1%), S133R(44.1%), I135L(100%), Q147K(61.8%), E148D(91.2%), E151A(70.6%), K154S(55.9%), A156N(38.2) and S157N(61.8%) were detected in Heptad repeat 1 (HR1) and heptad repeat 2 (HR2) of gp41. CXCR4 coreceptor associated use, mutation L34M in gp41 HR1 was detected in 2 samples (5%). Analysis of coreceptor usage showed (17/25) 64% use of CCR5, and a (9/25) 36% use of the CXCR4 coreceptor. A moderately high proportion of treatment experienced (salvage) participants had CXCR4 coreceptor using strains (X4 variants). X4 variants are associated with disease progression and ultimately death in absence of treatment. The use of maraviroc in Botswana would require coreceptor tropism testing. Non T20 treatment experience in Botswana reduces the prevalence of the major mutations that confer resistance to the drug. T20 is therefore a potential alternative drug for patients failing salvage therapy in Botswana.

Keywords

Multi drug resistance (MDR); gp41; gp120; Salvage therapy; HIV entry inhibitors; CCR5 antagonists; fusion inhibitors; maraviroc (MVC); enfuvirtide (T20)

ACKNOWLEDGEMENTS

My great acknowledgements to Dr Gaseitsiwe, Dr Moyo, Dr Seatla, Dorcas Maruapula, the HIV drug resistance group, and not leaving out anyone at Botswana Harvard Partnership for their guidance and mentorship during the research work. To my supervisor Dr Songwe, co-supervisor Mr Dowo I am grateful for the guidance during the study. My family especially mom (Mrs D Ndlovu) I am grateful for the love, support, encouragement my entire life and throughout the study. Last but not least to the Almighty God for the love, grace and favor, words are not enough to express my gratitude.

DEDICATIONS

This thesis is dedicated to my mom (Mrs. D. Ndlovu) and my babies, Tyrone Shannon, Tyra Sheena and Tamar Shylet. You inspire me to do better each day.

Table of Contents

CHAPTER 1: INTRODUCTION	1
1.1 Background of the study	1
1.4 Objectives.....	5
1.4.1 Main objective	5
1.4.2 Specific Objectives	5
CHAPTER 2: LITERATURE REVIEW	6
2.1 Molecular Epidemiology	6
2.2 Virology of HIV	7
2.3 Transmission	9
2.5 HIV life cycle and drug interactions	10
2.6 Drug resistance, a challenge in the management of HIV.....	12
CHAPTER 3: MATERIALS AND METHODS	14
3.2 Study Design and Study Population	14
3.3 Sample collection and storage	14
3.4 Materials	15
3.5 Methods.....	15
3.5.1 Determination of coreceptor usage, naturally occurring gp41 Polymorphisms and T20 associated mutations.....	15
3.5.1.1 Isolation of Viral DNA from buffy Coat Samples of the 38 HIV-1C Infected Patients.....	15
3.5.1.2 Amplification of the Isolated viral genome Using Nested PCR	15
3.5.1.3 Quantification of Amplified Viral DNA	16
3.5.1.4 PCR Product Purification.....	17
3.5.1.5 Sequencing Cycle PCR	18
3.5.1.6 Sequence Clean Up	18
3.5.1.7 Sequencing and Sequence base editing.....	19
3.5.1.8 Multiple Sequence Alignment, Sequence Editing and Phylogenetic Analysis	19
3.5.1.9 HIV subtype verification.....	19
3.5.2 Characterization of gp41.....	20
3.5.2.1 Analysis for Naturally Occurring Polymorphisms and T20 Resistance Associated Mutations....	20
3.5.3 Characterization of gp120 V3 Loop.....	20
3.5.3.1 Prediction of Coreceptor Usage.....	20
3.6 Determining the Associations of Age, Gender, viral load, and CD4 count with Coreceptor Usage .	21

CHAPTER 4: RESULTS	22
4.1 Determining Coreceptor Usage and naturally occurring gp41 Polymorphisms and T20 associated mutations.....	22
4.1.1 Viral isolation, amplification and quantification.....	22
4.1.1.1 Quantification of gp120 amplicons.....	22
4.1.1.2 Quantification of gp41 amplicons.....	23
4.1.2 Multiple sequence alignment (MSA) of the HIV-1 sequences.....	23
4.1.3 Phylogenetic analysis of the HIV-1 sequences.....	24
4.1.4 HIV-1 subtype verification using NCBI HIV and REGA subtyping tools.....	25
4.2 Characterization of gp41.....	26
4.2.1 Mutation Analysis for Naturally Occurring gp41 Polymorphisms and T20 Resistance Associated Mutations.....	26
4.3 Characterization of gp120 V3 loop.....	27
4.3.1 Coreceptor analysis using Geno2Pheno(g2p) and PSSM.....	27
4.3.2 Coreceptor analysis by the 11/24/25 Rule.....	28
4.3.3 Coreceptor Analysis Using the Insertion/ Deletion Rule.....	28
4.4 Effects of patients demographic (age and gender) and clinical parameters (viral load and CD4 count) on Coreceptor Use.....	29
4.4.1 Patients Demographics and Data.....	29
4.4.2: Gender Distribution of Coreceptor Used.....	31
4.4.3: Assessing the Median Difference Between CCR5 and CXCR4 on Age, Viral Load and CD4 Count.....	32
CHAPTER 5: DISCUSSION	32
5.2 RECOMMENDATIONS.....	36
5.3 CONCLUSIONS.....	37
APPENDICES	38
Appendix 1 Reagents, Buffers and Solutions.....	38
Appendix 2: Kits.....	39
Appendix 3: Equipment and Consumables.....	40
Appendix 4: List of sample numbers and PID's.....	41
Appendix 5: Gp41 chromatograms.....	41
Appendix 6: Gp 120 V3 loop chromatograms.....	Error! Bookmark not defined.
REFERENCES	45

List of Tables

Table 1.1A: Current treatment guidelines for pediatrics in Botswana.....	2
Table 1.1B: Current treatment guidelines for adults and adolescents (Including pregnant women) in Botswana.....	3
Table 3.4.1: Reagents, buffers and solutions.....	15
Table 3.4.2: Kits.....	16
Table 3.4.3: Equipment and consumables.....	16
Table 4.1: Patients demographics and data.....	31
Table 4.2: CCR5 and CXCR4 specific mean, median of age, viral load and CD4 counts.....	32
Table 4.3: Mann Whitney test results.....	33

List of Figures

Fig 2.1: HIV-1 envelop glycoprotein structure domains.....	8
Fig 2.2: HIV life cycle and drug interactions.....	11
Fig 4.1: Amplified gp41 results.....	23
Fig 4.2; Amplified gp120 results.....	24
Fig 4.3: Multiple sequence alignment of gp41 sequences against HxB2.....	25
Fig 4.4: Multiple sequence alignment of gp120 sequences against HxB2.....	25
Fig 4.5: Phylogenetic tree built based on gp41 sequences.....	26
Fig 4.6: Phylogenetic tree built based on gp120 sequences.....	27
Fig 4.7: Frequency of major and minor T20 resistance associated mutations.....	28
Fig 4.8: Frequency of gp41 polymorphisms.....	29
Fig 4.9: Coreceptor usage results.....	30
Fig 4.10: Gender distribution of coreceptor use.....	33

List of Abbreviations

- AIDS – Acquired immunodeficiency syndrome
- ARV - Antiretroviral
- BHHRL - Botswana Harvard HIV Reference Laboratory
- BHP - Botswana Harvard Partnership
- CCR5 – Chemokine receptor 5
- Cdna – Complementary deoxyribonucleic acid
- CRF – Circulating recombinant forms
- CT – Cytoplasmic tail
- CXCR4 – Chemokine receptor four
- DC-SIGN – Dendritic cell specific intercellular adhesion molecule-3-grabbing non-integrin
- DNA – Deoxyribonucleic acid
- DRAM – Drug resistance associated mutations
- DTG - Dolutegravir
- FDA – Food and drug administration
- FPR - False positive rate
- Gp41 – Glycoprotein 41
- Gp120 – Glycoprotein 120

HAART – Highly active antiretroviral therapy
HIV – Human immunodeficiency virus
HR1 – Heptad repeat 1
HR2 – Heptad repeat 2
HxB2 – HIV 1B reference sequence
IAS – International Aids Society
IDCC PMH – Infectious Disease Control Centre Princess Marina Hospital
INSTI – Integrase strand transfer inhibitor
Ks – Kennedy sequence
MDR – Multi drug resistant
MVC – Maraviroc
mLC – Mucosal Langerhans cells
MPER – Membrane proximal
NNRTS – Non-nucleoside reverse transcriptase inhibitors
NRTS – Nucleoside reverse transcriptase inhibitor
OI – Opportunistic infections
PI – Protease inhibitors
PID – Patient identification number
PSSM – Position specific scoring matrices
RNA – Ribonucleic acid
R5 – Chemokine receptor 5 using variant
RT PCR – Reverse transcriptase polymerase chain reaction
T20 – Enfuvirtide
TMD – Transmembrane domain
VL – Viral load
X4 – Chemokine receptor 4 using variant

List of Appendices

Appendix 1: Table of reagents, buffers and solutions.....	38
Appendix 2: Table of Kits.....	39
Appendix 3: Table of equipment and consumables.....	40
Appendix 4: List of sample numbers and PID's.....	41
Appendix 5: Gp41 chromatograms.....	42
Appendix 6: Gp120 V3 loop chromatograms.....	43

CHAPTER 1: INTRODUCTION

1.1 Background of the study

Botswana is the second highly HIV burdened country globally (Bussmann *et al.*, 2005; Farahani *et al.*, 2014; Gaolathe *et al.*, 2016) with a reported prevalence rate of 25.2% in adults (15-49 years) in 2016 (Gaolathe *et al.* 2016). Botswana has a generalized HIV epidemic (HIV infections are not restricted to a specific groups of people) (Marukutira *et al.* 2018), with HIV-1 subtype C (HIV-1C) being the most prevalent HIV infection (Bussmann *et al.* 2005). A total of 380 000 individuals were living with HIV in 2016, with 14 000 being new infections. Overall a total of 4 100 individuals died due to HIV related infections and there were 61 000 orphans due to the pandemic (UNAIDS 2016). Marukutira *et al.* (2018) states that the current prevalence of HIV in Botswana is at 18.5% with a 1.34% annual incidence. In the verge of fighting the high prevalence of HIV, Botswana became one of the first African countries to initiate the national HIV/AIDS treatment program in 2002 (Farahani *et al.*, 2014; Gaolathe *et al.*, 2017), and to introduce Integrase Strand Transfer Inhibitors (INST) to its arsenal of Antiretroviral (ARV) regimens which is currently being used as part of first line Highly Active Antiretroviral Therapy (HAART) salvage therapy, and also adopted the test and treat policy in 2016 (Marukutira *et al.*, 2018; Botswana MoH, 2016).

Treatment guidelines currently in use in Botswana are inclusive of combinations of Nucleoside Reverse Transcriptase Inhibitors (NRTI's), Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTI's), Integrase Strand Transfer Inhibitors (INST's), and Protease Inhibitors (PI's) for the management of HIV infections in pediatrics and adults (Table 1.1). Treatment combinations for second line treatment are based on drug resistance test results of the patients (Botswana MoH 2016) which are recommended for guiding the selection of active drugs (Imaz,

Falc and Ribera, 2011). When patients develop multi drug resistance (MDR) and do not respond to standard therapy for a long time, treatment options become limited, and they need to be put on salvage therapy (also known as rescue therapy/ last resort therapy)

http://www.aidsinfonet.org/fact_sheets/view/408, <https://aidsinfo.nih.gov/understanding-hiv-aids/glossary/643/salvage-therapy>, <https://aac.org/aachealth-library/salvage-treatment-for-hiv/>.

In cases of salvage therapy failure patients require alternative regimens that can suppress viral loads and maintain them for long periods of time (Gantner *et al.* 2016). Development of newer and more potent antiretroviral agents like entry inhibitors (CCR5 antagonist and fusion inhibitors) gives therapeutic options for HIV-1C infected patients failing therapy (Del Rio, 2006; Imaz, Falc and Ribera, 2011).

Table 1.1 A and B: Botswana treatment guidelines for infants and adults used since 2016

Table 1.1 A: Current treatment guidelines for pediatrics in Botswana

Age/weight	1 st Line	1 st Line Modifications for toxicities	2 nd Line
< 1month results	AZT/ ₃ TC/NVP (Switch infants to ABC/ ₃ TC/ALU at 3months).	<u>ABC or NVP Rash</u> CBV/ALU	Based on resistance testing and consultation with HIV specialist
> 1 month to 3years specialist	ABC/ ₃ TC/ALU	CBV/ALU	Based on resistance testing results and consultation with HIV specialist
> 3 years (<40kg) specialist	ABC/ ₃ TC/ENF	<u>ABC Rash</u> CBV/EFV <u>CNS Toxicities</u>	Based on resistance testing results and consultation with HIV specialist

Table 1.1B: Current treatment guidelines for adults and adolescents (Including pregnant women) in Botswana

Age/Weight	1 st Line	1 st Line Modifications	2 nd Line
All adults and adolescents (> 40kg)	Truvada and Dolutegravir	TDF renal toxicity w/o CDV risk: ABC ₃ TC/DTG TDF renal toxicity and insufficiency with CDV risk or DTG toxicity: Discuss with specialist	Based on resistance testing results and consultation with HIV specialist

Source: Botswana MoH, (2016)

Entry inhibitors, Maraviroc (MVC), a Chemokine receptor 5 (CCR5) coreceptor antagonist and Enfuvirtide (T20, Fuzeon), a fusion inhibitor have recently been approved by the Food and Drug Administration (FDA) and have offered clinicians ability to design regimens with three fully active drugs for multi class drug resistant HIV-1 infections (Cortez and Maldarelli 2011). CCR5 antagonists prevent entry of HIV into the host cell by binding to and stabilizing the CCR5 coreceptor in a conformation that prevents it from interacting with gp120 V3 loop (Cortez and Maldarelli, 2011; Recordon-Pinson *et al.*, 2010). Mutations on the crown of the gp120V3 loop, whose occurrence is associated with disease progression, gives the HIV virus

ability to switch from using the CCR5 coreceptor which is used by the virus at initial infection to using the Chemokine receptor 4 (CXCR4) coreceptor, reducing the potential of MVC as alternative therapy (Dimonte *et al.*, 2011; Lin *et al.*, 2011; Shen *et al.*, 2016; Zhang *et al.*, 2010). Fusion inhibitors competitively bind to the HR1 domain of gp41 preventing the formation of a six-helix bundle required for the fusion process thus preventing fusion of the host cell and viral membrane (Qian, Morris-Natschke and Lee, 2009; Ismael *et al.*, 2014), however, naturally occurring Enfuvirtide (T20) resistance associated mutations would reduce the efficacy of HIV-1 inhibition by T20 (Ismael *et al.* 2014). Polymorphism variances within the Heptad repeat 1 (HR1) and Heptad repeat 2 (HR2) domains of gp41 have also previously been reported in Mozambique (Ismael *et al.* 2014) and Korea (Jang *et al.* 2014).

1.2 Problem Statement

Previous studies have demonstrated that specific mutations in gp41 and gp120 V3 of MDR HIV variants reduce efficacy of T20 and MVC respectively, in the inhibition of HIV (Aquaro *et al.*, 2006; Pessoa *et al.*, 2011; Svicher *et al.*, 2014; Jang *et al.*, 2014). The understanding of the HIV envelop interactions with the host cell coreceptors is mostly based on HIV-1 subtype B (HIV-1B) studies though a majority of infections worldwide, including in Botswana are due to HIV-1 subtype C (HIV-1C) (Lin *et al.*, 2012).

1.3 Justification

To our knowledge, no study has been done to determine if multi class drug resistant HIV-1C variants infecting highly disease progressed patients not responsive to salvage therapy exhibit naturally occurring gp41 and gp120 V3 loop mutations conferring resistance to entry inhibitors. This research was aimed at determining coreceptor usage by multi class drug resistant

strains, identifying naturally occurring gp41 and gp120 V3 loop mutations associated with resistance to entry inhibitors, and gp41 polymorphism variations of heptad repeat 1 and 2 domains in MDR HIV-1C isolates from salvage therapy failing patients. We therefore characterized gp41 and gp120 V3 loop of the viral envelop for mutations that confer resistance to entry inhibitors, identified polymorphism variances of HR1 and HR2 of gp41 and determined the associations of patient demographics and clinical parameters (age, gender, viral load, and CD4 counts) on coreceptor usage.

1.4 Objectives

1.4.1 Main objective

- To characterise the gp41 and gp120 V3 loop of HIV 1C multi-drug resistant strains in salvage patients failing combination Dolutegravir (cDTG).

1.4.2 Specific Objectives

- To determine HIV-1C viral coreceptor usage in treatment experienced patients in Botswana.
- To identify gp41 and gp120 V3 loop mutations associated with resistance to T20 and MVC respectively.
- To determine the associations of age, gender, viral load, and CD4 counts on Coreceptor Usage.

CHAPTER 2: LITERATURE REVIEW

2.1 The burden of HIV

Acquired immunodeficiency syndrome (AIDS) was discovered in 1981 when an increasing number of young homosexual men presented unusual opportunistic infections and a rare malignancy (Sharp and Hahn, 2011; Weiss 2013). Human immunodeficiency virus (HIV), the causative agent of AIDS is said to have originated from non-human primates in Sub-Saharan Africa as Simian immunodeficiency virus (SIV) and evolved into HIV as it was transmitted from primates to humans (Sharp and Hahn, 2011). A total of 76.1 million (65.2million - 88 million) have been infected with HIV since its discovery (Africa and Africa, 2016; Weiss, 2013), and 35 million (28.9 million - 41.5 million) people have since died of HIV related illnesses. In the year 2016 alone 36.7 million (34.5m adults, 2.1m children <15 years, and 17.8m women 15+ years) people were living with HIV with 1.8 million new infections (UNAIDS., 2016), and a record of 1 million people died from HIV related illnesses (UNAIDS.,2017). Sub-Saharan Africa is responsible for 43% of new infections globally and carries the highest HIV burden with 19.4 million people in the region infected. Botswana is the second highly HIV infected country globally (Gaolathe *et al.* 2016) after Swaziland (UNICEF 2013).

2.1 Molecular Epidemiology

HIV type 1 (HIV-1) is the most common in humans (Sharp and Hahn, 2011). HIV-1 is divided into four quite distinct lineages: M, N, O and P groups. Among the four groups, Group M predominates and is largely responsible for the HIV-1 epidemic in most of the world. Group O is largely confined to Africa (Hartley, 2005) whilst group P has also been found only in

Africa. Group M has been reported to evolve into distinct phylogenetic subtypes A1, A2, B, C, D, F1, F2, G, H, J, K and also some circulating recombinant forms (CRFs) which consist of more than one subtype. Even though there are so many subtypes, the highest diversity of HIV is found in West-Central Africa where it originated. Only subtypes A, B, C, D, CRF01_AE and CRF02_AG have managed to increasingly spread and expand with 90% of the epidemic (Woodman and Williamson, 2009). In West and Central Africa subtype A and CRF_ A/G are the most common. Subtype B is mostly found in Europe, America, Japan and Australia. Subtype C is the most prominent subtype accounting for half of all the infections in the world, it is mostly found in Southern and East Africa, India and Nepal. Subtype D is also found in East and Central Africa (Hoy and Hahn, 2011).

2.2 Virology of HIV

HIV is a retrovirus (virus that carries out reverse transcription), (Hardin *et al.*, 2012) comprising a single stranded positive-sense ribonucleic acid (RNA) genome of about 9.7 kilobases (Hoy and Lewin, 2004). This genome consists of nine genes (*gag*, *env*, *pol*, *tat*, *rev*, *vpu*, *vif*, *vpr* and *nef*) that encode proteins that may be grouped into structural (*gag* and *env*), catalytic (*pol*), regulatory (*tat* and *rev*) and accessory (*vpu*, *vif*, *vpr* and *nef*). The gene products play an important role in the viral life cycle: *env* gene generates an envelope subunit that consists of two non- covalently linked membrane proteins glycoprotein (gp) 120 which is the outer protein, and gp41 which anchors the glycoprotein complex to the surface of the virion and is essential for HIV binding and entry into the host cell (Hoy and Lewin, 2004).

Gp120 is structurally divided into highly variable (V) and constant (C) regions (Figure1). The variable region comprises of the V1, V2, V3, V4 and V5 regions which play major roles in infectivity, transmission and resistance to neutralization (Cenci *et al.*, 2012). Studies on HIV-1B

and HIV-1C have proved that the most variable region is V1 and amongst its roles is to protect the CD4 binding site from neutralizing antibodies (Cenci *et al.*, 2012). The V2 region is less variable than V1 and its role is to bind Dendritic Cell Specific Intercellular adhesion molecule-3-Grabbing Non-integrin (DC-SIGN). DC-SIGN binds to different microorganisms by recognizing high-mannose-containing glycoproteins on their envelopes and most commonly those that functions as receptors for viruses like HIV. Binding to DC-SIGN is an important process for HIV infection as the binding can promote HIV to infect T-cells from dendritic cells (Geijitenbeek *et al.*, 2000). The V3 region loop has been shown to bind to coreceptor CCR5 to initiate infection and can acquire specific mutation that allow the use of the CXCR4 coreceptor (Cenci *et al.*, 2012; Le *et al.*, 2015; Zhang *et al.*, 2010). The roles of V4 and V5 have not been clearly identified but it is thought that they could play a role in envelope conformation and glycan packing (Cenci *et al.*, 2012)

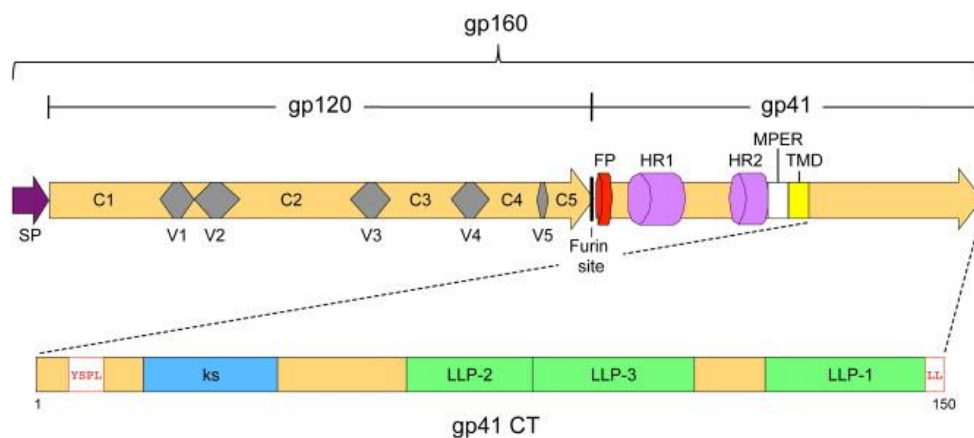


Figure 2.1: HIV-1 envelope glycoprotein structure domains. Precursor gp160 contains the signal peptide (SP), which is cleaved during translation. The remaining precursor is cleaved into the surface subunit (gp120) and transmembrane subunit (gp41) in the Golgi complex at the furin site indicated. Gp120 contains five variable domains (V1-V5) and five constant domains (C1-C5). Gp41 consists of an extracellular domain, containing the fusion peptide (FP), heptad-repeats

(HR1 and HR2), and the membrane-proximal external region (MPER), a transmembrane domain (TMD), and a cytoplasmic tail (CT). An enlarged representation of the gp41 CT is shown to highlight several motifs: the internalization signal YSPL, the Kennedy sequence (ks), the amphipathic α -helices LLP-1, -2, -3, and a C-terminal dileucine motif (LL) involved in endocytosis and intracellular distribution of *Env* (Checkely *et al.*, 2011).

2.3 Transmission

HIV-1 spreads by sexual, percutaneous, and perinatal routes, however, 80% of adults acquire HIV-1 following exposure at mucosal surfaces at the genital area, and AIDS is thus primarily a sexually transmitted disease (Sharp and Hahn, 2011). It has been shown that despite a swarm of closely related viruses present in the donor, there is a genetic bottleneck associated with transmission so that only a limited number of variants get transmitted to the recipient (Hoy and Lewis., 2004). Previous studies have shown that approximately 80% of productive infections were a result of a single virus or a single virus infected cell (Woodman and Williamson, 2009).

2.4 Pathogenesis of HIV-1

The HIV-1 infection causes a slow decline in CD4⁺ T- cells over time. Once a threshold of approximately 200×10^9 CD4 cell/l is passed immune deficiency and virally-induced tumors are increasingly likely to occur (Weber, 2001). HIV-1 has been found to first affect mucosal Langerhans' cells. Mucosa membranes are linings of endodermal origin covered in epithelium which are involved in absorption and secretion. The well understood mechanism that HIV-1 infection uses in mucosal surfaces is whereby the free infectious virions are bound to DC- SIGN on the surfaces of mucosal Langerhans cells (mLC) (Weber, 2001). This occurs by the virion interlocking with the DC-SIGN and migrating within the vaginal mucosa. Then the virus that is bound to mLC moves away from the mucosal surface and comes close to the CD4⁺ T- cell where

the infection will occur. The DC-SIGN bound virus will then infect CD4⁺ /CCR5⁺ T-cells which will move to their regional lymph nodes. The HIV-1 infected T- cells stay in the regional lymph until certain replication amount is reached and the bursting of plasma viremia occurs (Weber, 2001).

The virus will then widely spread out through the body and the level of the viral RNA can be detected by quantitative RT- PCR. This is termed primary infection and is associated with symptoms like fever, rash and painful lymphadenopathy. The HIV-1 infection is further characterized by prolonged asymptomatic period which is associated with slow decline in CD4⁺ T- lymphocytes. There is also an increase in the amount of the virus detected in plasma signifying a direct relationship between viral replication and CD4⁺ T- cell destruction. The late disease stage is whereby the HIV- specific cellular immunity is exhausted and CD4 count is $<200 \times 10^9/L$ (Weber, 2001).

2.5 HIV life cycle and drug interactions

The HIV entry process entails viral proteins gp120 binding to the CD4 and its third variable loop (gp120 V3 loop) further binding to a chemokine coreceptor (CCR4 or CXCR4) of the host cell. After coreceptor binding glycoprotein 41 (gp41) undergoes conformational changes that puts viral and host membranes at juxta position, creating a pore to allow viral capsid to be deposited into the host cell hence initiating infection. (Zhang *et al.*, 2010; Le *et al.*, 2015). Viral RNA is reverse transcribed into cDNA which is transported to the cell's nucleus where it is integrated into host DNA by the enzyme integrase creating a provirus. Viral cDNA is transcribed and translated into viral proteins, including the viral RNA, which are assembled into noninfectious virus that bud off the host. Enzyme protease then cleaves the long protein chains

that form the noninfectious virus and small proteins combine to form infectious viral particles (Nolan *et al.* 2009)

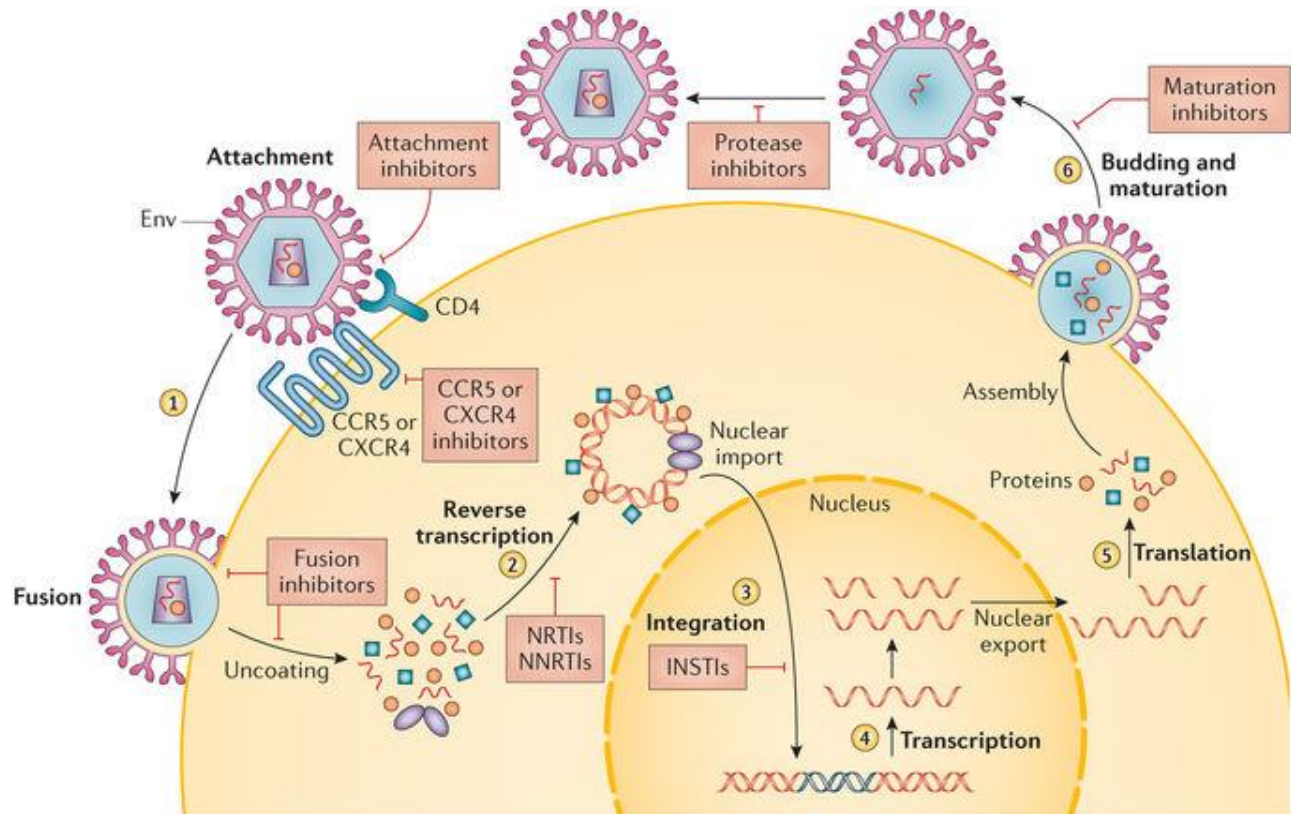


Fig 2.2 Life cycle of HIV and different drug responsible for the inhibition of the HIV virus interactions.

Source: WHO (2017), Tsibris (2007)

The Food and Drug Administration (FDA) approved more than 25 HIV medications from different drug classes: the Nucleoside Reverse Transcriptase Inhibitors (NRTI's), Non-Nucleoside Reverse Transcriptase Inhibitors(NNRTI's), Integrase Strand Transfer Inhibitors (INST's), and Protease Inhibitors (PI's) which suppress viral replication by inhibiting enzymes involved in viral replication and the entry inhibitors (Fusion inhibitors and CCR5 antagonists) (Chesna and Fellner 2017). CCR5 antagonists prevent viral replication by binding the CCR5

coreceptor preventing its interaction with the V3 loop. Fusion inhibitors prevent fusion of the viral and host cell membranes thus preventing infection (Pang, *et al.*, 2009)

2.6 Drug resistance, a challenge in the management of HIV

Drug resistance poses a major challenge in the treatment of HIV as it occurs in a substantial proportion of treated patients and accumulates over time on therapy (Green *et al.*, 2012). Drug-resistant variants are transmitted when new infections occur, effectively multiplying the individual drug failures and creating a growing public health concern. Although its frequency has reduced with the introduction of better tolerated regimens, resistance is still reported in 7–15% of patients initiating first line antiretroviral therapy (Cortez and Maldarelli 2011) and treatment options are limited for highly ART experienced salvage patients harboring viruses with past drug-resistance-associated mutations (DRAMs) who are currently controlled under ineffective treatment (Gantner *et al.* 2016). Exhaustion of therapeutic options results in HIV-1 disease progression and death (Cortez and Maldarelli 2011). Risk of Disease progression is associated with the emergence of viruses that have acquired the ability to use the CXCR4 (X4 variants) coreceptor (Lin *et al.*, 2011; Zhang *et al.*, 2010). X4 variants are more virulent and are associated with an accelerated rate of CD4+ T cell loss and a relatively rapid progression to AIDS and death (Zhang *et al.* 2010).

Previous studies on HIV-1B have shown X4 frequencies of 30-50% (Lin *et al.*, 2011). Lin *et al* (2011) observed a 15% frequency of CXCR4 coreceptor usage in HIV-1B treatment naive patients while Arruda *et al* (2014) observed 21.4%. Green *et al*, (2012) observed X4 or (R5X4)/mixed (R5, X4) (D/M)-tropic viruses 54.3% of HAART-failing and 9.4% of HAART-naive children. HIV-1C is said not to switch coreceptor usage as much as HIV-1B thus variations in the coreceptor usage results are affected by difference in HIV subtypes (Lin *et al.*, 2011).

Previous studies by Ismael *et al* (2014) and Jang *et al* (2014) observed low frequencies of T20 associated mutations and low to high frequencies of polymorphisms of gp41 with the HR2 region showing greater variability. Ismael *et al* (2014) observed the minor T20 associated mutation S138A more prevalent in HIV-1C variants without mutations of HR1. Variations in results obtained by Ismael *et al* (2014) and Jang *et al* (2014) could be because the study by Ismael *et al* (2014) focused on both, treatment experienced and treatment naive patients while Jang *et al* (2014) focused on treatment naive patients. Ismael *et al* (2014) reported a 74% frequency of polymorphism L34M, 13% for L130I, and 31% for S129T, further identifying E148D, E151D, E119E, Q147K and A156N at high frequencies in HIV naive patients which was consistent with Jang *et al* (2014). Jang *et al* (2014) reported polymorphism L34M at a 1.2% frequency which was less than the 74% frequency reported by Ismael *et al* (2014). Variations of the frequencies of the same polymorphisms arise as a consequence of different subtypes and treatment experience. Polymorphism L34M is said to be most frequent in non subtype-B variants (Jang *et al.*, 2014).

CHAPTER 3: MATERIALS AND METHODS

3.1 Study Site

The study was carried out at Botswana Harvard HIV Reference Laboratory (BHHRL), an affiliation of Botswana Harvard AIDS Initiative Partnership (BHP) which is situated at the heart of Princess Marina Hospital, Gaborone, Botswana, corner North Ring and Notwane Road.

3.2 Study Design and Study Population

A cross sectional analysis was done on 38 consecutive, non-selected, highly treatment experienced HIV-1C infected patients failing salvage therapy, who were under clinical monitoring at the Infectious Disease Control Centre Princess Marina Hospital (IDCC PMH). All patients had multiple HAART failures and were currently on cDTG which they were non-responsive to.

3.3 Sample collection and storage

Blood samples were anonymously collected at IDCC PMH during 2015 and also in 2017. Blood sample processing to buffy coat was done at the Botswana Harvard HIV Reference Laboratory (BHHRL) where each buffy coat sample was labelled with the same patient identification number (PID) on the blood sample. Samples were stored at -70°C .

3.4 Materials

Refer to appendices 1, 2 and 3.

3.5 Methods

3.5.1 Determination of coreceptor usage, naturally occurring gp41 Polymorphisms and T20 associated mutations

3.5.1.1 Isolation of Viral DNA from buffy Coat Samples of the 38 HIV-1C Infected Patients

Deoxyribonucleic Acid (DNA) was extracted from 200µl buffy coat using the QIAGEN DNA Blood Extraction Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Extraction was done commercially using the EZI QIAGEN extraction machine.

3.5.1.2 Amplification of the Isolated viral genome Using Nested PCR

Nested PCR was performed for the amplification of isolated HIV genome using the KAPA kit. First round PCR targeted the near full length of genome using the outer primers OFM19 forward primer (For) 5'-CGACCTAAGGCAAGCTTTATTGAGGCTTA-3' and SK145 reverse primer (Rev) Rev 5'-AGTGGGGGGACATCAAGCAGCCAT-3'. Working in a clean room, master mix reagents were thawed while in ice for 20 to 30 minutes, briefly spun before use and kept in ice for the whole process. For 38 samples with a PCR reaction mixture of 25µl for each sample, master mix was prepared by pipetting 473µl dH₂O, 76µl dNTP, 190µl 5X PrimeSTAR GXL Buffer, 19µl 5X PrimeSTAR GXL DNA Polymerase and 19µl of each primer (0.2µM) into an aliquot tube and mixing by flicking the tube, then aliquoting 20µl into each well of the reaction plate. Working under a lamina flow 5µl of template DNA (Isolated DNA) from

each sample was pipetted into the reaction mixture and mixed by pipetting up and down, then PCR reaction plate was covered. PCR was performed with a thermocycler using the following steps: 40 cycles of denaturation at 98°C for 10 seconds, annealing at 62°C for 15 seconds and final elongation at 68°C for 9 minutes, holding at 4°C until reaction plate was taken storage at -20°C.

Second round PCR targeted the gp41 and the gp120 regions separately. For both gp41 and gp120 master mix was prepared as mentioned above. PCR reaction mixture of 25µl contained 1µl of first round PCR products that were used as template, 15.5µl dH₂O, 2µl dNTP, 5µl 5X PrimeSTAR GXL Buffer (2mM MgCl₂), 0.5µl 5X PrimeSTAR GXL DNA Polymerase and 0.5µl (0.2µM) of each primer. Inner primers for the amplification of gp41 were gp41A (For) 5'TAGGAGCTTTGTTTCCTTGGGTTC-3' and gp41B (Rev) 5'GGTGAATATCCCTGCCTAACTCTATT-3'. PCR was performed using the following steps: Initial denaturation at 97°C for 7minutes, 10 cycles of 94°C for 10 seconds, 52,5°C for 30 seconds, and 68°C for 2 minutes, 35 cycles of 94°C for 10 seconds, 53°C for 30 seconds, 68°C for 2 minutes increasing by 10 seconds each cycle, and final elongation at 69°C for 5 minutes, holding at 4°C until products were stored at 20°C. Inner primers used for gp120 were ED5 (For) 5'-ATGGGATCAAAGCCTAAAGCCATGTG-3' and ED12 (Rev) 5'AGTGCTTCCTGCTGCTCCCAAGAACCCAAG-3'. PCR was performed using the following steps: 25 cycles of denaturation at 98°C for 10 seconds, annealing at 57°C for 15 seconds, and extension at 68°C for 2 minutes, holding at 4°C. Amplicons were stored at -20°C.

3.5.1.3 Quantification of Amplified Viral DNA

DNA quantification was done using 1% agarose gel electrophoresis. For each gel run one litre of 1X TBE buffer was prepared from 10X stock by diluting 100ml of 10X TBE with 900ml

of distilled water. Agarose gel was prepared by weighing 10g of agarose powder and dissolving it in 100ml of 1X TBE buffer. The solution was heated to boiling point while slowly swilling the beaker at regular intervals until solution was clear. Gel solution was cooled down by running tap water on the beaker to about 60°C and was slowly poured into the gel stand with a comb and allowed to set at room temperature for 30 minutes. Electrophoresis tank was filled to the full mark with 1X TBE buffer and the comb was removed from the gel. Eight micro litres of molecular weight maker (DNA ladder) was loaded on the first well. For each specimen, including the negative control, 5µl of the PCR products were mixed with 3µl of loading dye and pipetted into the wells. Gel was run for 30-45 minutes from a negative to a positive current at high voltage. Gel was viewed using ultra violet light machine and the pictures were saved.

3.5.1.4 PCR Product Purification

Amplified PCR products were purified using the QIAquick Purification Kit (Qiagen, Hilda, Germany). According to the manufacturer instructions, 100µl of PB buffer was added to 20µl of second round PCR products and mixed by pipetting up and down. QIAquick spin columns equaling the number of amplified samples were placed in 2ml collection tubes and labelled with sample identification numbers. Each DNA sample was pipetted into the silica membrane of the QIAquick columns to bind the DNA and then centrifuged at 13 000rps for 30 seconds. Flow through was discarded and the QIAquick placed in the same collection tube. To wash, 75µl of Buffer PE was pipetted to the QIAquick column and centrifuged at 13 000rps for 30 seconds. Again, flow trough was discarded and the QIAquick column placed back in the same collection tube. Column was centrifuged for an additional 1 minute then placed in a clean 1.5ml micro-centrifuge tube. To elute DNA, 30µl of Buffer EB (10mM Tris-Cl, pH 8.5) was pipetted to the centre of the QIAquick membrane and allowed to stand for 1 minute, column was then

centrifuged at 13 00rps for 1 minute and the QIAquick membrane column discarded. Purified DNA was stored at -20°C before sequencing cycle was done.

3.4.1.5 Sequencing Cycle PCR

Sequencing cycle PCR was performed for all purified samples using BigDye terminator v.3.1 Cycle Sequencing Kit: Applied Biosystems. PCR was performed for a 10µl reaction mixture which contained 3.8µl dH₂O, 3µl BigDye 5X Sequencing Buffer, 1µ BigDye Terminator, 0.2µl (10uM) primer and a 2µl aliquot of each purified DNA template. Two different overlapping sequence specific primers were used for each sample (forward and reverse primer). PCR was performed using the following steps: 25 cycles of denaturation at 96°C for 10 seconds, annealing at 50°C for 5 seconds, and extension at 60°C for 4 minutes, holding at 4°C until taken for sequence clean up.

3.5.1.6 Sequence Clean Up

Zymo Research Sequence Clean up Kit was used for the sequence clean up, 240µl of Sequencing Binding Buffer was added to 10µl sequencing reaction and mixed by pipetting up and down. Mixture was transferred to a Zymo-Spin 96 Plate mounted onto a Collection Plate and centrifuged at 4000 x g for 2 minutes. In each well of the Zymo-spin plate 300 µl Sequencing Wash Buffer was added. Plate was centrifuged at 4000 x g for 5 minutes. Zymo-spin 96 well plate was placed on top of the 96 well collection plate, 20 µl of 20% formamide solution was then pipetted to each column matrix of the filter plate, then centrifuged at 4000 x g for 2 minutes to elute the DNA.

3.5.1.7 Sequencing and Sequence base editing

Pure DNA was sequenced using the ABI Prism 3130XL Genetic Analyzer: Applied Biosystems according to manufacturer's instructions. Obtained gp41 and gp120 raw sequence base editing was done in Sequencer 5.0 sequencing software. Base editing was done to determine if base representing peaks (blue- C, red-T, Green-A, and Black-G) matched sequence bases. Where there were mixed bases (double peaks or more) it was checked to see if the peaks were true and represented by the correct mixed base. After base editing, contigs were formed and saved in FASTA format.

3.5.1.8 Multiple Sequence Alignment, Sequence Editing and Phylogenetic Analysis

Multiple sequence alignment was done against HIV-1B reference sequence (HxB2) using ClustalX algorithm using the following protocol: Pairwise Alignment (Gap Opening Penalty=15, Gap Extension Penalty=6.66), Multiple Alignment (Gap Opening Penalty=15, Gap Extension Penalty=6.66), DNA Weight Matrix=IUB, Transition Weight=0.5, Delay Divergent Cutoff=30%. Aligned sequences were manually edited in BioEdit software (Hall; 2011). Edited gp41 and gp120 sequences were used to create phylogenetic trees using the MEGA software using the following protocol; Distance Estimation (Pairs of taxa, Maximum Composite Likelihood, Uniform Rates, Complete Deletion for Missing Data) and Neighbor-Joining Algorithm (Maximum Composite Likelihood, Uniform Rates, Complete Deletion for Missing Data). Sequences were analyzed for homology which was used to determine if no two or more sequences obtained where possibly from the same individual.

3.5.1.9 HIV subtype verification

Obtained sequences were subtyped to verify if they are type C using the NCBI HIV and REGA (Abecasis *et al.* 2010) viral subtyping tools. Sequences were uploaded to the software's and subtype results were obtained.

3.5.2 Characterization of gp41

3.5.2.1 Analysis for Naturally Occurring Polymorphisms and T20 Resistance Associated Mutations

Mutation analysis was done in BIOEDIT using aligned gp41 sequences. International AIDS Society United States (IAS USA) 2017 HIV Resistance Mutation Update Report was used as a reference to determine T20 resistance causing mutations G36D/S, I37V, V38A/M/E, Q39R, Q40H, N42T, and N43D in gp41 Heptad Repeat 1 (HR1) domain for major resistance causing mutations. Sequences were also analyzed for minor mutations N126K, E137K, and S138A in the Heptad repeat 2 (HR2) domain of gp41 that cause resistance to T20 when found in association with certain mutations in HR1 (De Feo and Weiss, 2012; Jang *et al.*, 2014; Pessoa *et al.*, 2011). GP41 mutations associated with CXCR4 coreceptor use were determined by identifying the mutation A30T and L34M in the gp41 sequences (Thielen *et al.* 2011).

3.5.3 Characterization of gp120 V3 Loop

3.5.3.1 Prediction of Coreceptor Usage

Viral DNA gp120 sequences in FASTA format were used for the analysis of coreceptor usage using Geno2Pheno (g2p). G2P assigns a score from 0% -100% called a False Positive Rate (FPR), which presents probability of sequence being R5 (Dimonte *et al.* 2012). According to g2p coreceptor FPR is the probability of classifying an R5 virus falsely as X4. Samples with a $\leq 10\%$ FPR were regarded as CXCR4 using and the ones with an FPR $>10\%$ were regarded as CCR5 using (according to the European Consensus group on clinical management of HIV-1 tropism testing (10% FPR). Sequences with an FPR $>10\%$ were predicted as R5 variants, and the ones with $\leq 10\%$ FPR as X4. V3 loop sequences were extracted from aligned gp120 sequences and used for the analysis of coreceptor usage in PSSM_{sinsi}. PSSM gives a score of 0 or 1, where sequences with a score of 0 are regarded as CCR5 using variants and those with a 1 are regarded as CXCR4 using. The 11/24/25 rule, the insertion or deletion mutations before and after the

GPGQ motif of the V3 sequences were also used for coreceptor use analysis. The 11/24/25 and 11/25 rules determine positively charged amino acids in position 11, 24 and 25 of the V3 loop amino acid sequences. Combination of positively charged amino acids in these positions determine CXCR4 using strains (Dimonte *et al.*, 2012; Shen *et al.*, 2016). Comparison of coreceptor use results was done across results obtained from g2p, PSSM, 11/24/25 rule, 11/25 rule and the insertion or deletion mutations as verification of correct results.

3.6 Determining the Associations of Age, Gender, viral load, and CD4 count with Coreceptor Usage

Fisher's exact test was used to assess the association between gender and the co-receptor used. Mann-Whitney test was used to assess the association between the co-receptor used and continuous independent variables such as age, viral load and CD4 count.

CHAPTER 4: RESULTS

4.1 Determining Coreceptor Usage and naturally occurring gp41 Polymorphisms and T20 associated mutations

4.1.1 Viral isolation, amplification and quantification

Viral DNA was isolated from 38 buffy coat samples and amplified using nested PCR.

Results are shown in Figure 4.1 and 4.2 below:

4.1.1.1 Quantification of gp120 amplicons

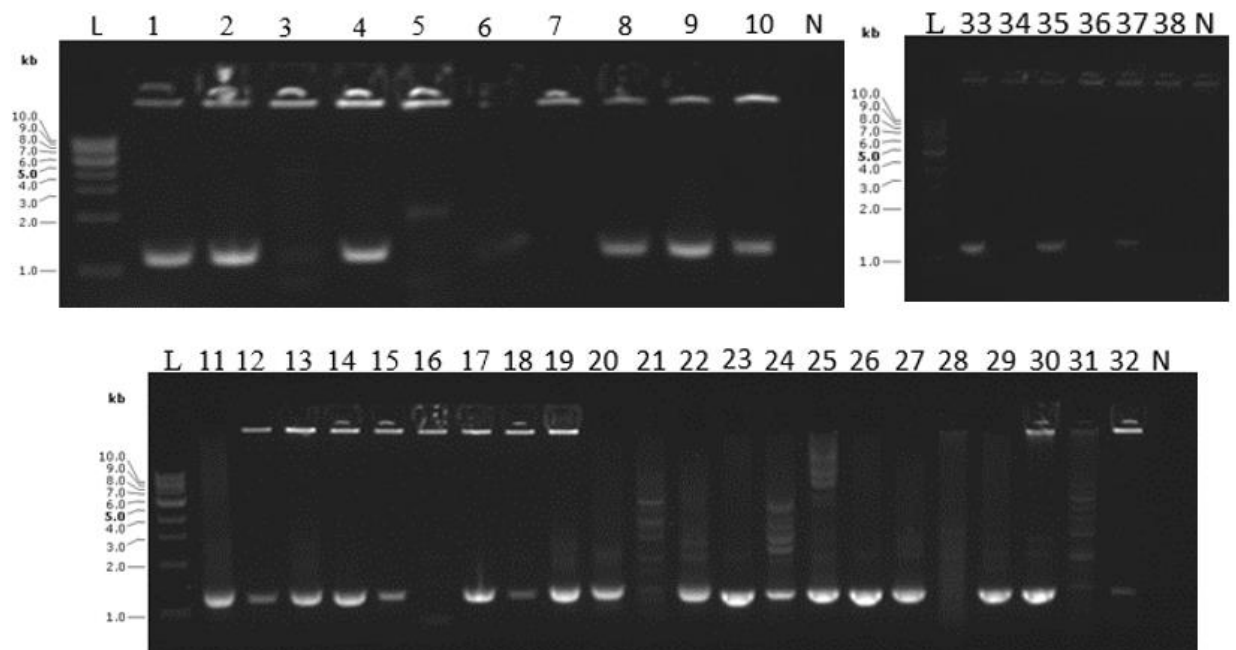


Fig 4.1: Gel electrophoresis results showing amplified gp120 amplicons (about 1,2kb).

Molecular weight marker (Ladder) is represented by L, N is the negative control and numbers 1-38 represents the patient's samples. Each band is representative of an amplified DNA sample. Samples 3,5,6,7,16, 21, 28, 31, 32, 34, 36 and 38 did not amplify (refer to appendix 4 for PID's).

4.1.1.2 Quantification of gp41 amplicons

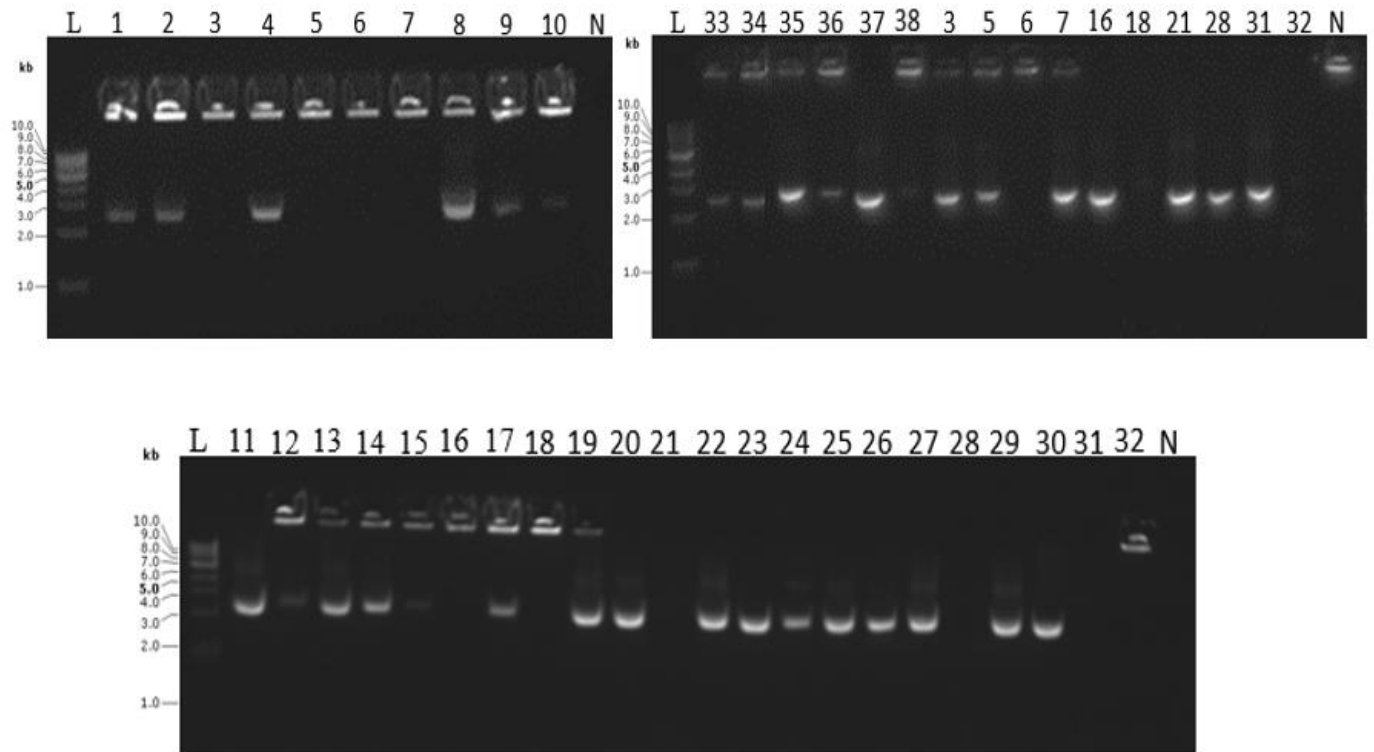


Fig 4.2: Gel electrophoresis results showing some of the amplified gp41 amplicons (about 2.5kb). L represents the molecular weight marker (Ladder), N is the negative control and numbers 1-38 represents the patient's samples. Each band is representative of an amplified DNA sample. Samples 6,18, 32 and 38 did not amplify (refer to appendix 4 for PID's).

4.1.2 Multiple sequence alignment (MSA) of the HIV-1 sequences

Alignment of gp41 and gp120 sequences against HIV-1B (HxB2) reference strain deduced homology between the sequences and HxB2, this verified that obtained sequences were from an HIV species. V3 loop sequences were extracted from the gp120 sequences after MSA

giving 105bp for all sequences. (Refer to appendix 5 and 6 for gp41 and gp120 V3 loop sequence chromatograms).

4.1.3 Phylogenetic analysis of the HIV-1 sequences

Isolates were analyzed by phylogenetics to confirm if there were no different samples obtained from the same individual. All sequences from gp41 and gp120 V3 loop regions were found to be phylogenetically distinct from each other, therefore no different samples were from the same patient. All sequences clustered with the HIV-1C subtype [Fig 4.6 and 4.7].

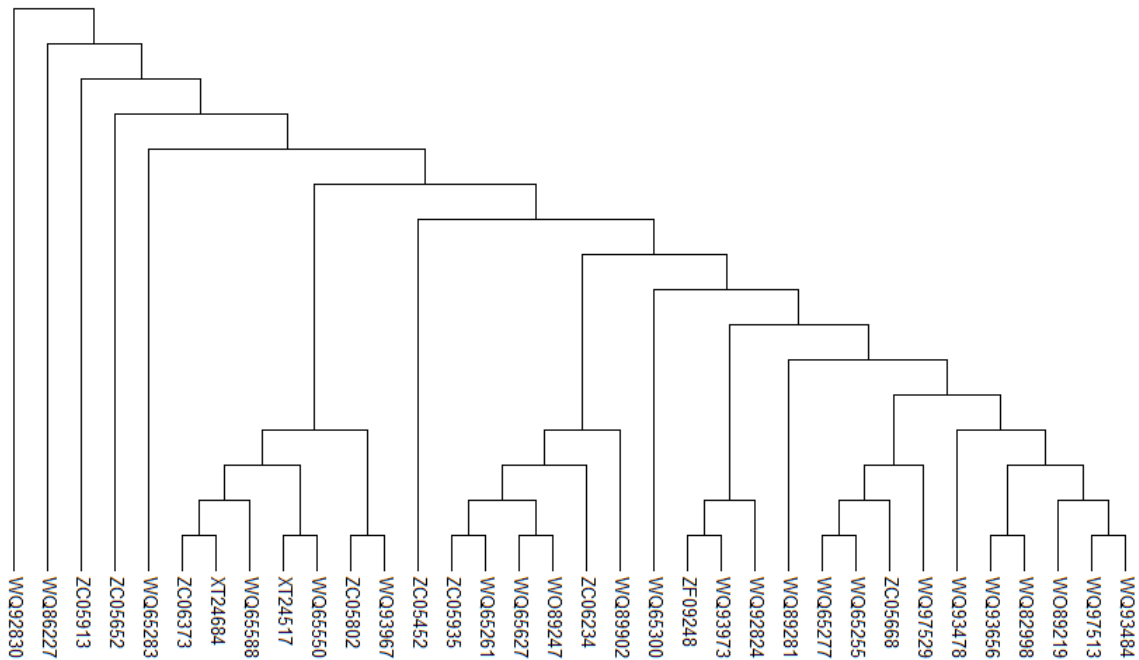


Fig 4.5: Phylogenetic Tree based on alignment of 500 base pairs (bp) of gp41 region nucleotide sequences of isolates from salvage patients failing therapy, compared with HIV-1B (HxB2)

reference strain. Tree was created using MEGA software: Distance Estimation (Pairs of taxa, Maximum Composite Likelihood, Uniform Rates, Complete Deletion for Missing Data) and Neighbor-Joining Algorithm (Maximum Composite Likelihood, Uniform Rates, Complete Deletion for Missing Data)

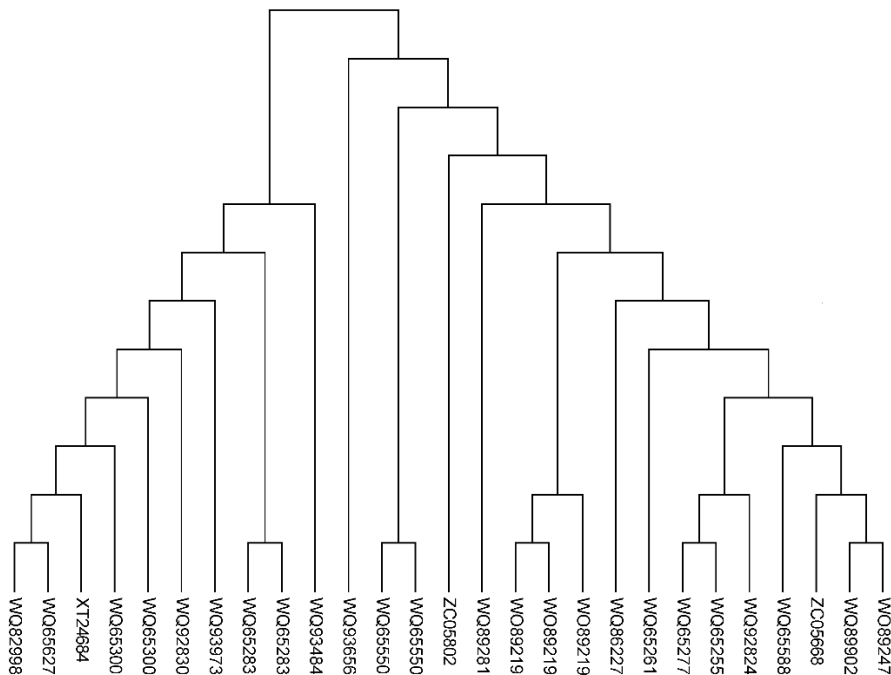


Figure 4.6: Phylogenetic Tree based on alignment of 105 base pairs (bp) of gp120 V3 region nucleotide sequences of isolates from salvage patients failing therapy compared with HIV-1B (HxB2) reference strain. Tree was created using MEGA software: Distance Estimation (Pairs of taxa, Maximum Composite Likelihood, Uniform Rates, Complete Deletion for Missing Data) and Neighbor-Joining Algorithm (Maximum Composite Likelihood, Uniform Rates, Complete Deletion for Missing Data)

4.1.4 HIV-1 subtype verification using NCBI HIV and REGA subtyping tools

All samples belonged to HIV-1 subtype C.

4.2 Characterization of gp41

4.2.1 Mutation Analysis for Naturally Occurring gp41 Polymorphisms and T20 Resistance Associated Mutations

A CXCR4 use associated mutation L34M of gp41 HR1 was found in 7.7% (2/26) of the isolates. One isolate was determined to be an X4 variant (50%) and the other a R5 variant by g2p, PSSM, and the 11/24/25 rule. Major mutation G36S, and minor mutations E137K and S138A were observed in different isolates [Figure 8]. Polymorphisms were observed in both HR1 and HR2 of gp41 with a greater variability observed in HR2. Polymorphisms N42S, L54M, A67T, I69V, and V72I, were noted in HR1 and E119Q, N125S, S129D, S129N, L130T, H132Y, S133R, I135L, Q147K, E148D, E151A, K154S, A156N, and S157N in HR2 domains [Figure 9].

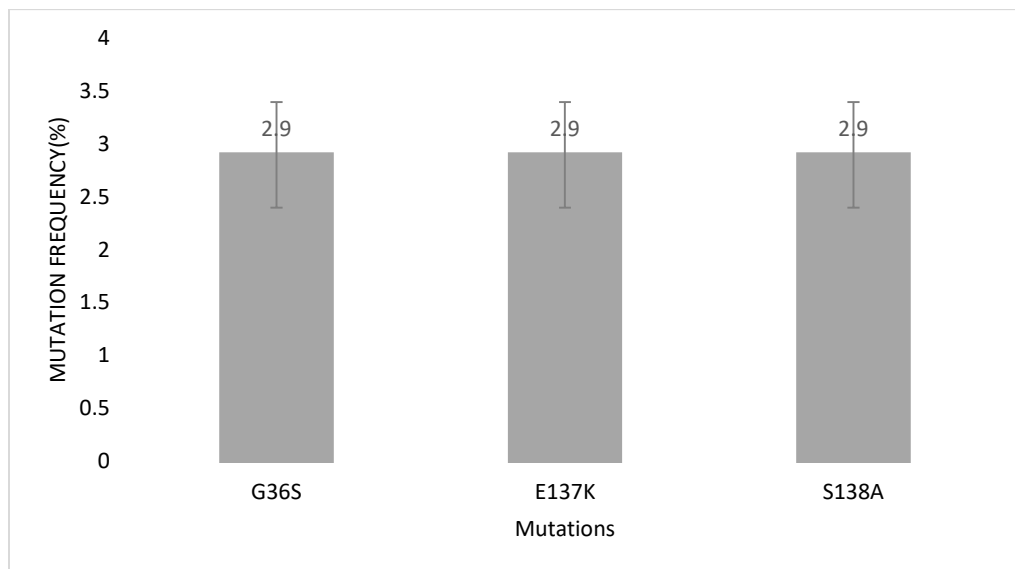


Figure 4.7: Frequency of naturally occurring major and minor T20 resistance associated mutations in T20 patients. All obtained mutations (G36S, E137K and S138A) showed a 2.9% prevalence rate.

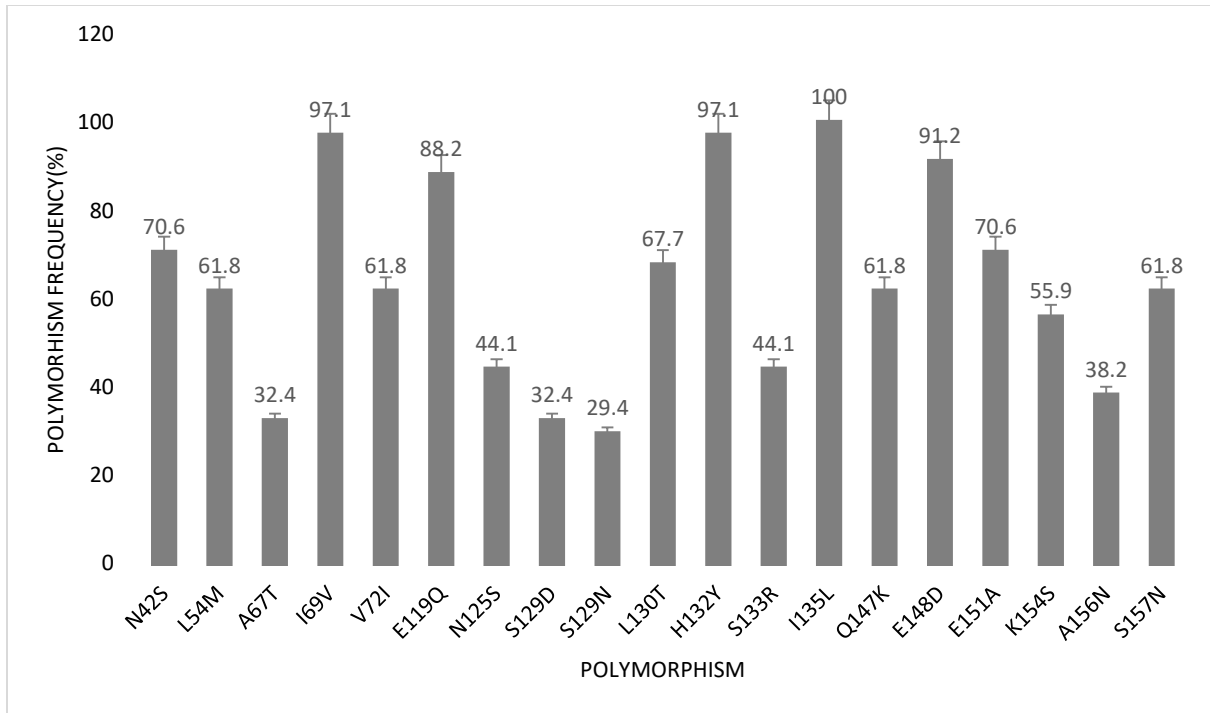


Figure 4.8: Frequency of polymorphisms of the HR1 and HR2 domains of gp41.

4.3 Characterization of gp120 V3 loop

4.3.1 Coreceptor analysis using Geno2Pheno(g2p) and PSSM

Coreceptor analysis by Geno2pheno showed a 64% (16/25) prevalence rate of R5 variants and 36% (9/25) X4 variants at a Falls Positive Rate (FPR) of 10%. PSSM showed the same coreceptor use variation as Geno2pheno for the same sequences [Table 4.8].

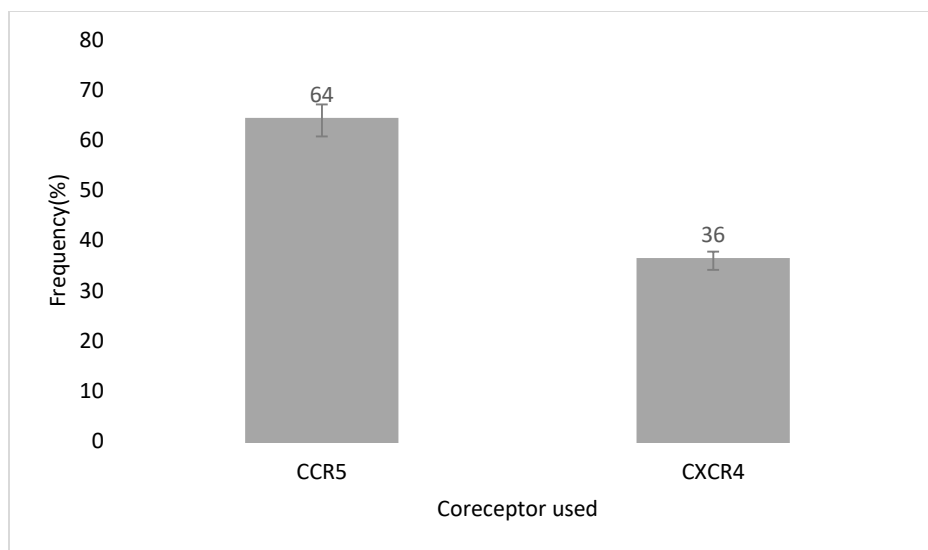


Fig 4.9: Prevalence of the CCR5 and CXCR4 coreceptor using variants by triple drug resistant strains. Analysis done using the geno2pheno (FPR-10%) and PSSM software.

4.3.2 Coreceptor analysis by the 11/24/25 Rule

Positively charged amino acids at positions 11 and/or 24 and/or 25 of the V3 loop denote an X4 variant, this is the so called 11/24/25 rule. Of the nine X4 variants 6 (66.7%) were found to have a positively charged amino acid residues on at least one of the positions 11, 24 and 25 of the 9 (22.2%) had a deletion at either position 11 or 24. All R5 variants had no positive amino acid residues or deletions at these positions.

4.3.3 Coreceptor Analysis Using the Insertion/ Deletion Rule

The amino acid sequences of the V3 loops were aligned and variations in the GPGQ motif (within the V3 loop) were found. All R5 (17/17, 100%) and 66.7% (6/9) X4 variants presented a GPGQ crown motif. The remaining X4 presented GRGQ (1/9; 11.1%), GPGR (1/9, 11.1%) and 1(11.1%) had an insertion (L) mutation after the motif. A double amino acid deletion

prior to the GPGQ was noted in 2 variants (2/26, 7.7%), both of which were determined to be X4.

4.4 Effects of patients demographic (age and gender) and clinical parameters (viral load and CD4 count) on Coreceptor Use

4.4.1 Patients Demographics and Data

For the 25 patients analyzed for coreceptor usage 13(52%) were female, mean age 40.5(20-62) and 48% male, mean age 32.8(21-55). Overall mean age was 36.8(20-55) years. Viral load (VL) counts were collected for 22/25 patients (88%), their median HIV-1 RNA load was 16651(128-359601) copies/ml. CD4 counts were available for 13/25patients (52%) and had a median CD4 count of 333 (128-2580) cells/ μ l. Full demographics and data on Table 4.1 and 4.2. The empty boxes represent absence of data.

Table 4.1: Demographics for the patients under study.

PID	Gender	Age	Viral Load	CD4 count	Coreceptor usage
WO89219	M	22	2616	140	CCR5
WO89247	M	21	1065	231	CCR5
WQ65255	F	-	35447	-	CCR5
WQ65261	F	-	2816	-	CCR5
WQ65277	-	-	75887	-	CCR5
WQ65300	M	45	946	-	CCR5
WQ65550	M	45	128	-	CCR5
WQ65627	F	18	3692	-	CCR5
WQ89281	M	55	214799	367	CCR5
WQ92824	-	-	-	-	CCR5
WO92830	M	40	359601	234	CCR5
WQ93973	F	33	36063	380	CCR5

ZC05802	F	49	-	-	CCR5
ZC05913	F	20	78412	-	CCR5
ZC06234	F	34	114102	254	CCR5

Table 4.1 Continued

ZF09248	M	53	3690	670	CXCR4
WQ65283	F	-	33891	-	CXCR4
WQ82998	M	25	17906	204	CXCR4
WQ86227	M	42	15396	473	CXCR4
WQ89902	F	53	1741	2580	CXCR4
WQ93484	M	29	94857	730	CXCR4
WQ93656	F	42	-	333	CXCR4
WQ97513	F	20	78412	-	CXCR4
XT24684	F	40	1079	128	CXCR4

Table 4.2: CCR5 and CXCR4 specific mean and median for age, viral load and CD4 count

	CCR5(n =16)			CXCR4		
	Age	Viral load	CD4 count	Age	Viral load	CD4 count
Mean	36.35	66376.00	326.19	40.22	30993.13	741.22
Median	36.65	35755.00	326.00	42.00	17906.00	730.00

4.4.2: Gender Distribution of Coreceptor Used

There was no statistically significant difference between male and female on the co-receptor used (unadjusted odds ratio (OR) = 2.0, 95% Confidence interval 0.35 to 11.4; p = 0.669.)

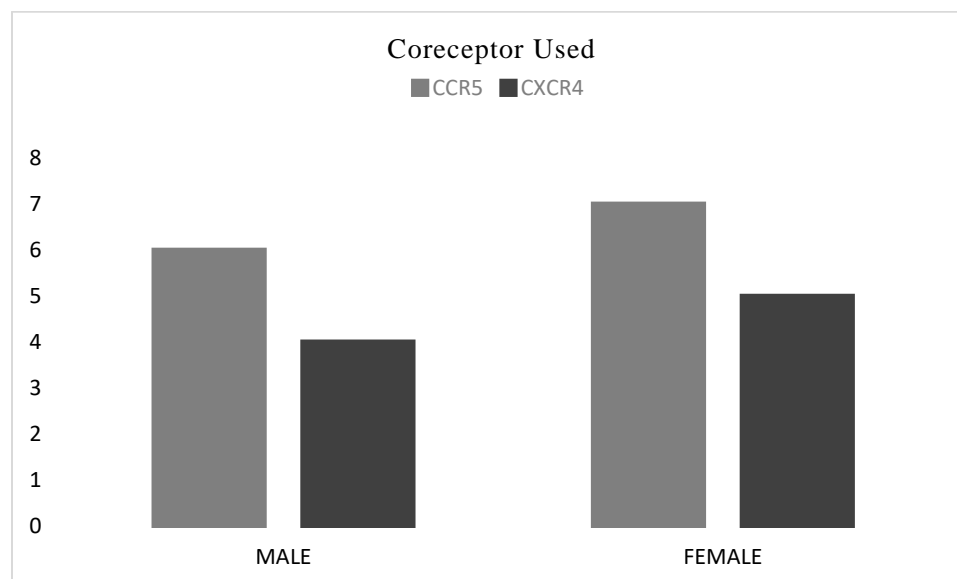


Fig 4.10: Gender distribution of co-receptor used

4.4.3: Assessing the Median Difference Between CCR5 and CXCR4 on Age, Viral Load and CD4 Count

There were no statistically significant differences in the median of age ($p = 0.443$) and viral load (0.760) between CCR5 and CXCR4 [Table 4.3]. However, there was statistically significant difference in the median of CD4 count between CCR5 and CXCR4 ($p = 0.037$) [Table 4.3].

Table 4.3: Mann-Whitney test results to assess the median difference between CCR5 and CXCR4 on age, viral load and CD4 count

	Co-receptor Used	Mean Rank	Sum of Ranks	Z score	P value
Age	CCR5 (n = 16)	12.16	194.50	-0.767	0.443
	CXCR4 (n = 9)	14.50	130.50		
Viral load	CCR5(n = 16)	13.34	213.50	-0.311	0.760
	CXCR4(n = 9)	12.39	111.50		
CD4count	CCR5(n = 16)	10.69	171.00	-2.147	0.037
	CXCR4 (n = 9)	17.11	154.00		

CHAPTER 5: DISCUSSION

This study sought to determine coreceptor usage by MDR HIV strains, identify naturally occurring gp41 and gp120 V3 loop mutations associated with resistance to entry inhibitors, identify gp41 polymorphism variations of heptad repeat 1 and 2 domains and determine the associations of patients demographics (age and gender) and clinical parameters (viral load and CD4 count) on coreceptor usage in isolates from patients failing salvage therapy. Our results showed a moderately high proportion of X4 variants in multi-drug resistant strains of salvage patients failing treatment in Bots. X4 variants are highly virulent and increase the risk of acquiring opportunistic infections (OI), leading to disease progression and death. High X4 rates

reduce the potential use of CCR5 antagonists as alternative therapy (Le *et al.*,2015). In analyzing gp41 HR1 and HR2 the study showed low frequencies of major and minor T20 resistance associated mutations distributed between different individuals. Non-experience to T20 reduces the occurrence of T20 resistance associated mutations, increasing the potential use of fusion inhibitors for the control and management of HIV in patients who are non-responsive to their current treatment (Jang *et al.*,2014) Polymorphisms at different variances were observed in HR1 and HR2 domain of gp41 HR1, with HR2 presenting higher variations in mutations and polymorphisms.

Previous studies suggested that there were lower frequencies of X4 variants in HIV-1C infection compared to HIV-1B where patients with highly progressed disease show a $\geq 50\%$ CXCR4 coreceptor usage (Lin *et al.* 2011). Phuphuakrat *et al* (2014) reported that R5 variants were harbored by 80–90% of treatment naive HIV-1 infected patients and 50–60% of treatment experienced. According to Lin *et al* (2011) the overall prevalence of R5X4 and X4 variants among ART naive HIV-infected patients in Botswana was considered lower than 15%. Our study showed a 36% switch to using the CXCR4 coreceptor in highly treatment experienced HIV-1C patients which is consistent with results from Green *et al* (2012). In this study we observed the polymorphism at high frequencies: E119Q (88.2%), Q147K (61.8%) and A156N (38.3%) which is consistent with the study by Ismael *et al* (2014). Specific mutations of the V3 loop are responsible for giving the virus ability to switch coreceptor usage from CCR5 to CXCR4 (Tsibris 2007). This limits the efficacy of CCR5 antagonists in the treatment of variants that have acquired the ability to use the CXCR4 coreceptor as the use of CCR5 antagonists on R5X4 variants leads to the emergence of X4 variants that are more virulent and are associated with disease progression (Abbate *et al.*, 2011).

This study was limited by the fact that coreceptor prediction identified only X4 and R5 variants and not R5X4, however this did not have an effect on the outcome of the coreceptor results because geno2pheno identifies R5X4 variants as X4. As T20 use becomes more common, patients are becoming core-infected and superinfected with HIV variants already resistant to T20 (Carmona *et al.*, 2005). A study in France, revealed a 6% prevalence of T20 resistant strains in treated patients, with plasma RNA >1,000 copies/ml and suggested that it could potentially be the transmission source of T20-resistant viruses (Peuchant *et al.*, 2007). Single mutations G36D/S, V38A/M, Q40H, N42S/T/D/E, N43D/K/S, L44M and L45M cause a 5 to 10 fold susceptibility to T20, and double mutations of HR1 G36S/L44M, G36S/V38M, V38A / N42D, V38A / N42T , V38E / N42S , N42T / N43S and N42T/N43K have highest level of T20 susceptibility reduction with some showing an average reduction of 100 fold (Peuchant *et al.*, 2007).

The low prevalence of T20 resistance mutations in T20 naive individuals may suggest that the virus replication capacity may select for the occurrence of T20 mutations even without exposure to the drug. In this study we observed minor mutations E137K and S138A, which are said to reduce the efficacy of T20 in inhibiting viral replication when found in association with major T20 resistance mutations (Ismael *et al.*, 2014). Mutation S138A in association with mutations in HR1 position 42 and 43 increase the resistance levels to 3 fold by disturbing the interaction of HR1 and HR2, subsequently influencing the fusion and inhibitory properties of T20 (Ismael *et al.*, 2014, Jang *et al.*, 2014). A study by Ismael *et al* (2014) states that minor mutation G36S might still contribute to causing resistance to T20 even when not found in associations with other mutations. Overall low prevalence of T20 resistance associated mutations

in Botswana is with regards to no experience to the drug. T20 therefore can be used as an alternative drug for patients failing salvage therapy in Botswana.

Polymorphisms of the HR1 domain, N42S (observed variation of 70.6% in this study), and L54M (observed variation of 61.8% in this study) have previously been associated with increased susceptibility to T20 (Aquaro *et al.*, 2006). HR2 domain polymorphisms S129N is also associated with decreased susceptibility to T20 (Ismael *et al.*, 2014), a 32.4% frequency of S129N was observed in this study. Polymorphisms E148D, E151D, E119Q, Q147K, and A156N have previously been reported in T20 naive and experienced patients at a high. The genetic diversity of the multi drug resistant strains leaves the requirement of continuous assessment of the variants for diagnosis, primary drug resistance and vaccination. There is still need to determine the impact imposed by the described major and minor mutations and polymorphisms on HIV 1C variants and their response to T20 therapy.

This study demonstrates that multi-class drug resistant variants of HIV-1C have the ability to switch coreceptor usage at a frequency more or less constant with that observed in multi-class drug resistant HIV-1B infections. Therefore, the knowledge of viral coreceptor usage may contribute to selection of appropriate coreceptor antagonists as alternative drugs for patients who are not responsive to treatment. The study also demonstrates the natural genetic diversity of gp41 as mutations and polymorphisms, from which it can be determined if fusion inhibitors can be used as alternative therapy for patients infected with multi-class drug resistant HIV variants. In summary this study shows a moderately high frequency of X4 variants, low frequencies of T20 resistance associated mutations and gp41 polymorphisms in multi-class drug resistant variants infecting salvage patients failing therapy.

5.2 RECOMMENDATIONS

In order to fight off drug resistance, reduce the high morbidity and mortality rates of HIV and work towards achieving the 90-90-90 targets set by the UNAIDS the Ministry of Health in Botswana should consider increasing the available therapeutic options. MVC and T20 can be adopted as alternative therapy for patients who are failing treatment and can as well be used for first and second line treatment. Because of the high frequency of CXCR4 using variants in

highly treatment experienced patients in Botswana, tropism testing should be done prior to initiating therapy with MVC and during the course of treatment by the drug so as to determine if patients are eligible for the drug. Administration of drugs effective in therapy would mean efficient management of the HIV infection. Single genome sequencing can be employed for the determination of coreceptor usage by multi drug resistant variants so that low levels of R5X4 and X4 variants can be determined in order to avoid misuse of Maraviroc by preventing emergence of X4 variants during treatment. Studies should also be done to determine levels of R5X4 and X4 variants that can be tolerated and kept under suppression for a long time by CCR5 antagonists. Gp41 has a great genetic variation and would require studies that would determine the effects of the naturally occurring mutations and polymorphisms to treatment by T20.

5.3 CONCLUSIONS

Entry inhibitors can be used as alternative therapy for patients failing salvage therapy. Moderately high frequency of X4 variants in highly treatment experienced patients who are failing salvage therapy would mean use of MVC in Botswana requires tropism testing prior to use of the drug and during course of treatment with the drug. Non T20 treatment experience in

Botswana reduces the prevalence of the major mutations that confer resistance to the drug. T20 is therefore a potential alternative drug for patients failing salvage therapy in Botswana.

APPENDICES

Appendix 1

Reagents, Buffers and Solutions

Items/ Reagent	Manufacturer/Solution
-----------------------	------------------------------

Molecular weight marker	KAPA Biosystems, South Africa
Agarose	
Primers (OFM12, SK145, GP41A, GP41B, ED5, ED12)	Inqaba Biotec, South Africa
10X TBE Buffer	890mM Tris base, 890mM Borate, 20Mm EDTA(Ph8.0)
6X Agarose gel loading dye	0.2% Bromophenol blue, 0.2% Xylene, cyanol 25% glycerol
RNAses	
DNAses	
Formamide solution	90% in DMSO-d6 (99.9 atom % D)
Ethidium bromide	
Ethanol	100% Ethanol

Appendix 2

Kits

Kit	Supplier
-----	----------

KAPA	KAPA Biosystems, South Africa
QIAquick Purification Kit	Qiagen, Hilda, Germany
BigDye terminator v.3.1 Cycle Sequencing Kit	Thermo Fisher Scientific
Zymo Research Sequence Clean up Kit	Zymo Research

Appendix 3

Equipment and Consumables

Equipment	Consumables
-----------	-------------

Themocycler	96 Well PCR Reaction Plates
UV Machine	PCR reaction Plate Covers
Electrophoresis Tanks	Ependorf Tubes
Electrophoresis Machine	Kleenix Paper
3130XL ABIprism Genetic Analyzer	Disposable Glooves
Lamina Flow	Aliquot tubes
Micro pipettes (p10, p20, p200, p1000)	Pipette Tips (10µl, 20µl, 200µl and 1000µl)
Centrifuge	Fine Tip Permanent Markers
Vortex	
Spinner	
Balance scale	

Appendix 4

Sample

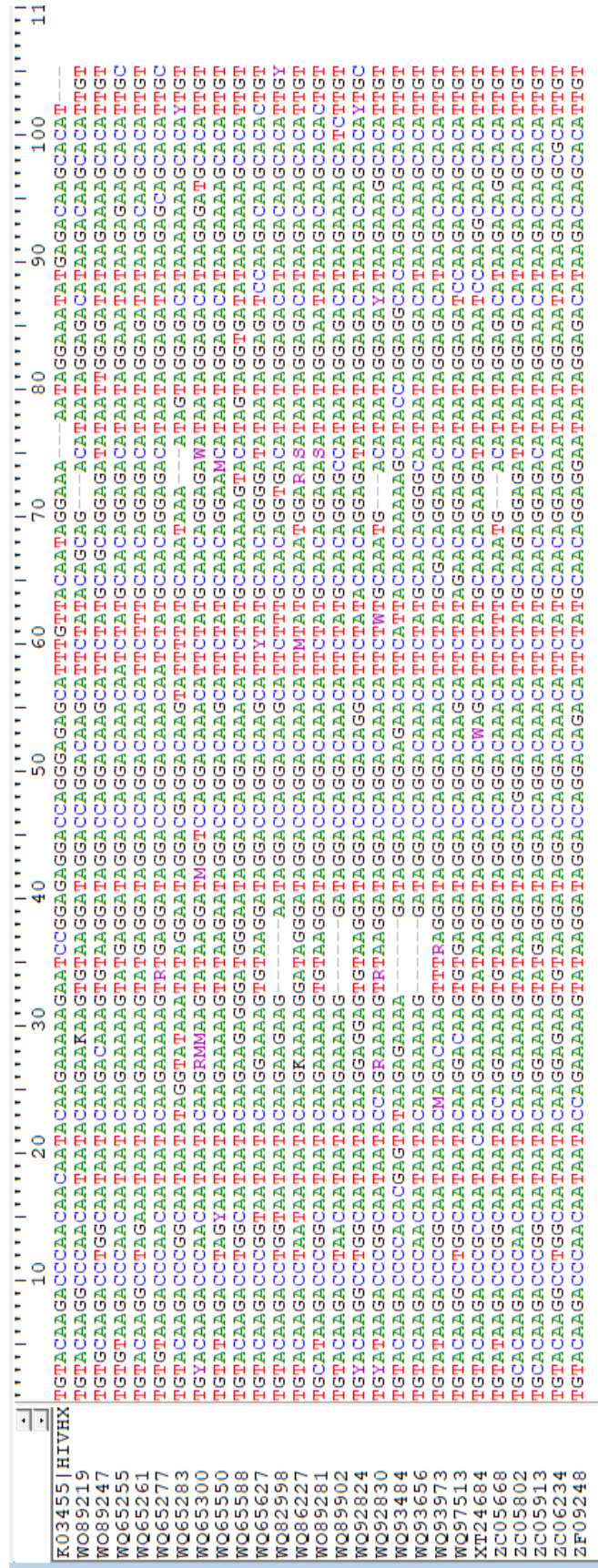
Number	PID
1	WQ65550
2	WQ86227

3 WQ65627
4 WQ65277
5 WQ65299
6 WQ65283
7 XT24517
8 WQ82998
9 WQ93478
10 WQ65261
11 WQ65300
12 WQ93484
13 WQ89902
14 WQ92824
15 WQ65255
16 WQ65594
17 WQ65588
18 XT24684
19 WQ92830
20 WQ93656
21 WQ93967
22 WQ93973
23 ZC05913
24 ZC05935
25 ZC06234
26 ZC05668
27 ZC05652
28 ZC06373
29 ZC05452
30 WQ97513
31 ZC05802
32 WQ97529
33 WQ89281
34 WR08912
35 WO89247
36 WO89219
37 ZF09248
38 ZF09610

	10	20	30	40	50	60	70	80	90	100	110	120	130
K03455 HIVHX	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W089219	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W089247	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W065255	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W065261	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W065277	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W065283	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W065300	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W065550	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W065588	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W065627	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W082998	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W086227	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W089291	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W089902	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W092824	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W093478	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W093484	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W093656	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W093830	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W093967	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W093973	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W097513	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
W097529	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
XT24517	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
XT24684	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
ZC05452	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
ZC05652	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
ZC05668	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC
ZC05802	CTGACGGTACAGGCCAGACA	TTATTTCTGGTATATAGTCAGCAGCAGAGAACAA	TTTTGCTCAGGGCTATTTGAGGCGCAACACACAT	CTGTTGGCAA	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC	CTC

Appendix 5: Multiple sequence alignment of gp41 198 basepairs (bp) sequences against HxB2. Alignment by Clustal X algorithm using

the following protocol: Pairwise Alignment (Gap Opening Penalty=15, Gap Extension Penalty=6.66), Multiple Alignment (Gap Opening Penalty=15, Gap Extension Penalty=6.66), DNA Weight Matrix=IUB, Transition Weight=0.5, Delay Divergent Cutoff=30%. Sequences were highly similar.



Appendix 6: Multiple sequence alignment of gp120 V3 105 base pairs (bp) against HxB2. Alignment by Clustal X algorithm using the following protocol: Pairwise Alignment (Gap Opening Penalty=15, Gap Extension Penalty=6.66), Multiple Alignment (Gap Opening Penalty=15, Gap Extension Penalty=6.66), DNA Weight Matrix=IUB, Transition Weight=0.5, Delay Divergent Cutoff=30%. Gaps (-) on alignment represent deletion mutations. Sequences were highly similar.

REFERENCES

Abbate I, Rozera, G, Tommasi C, Bruselles A, Bartolini B, Chillemi G, Nicastrì E, Narciso P, Ippolito G, and Capobianchi G. R. (2011). “Analysis of Co-Receptor Usage of Circulating Viral and Proviral HIV Genome Quasispecies by Ultra-Deep Pyrosequencing in Patients Who Are Candidates for CCR5 Antagonist Treatment.” *Clinical Microbiology and Infection* **17** (5): 725–31.

Abecasis A B., Wang Y, Libin P, Imbrechts S, Tulio de Oliveira, Camacho R. J, and Vandamme Anne-Mieke A. M. (2010). “Comparative Performance of the REGA Subtyping Tool Version 2 versus Version 1.” *Infection, Genetics and Evolution* **10** (3). Elsevier B.V.: 380–85.

Aquaro S, D’Arrigo R, Svicher V, Perri G. P, Caputo S. L, Visco-Comandini U, Santoro M. (2006). “Specific Mutations in HIV-1 Gp41 Are Associated with Immunological Success in HIV-1-Infected Patients Receiving Enfuvirtide Treatment.” *Journal of Antimicrobial Chemotherapy* **58** (4): 714–22.

Arruda L B, Ladeira de Araújo M, Martinez M. L, Gonzalez C. R, José da Silva Duarte A, Coakley E, Lie Y, and Casseb J.(2014). “Determination of Viral Tropism Viral through Genotypic Phenotypic Assays in HIV-1 infected Patients in Brazil.” *Institute of Medicina, Sao Paulo* **56** (4): 287–90.

Botswana MoH. (2016). “Handbook of the 2016 Integrated HIV Clinical Care Guidelines,” 44. <https://doi.org/10.1037/xhp0000318>.

Bussmann H, Novitsky V, Wester W, Peter T, Masupu K, Gabaitiri L, Kim S. (2005). “HIV-1 Subtype C Drug-Resistance Background among ARV-Naive Adults in Botswana.” *Antiviral Chemistry and Chemotherapy* **16** (2): 103–15.

Cenci J.A, Tavoschi L, Avenio G.D, Narino P, Becattini S, Bernasconi D, Chiappi M, Torre L. La, Sukati H, Vardas E, Presti A. Lo, Cella E, Ciccozzi M, Picconi O, Monni P, Ensoli B and Butto S, (2012) 'Characterization of Variable region of the Gp120 Protein from HIV-1 Subtype C Virus Variants Obtained from Individuals at Different Disease Stages in Sub-Saharan Africa,' *J AIDS Clinic Res*, **S8** (006):1-9.

Checkely M.A, Luttge B.G and Freed E.O, (2011), 'HIV-1 envelope glycoprotein biosynthesis, trafficking and incorporation', *J Mol Biol*, **410** (4): 582-608.

Chesna L. K, and Fellner C. (2017). "Promising HIV Treatments in Late-Stage Clinical Development" **42** (10): 647-49.

Cortez I. J., and Maldarelli F. (2011). "Clinical Management of HIV Drug Resistance." *Viruses* **3** (4): 347-78. De Feo, C. J., and Weiss C. D. (2012). "Escape from Human Immunodeficiency Virus Type 1 (HIV-1) Entry Inhibitors." *Viruses* **4** (12): 3859-3911.

De Feo, C. J., and Weiss C. D. (2012). "Escape from Human Immunodeficiency Virus Type 1 (HIV-1) Entry Inhibitors." *Viruses* **4** (12): 3859-3911.

Del Rio, C. (2006). "Current Concepts in Antiretroviral Therapy Failure." *Top HIV Med* **14** (3): 102-6.

Dimonte S, Muhammed B. M, Fabio M, Domenico D P, Francesca C. S, Valentina S, and Carlo F. P. (2012). "Selected Amino Acid Changes in HIV-1 Subtype-C Gp41 Are Associated with Specific Gp120V3signatures in the Regulation of Co-Receptor Usage." *Virus Research* **168** (1-2). Elsevier B.V.: 73-83.

Dimonte S, Fabio M, Valentina S, Arrigo R, Carlo-federico P, and Francesca C. (2011). “Selected Amino Acid Mutations in HIV-1 B Subtype Gp41 Are Associated with Specific Gp120 V3 Signatures in the Regulation of Co- Receptor Usage.” *Retrovirology* **8** (1). BioMed Central Ltd: 33.

Farahani M, Anusha V, Lebelonyne R, Seipone K, Anderson M, Avalos A, Chadborn T. (2014). “Outcomes of the Botswana National HIV/AIDS Treatment Programme from 2002 to 2010: A Longitudinal Analysis.” *The Lancet Global Health* **2** (1): 44–50.

Gantner P, Laurence M. J, Charlotte Sueur, Raffi F, Fagard C, Lascoux-Combe C, Salmon D. (2016). “Drug Resistance and Tropism as Markers of the Dynamics of HIV-1 DNA Quasispecies in Blood Cells of Heavily Pretreated Patients Who Achieved Sustained Virological Suppression.” *Journal of Antimicrobial Chemotherapy* **71** (3): 751–61.

Gaolathe T, Wirth K. E, Holme M. P, Makhema J, Moyo S, Chakalisa U, Kadima E. (2016). Botswana's Progress Towards Achieving the 2020 UNAIDS 90-90-90 Antiretroviral Therapy and Virological Suppression: A Population-Based Survey.” *Lancet* **3** (5): 1–21.

Geijitenbeek T. B, Kwon D. S, Torensman R, Van Vliet SJ, van Duijnhoven G.C, Middle J, Cornelissen I.L, Nottet HS, KewalRamani V.N, Littman D.R, Figdor C.G and van Kooyk Y, (2000), “DC-SIGN, a dendritic cell- specific HIV-1 binding protein that enhances trans-infection of T cells”, *Cell*, **100** (5): 587-597.

Green T. N., Archary M, Gordon M. L, Padayachi N, Lie Y, Anton E. D, Reeves J. D,(2012). “Drug Resistance and Coreceptor Usage in HIV Type 1 Subtype C-Infected Children Initiating or Failing Highly Active Antiretroviral Therapy in South Africa.” *AIDS Research and Human Retroviruses* **28** (4): 324–32.

Hall T, Biosciences I, and Carlsbad C. (2011), Bioedit; an important software for molecular biology. *GERF Bull Biosci*, **2** (1), pp. 60-61

Hardin J, Bertoni G and Kleinsmith L. J, (2012), *Becker's World of the Cell*, Pearson Education Inc, San Francisco, 648.

Hartley O, Klasse PJ, Sattentau QJ and Moore JP, (2005), “V3: HIV’s switch-hitter”, *AIDS Res Hum Retroviruses* **21**: 171–189.

Hoy J and Lewin, (2004), “HIV Management in Australasian a guide for Clinical Care”, *Australation Society for HIV Medicine Inc*.

http://www.aidsinfonet.org/fact_sheets/view/408. (Accessed 17/11/2017)

<https://aidsinfo.nih.gov/understanding-hiv-aids/glossary/643/salvagetherapy>. (Accessed 17/11/2017)

<https://aac.org/aachealth-library/salvage-treatment-for-hiv/>. (Accessed 17/11/2017)

Imaz A, VicenFa lc, and Ribera E. (2011). “Antiretroviral Salvage Therapy for Multiclass Drug-Resistant HIV-1-Infected Patients: From Clinical Trials to Daily Clinical Practice.” *AIDS Rev* **13** (3): 180–93.

Ismael N, Bila D, Mariani D, VubiAl, Mabunda N, Abreu C, Jani I, and Tanuri I. (2014). “Genetic Analysis and Natural Polymorphisms in HIV-1 Gp41 Isolates from Maputo City, Mozambique.” *AIDS Research and Human Retroviruses* **30** (6): 603–9.

Jang D. H, Yoon C. H, Choi B. S, Chung Y. S, Kim H. Y, Chi S. G, and Kim S. S. (2014). “Characterization of Gp41 Polymorphisms in the Fusion Peptide Domain and T-20 (Enfuvirtide) Resistance-Associated Regions in Korean HIV-1 Isolates.” *Journal of Korean Medical Science* **29** (3): 456–59.

Le A. Q, Taylor J, Dong W, McCloskey R, Woods C, Danroth R, Hayashi K, Milloy M. J, Poon A. F. Y, and Brumme Z. L. (2015). “Differential Evolution of a CXCR4-Using HIV-1 Strain in CCR5wt/Wt and CCR5 Δ 32/ Δ 32 Hosts Revealed by Longitudinal Deep Sequencing and Phylogenetic Reconstruction.” *Scientific Reports* **5**

Lin N. H, Smeaton L. M, Giguel F, Novitsky V, Moyo S, Mitchell R. M, Makhema J, Essex M, Lockman S, and Kuritzkes D. R. (2011). “Prevalence and Clinical Associations of CXCR4-Using HIV-1 Among Treatment-Naive Subtype C-Infected Women in Botswana.” *JAIDS Journal of Acquired Immune Deficiency Syndromes* **57** (1): 46–50.

Marukutira T, Stoové M, Lockman S, Mills L. A, Gaolathe T, Lebelonyane R, Jarvis J. N. (2018). “A Tale of Two Countries: Progress towards UNAIDS 90-90-90 Targets in Botswana and Australia.” *Journal of the International AIDS Society* **21** (3)

Nolan K. M, Del Prete G. Q, Jordan A. P. O, Haggarty B, Romano J, Leslie J. G, and Hoxie J.A. (2009). “Characterization of a Human Immunodeficiency Virus Type 1 V3 Deletion Mutation That Confers Resistance to CCR5 Inhibitors and the Ability to Use Aplaviroc-Bound Receptor.” *Journal of Virology* **83** (8): 3798–3809.

Pang W, TomS. C, and Zheng Y. T. (2009). “Current Peptide HIV Type-1 Fusion Inhibitors.” *Antiviral Chemistry and Chemotherapy* **20** (1): 1–18.

Pessoa, L. S, Valadão A. L. C, Abreu C. M, Calazans A. R, Martins A. N, Azevedo S. S. S. D, Couto-Fernandez J. C, Azevedo M. C .M .V, and Tanuri A. (2011). “Genotypic Analysis of the Gp41 HR1 Region from HIV-1 Isolates from Enfuvirtide-Treated and Untreated Patients.” *Journal of Acquired Immune Deficiency Syndromes* **57** (SUPPL. 3): 197–201.

Peuchant O, Capdepon S, Ragnaud J. M, Aurillac-Lavignolle V, Thiébaud R, Fleury H, and Masquelier B. (2007). “Primary Resistance to Enfuvirtide (T20) in Recently HIV-1 Infected, Antiretroviral-Naive Patients from the ANRS Aquitaine Cohort.” *Antiviral Therapy* **12** (4): 559–62.

Phuphuakrat, A., S. Phawattanakul, E. Pasomsab, S. Kiertiburanakul, W. Chantratita, and Sungkanuparph S.(2014). “Coreceptor Tropism Determined by Genotypic Assay in HIV-1 Circulating in Thailand, Where CRF01_AE Predominates.” *HIV Medicine* **15** (5): 269–75.

Qian K, Morris-Natschke S. L, and Kuo-Hsiung L. (2009). “HIV Entry Inhibitors and Their Potential in HIV Therapy.” *Medicinal Research Reviews* **29** (2): 369–93.

Recordon-Pinson P, Souli C, Flandre P, Descamps D, Lazrek M, Charpentier C, Montes B. 2010. “Evaluation of the Genotypic Prediction of HIV-1 Coreceptor Use versus a Phenotypic Assay and Correlation with the Virological Response to Maraviroc: The ANRS Genotropism Study.” *Antimicrobial Agents and Chemotherapy* **54** (8): 3335–40.

Sharp and Hahn, (2011), “Origin of HIV and AIDS Pandemic”, *Cold Spring Harb Perspect Med.*, 1 (1).

Shen H. S, Yin J, Leng F, Teng R. F, Xu C, Xia X. Y, and Pan X. M. (2016). “HIV Coreceptor Tropism Determination and Mutational Pattern Identification.” *Scientific Reports*. Nature Publishing Group: 1–11.

Svicher V, Caputo S. L , Ubaldo V, Santoro M, and Bertoli A. (2014). “Specific Mutations in HIV-1 Gp41 Are Associated with Immunological Success in HIV-1-Infected Patients Receiving Enfuvirtide ... Success in HIV-1-Infected Patients Receiving Enfuvirtide Treatment,” no. May.

Thielen A, Lengauer T, Swenson L. C, Dong W. W. Y, and Mcgovern R. A.(2011). “Original Article Mutations in Gp41 Are Correlated with Coreceptor Tropism but Do Not Improve Prediction Methods Substantially” **328**: 319–28.

Tsibris, A. (2007). “Update on CCR5 Inhibitors: Scientific Rationale, Clinical Evidence, and Anticipated Uses.” *PRN Notebook* 12.

UNAIDS. 2016. “Botswana | UNAIDS.” Available at <http://www.unaids.org/en/regionscountries/countries/botswana>.

UNAIDS (2017). “The Joint United Nations Programme on HIV and AIDS(UNAIDS);Regional HIV Statistics.” *Fact Sheet*, no. July: 1–8.

UNICEF. 2013. “UNICEF HIV and AIDS in Eastern and Southern Africa.” Available at http://www.unicef.org/esaro/5482_HIV_AIDS.html.

Weber J, 2001), “The pathogenesis of HIV-1 infection, *British Medical Bulletin*, **58**: 61-72.

Weiss R. A. (2013). “Thirty Years on: HIV Receptor Gymnastics and the Prevention of Infection.” *BMC Biology* 11. <https://doi.org/10.1186/1741-7007-11-57>.

Woodman Z and Williamson C, (2009), “HIV molecular epidemiology: transmission and adaptation human populations”, *Curr Opin HIV AIDS*, **4** (4): 247-252.

WHO. 2017. *Hiv Drug Resistance Report 2017 Trends Quality Action*.<http://apps.who.int/iris/bitstream/10665/255896/1/9789241512831-eng.pdf?ua=1>.

Zhang H, Tully D. C, Zhang T, Moriyama H, Thompson J, and Wood C. 2010. “Molecular Determinants of HIV-1 Subtype C Coreceptor Transition from R5 to R5X4.” *Virology* **407** (1). Elsevier Inc.: 68–79.

