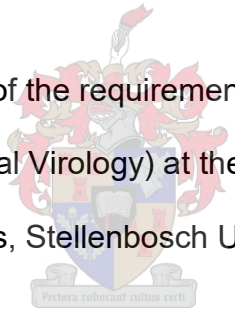
HIV-1 molecular diversity and drug resistance mutations amongst immuno-competent, therapy naïve infants/children and adults in Yaounde, Cameroon

Josiah Otwoma Gichana



Thesis submitted in fulfilment of the requirements for the award of a Master of Medical Science Degree (Medical Virology) at the Faculty of Medicine and Health Sciences, Stellenbosch University

Supervisor: Dr. Graeme Brendon Jacobs

Co-supervisor: Dr. George Mondinde Ikomey

December, 2017

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own original work, that I am the authorship owner thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: December 2017

Abstract

Background: In Cameroon, HIV infections range between 550, 000 to 690, 000 for adults aged 15 to 49 years and a prevalence rate at 4.5%. In children of 0 – 14 years, HIV infections range between 34, 000 to 44, 000. The country harbors both HIV type 1 (HIV-1) and HIV type 2 (HIV-2). HIV groups found in Cameroon include M, N, O, P variants. Group M subtypes are the most prevalent, with CRF02_AG accounting for approximately 40% of all HIV infections. This is unlike other regions globally where other group M subtypes like C are the predominant ones. The high genomic diversity of HIV-1 and the emergence of drug resistant associated mutations (RAMs) continue to be a major challenge in designing standardized laboratory protocols for HIV testing, vaccine development and providing successful lifelong therapy to HIV infected patients. In Cameroon, drug resistance rates for therapy naïve individuals are currently at 3.8% in adults and 3.6% in children. This study aimed at identifying HIV-1 diversity and evaluate drug resistant mutations (DRMs) in two different cohorts of therapy-naïve infants/children and adults in Cameroon. **Methods:** A total of 180 plasma samples were collected from therapy naïve HIV positive patients that included: (1) 55 plasma samples from proxy-consented infants/children aged 9-72 months old with unknown prevention of mother-to-child transmission (PMTCT) exposure and (2) 125 plasma samples from adults of 15 to 50 years old. The CD4+ T-cell count was performed using standard methods following manufacturer's instructions. To study the HIV-1 diversity and resistance in the two cohorts, partial *pol* Protease (PR), Reverse Transcriptase (RT) and Integrase (IN) regions of the HIV-1 genome were targeted for conventional PCR amplification and Sanger DNA

sequencing. Phylogenetic inference using Neighbor-Joining (NJ) trees were used to cluster and infer subtypes. **Results:** In the infants/children cohort, the CD4+ T-cell count ranged between 500-2000 cells/m3 (a median of 33.0%) and the HIV-1 viral load between 3000-6000 copies/ml (a median of 4.96 RNA copies/ml). A total of 37/55 (67.3%) paediatric cohort samples were amplified for at least one of the HIV-1 *pol* fragments. These include 29/55 (52.0%) for the PR, 27/55 (49.0%) for the RT and 28/55 (51.0%) for IN. The most predominant HIV-1 strain was G/CRF02_AG at 62.5% (n = 15). Other subtypes detected include subtype A (20.8%; n = 5), C (8.3%; n = 2) and F2 (8.3%; n = 2). Three sequences (11.1%) could not be assigned to any subtype with confidence. Levels of DRMs to Protease inhibitors (PIs), nucleoside reverse transcriptase inhibitors (NRTIs) and non-NRTI were 27.6% (only minor DRMS were observed for PR), 3.7% and 40.7%, respectively. The NRTI mutations observed showed high-level resistance to Zidovudine (AZT), Tenofovir (TDF), Didanosine (DDI) and Stavudine (D4T), and low to intermediate-level resistance to Lamivudine (3TC), Abacavir (ABC), and Emtricitabine (FTC). The NNRTI mutations observed showed high level resistance to Nevirapine (NVP) and Efavirenz (EFV) with reduced susceptibility to Etravirine (ETR) and Rilpivirine (RPV). In the adult cohort, the RT fragment (n = 55) was used for phylogenetic analysis with majority of the sample sequences clustered with HIV-1 subtype G/CRF02_AG which accounted for 40% (n = 22), CRF22_01A1 (10.9%; n = 6), C (1.8%; n = 1), B (1.8%; n = 1), other complex forms – 37_cpx/11_cpx (3.6%; n = 2). Twenty three samples (41.8%) could not be assigned to any subtype with confidence. The levels of drug resistance for adults was 5.4% for both NRT and NNRT inhibitors - 4.0% had low level resistance to EFV, ETR, NVP and RPV while 1.4% had intermediate to high level resistance to

ABC, FTC, TDF, EFV and NVP. **Conclusion:** Cameroon continues to harbor many HIV-1 subtypes and circulating recombinant forms (CRFs) as observed in both cohorts. Furthermore, rare group O and other group M subtypes like C were noticed within the study cohorts suggesting an improvement in sensitivity of detection methodologies currently used. Drug resistance is a major challenge to current antiretroviral drug regimens as illustrated by the detection of RAMS in both cohorts of this study. Continuous surveillance of the HIV diversity and drug resistance is therefore necessary to better manage the HIV-1 pandemic.

Acknowledgments

I most sincerely thank my supervisor Dr. Graeme Brendon Jacobs for believing in me and giving me a chance to pursue this course under his supervision. You have nurtured me with a wide range of skills in the realm of scientific research that forever will be engrained in my memory. My gratitude also goes to my co-supervisor Dr. George Mondinde Ikomey and your team at CSCCD in Cameroon, you made it possible for me to access the samples and your input in various stages of my project cannot be taken for granted. I thank you all (Merci!). Professor Susan Engelbrecht, nothing can describe you better than the knowledge in phylogenetic analyses that largely I gained from you, with much appreciation. Thank you Professor Wolfgang Preiser and Professor Andrew Whitelaw, you made it possible for me to register late - giving me an opportunity that I had thought was long gone. I salute you! Professor Loice Gathece and Dr. EAO Dimba - I deeply appreciate every effort you made for me to join this program. Mr. Njenda Duncan (our research assistant), Ms. Cynthia Tamandjou, Ms. Poovan and Comrades - Mr. Given Mikasi, Mr. Emmanuel Obasa, Mr. Ruhanya and Dr. Laundry - your help, suggestions and positive criticisms in my laboratory work and thesis writing made my project a success. Thank you all. Mrs. Sal van Zyl (our division secretary), Mrs. Rachel Pullen (international office), Mrs. Du Plessis Nina, fellow students and other staff in the Division and National Health Laboratory Services (NHLS) - thank you for supporting me and working with me. Dr. David Otieno Awange, Dr. Wakoli and Mrs. Alice Limo - you took my duties in our pathology laboratory section while I was away and supported me in many different ways. Thank you. Finally, the P4HPT (intra-acp), University of Nairobi and

Stellenbosch University management - thank you for the financial support and the opportunity to study.

Dedications

To my dear wife Gladys Kerubo, our two beautiful daughters - Marion Moraa and Faith Meroka and our son, Emmanuel Gichana, I dedicate this to you all for your unwavering support and understanding. It was not easy being far away from you for two and half years but your love gave me a reason to be strong and courageous in all I do. Forever God bless you**.**

To my dear parents – Mr. Solomon Gichana and Mrs. Jemimah Meroka - God bless you for all you have done for me. The virtues you instilled in me, is what has made me to be who I am today. My dear brothers and sisters- your love, financial support, words of hope and encouragement endeavored the smooth running of my career. God bless you all.

Table of contents

Work submitted for publication

Ikomey G.M., Assoumou O., **Gichana J.O**., Njenda D., Mikasi S.G., Mesembe M., Lyonga E., Jacobs G.B.: *Observed HIV drug resistance associated mutations amongst naïve immunocompetent children in Yaoundé, Cameroon*; Ref: GERMS-OA-2017-0019.R1. Accepted, in press. 2017.

Manuscript for submission

Gichana J.O., Ikomey G.M, Mesembe M., Lyonga E., Jacobs G.B.: *HIV-1 Pol-Phylogenetic diversity and drug resistance mutation profile in adults from Yaounde ,Cameroon – to be submitted to GERMS in October, 2017.*

Conferences and workshop attended

17th – 21st July, 2017. International Union of Microbiological Societies (IUMS) 2017 Congresses - 17th International Congress of Virology, Singapore. Poster presentation: *HIV-1 Pol-phylogenetic diversity and drug resistance mutation profile in two different cohorts (adult and infants) from Yaoundé, Cameroon*: **Gichana J.O**., Ikomey G.M., and Jacobs G.B.

19th - 23rd June, 2017. Introduction to phylogenetics and Advanced phylogenetics workshop

24th May, 2016. American Society of Microbiology (ASM) at Centre for the Study and Control of Communicable Diseases Institute, Yaoundé Cameroon. Oral

presentation: *HIV-1 diversity and resistance in a cohort of infants from Yaoundé, Cameroon*.

1st – 3rd December, 2015. Virology Africa 2015 conference in Cape Town, South Africa. Poster presentation: HIV-1 diversity and resistance in a cohort of infants from Yaoundé, Cameroon: **Gichana J.O.**, George Ikomey G.M., Njenda D., Graeme B. J.

List of Abbreviation

3TC:	Lamivudine
AIDS:	Acquired immunodeficiency syndrome
ART:	Antiretroviral therapy
ARV drugs:	Antiretroviral drugs
aa:	Amino acid
APOBEC3G:	Apolipoprotein B mRNA-editing enzyme-catalytic Polypeptide-like 3G
ABC:	Abacavir
ATV/r:	Atazanavir/ritonavir
AZT:	Zidovudine
BLAST:	Basic Local Alignment Search Tool
Bp:	Base pairs
Bpi:	Boosted protease inhibitors
C1 to C5:	Constant regions 1 to 5
CA:	Capsid
cART:	Combination antiretroviral therapy

CD:	Cluster of differentiation (CD3, CD4, CD8)
cDNA:	Complementary deoxynucleic acid
CCR5:	C-C chemokine receptor type 5
CM:	Cameroon (samples)
CRFs:	Circulating recombinant forms
COMET:	Context-based Modeling for Expeditious Typing
CTLs:	Cytotoxic T-lymphocytes
CXCR4:	C-X-C chemokine receptor type 4
d4T:	Stavudine
dNTPs:	Deoxynucleoside triphosphates
ddNTPs:	Dideoxynucleoside triphosphates
ddNTPs - A,G,C,T:	Adenosine, Guanine, Cytosine, Thymine
DC:	Dendritic Cells
ddI:	Didanosine
DNA:	Deoxyribonucleic acid
DRV:	Darunavir
DRC:	Democratic Republic of Congo

DTG:	Dolutegravir
EDTA:	Ethylene diamine tetra-acetic acid
EFV:	Efavirenz
Env :	Envelope gene
Env :	Envelope protein
ER:	Endoplasmic reticulum
ETR:	Etravirine
EVG:	Elvitegravir
FACS:	Fluorescent Assorted Cells Sorter
FSW:	Female sex worker
FTC:	Emtricitabine
GACB :	Germany Advisory Committee Blood
gag:	Group-specific antigen gene
gp:	Glycoprotein
GTR:	General time reversible
HHV8:	Human herpes virus 8
HIV:	Human Immunodeficiency Virus

HIV-1:	Human Immunodeficiency Virus type 1
HIV-2:	Human Immunodeficiency Virus type 2
HIVDB:	HIV drug resistance database
HIVDR:	HIV drug resistance
HLA:	Human Leukocyte Antigen
HREC:	Human Research Ethics Committee
HTLV:	Human T-Lymphotropic virus
IDV/r:	Indinavir/r
IDUs:	Intravenous drugs users
IN:	Integrase
INIs:	Integrase inhibitors
jpHMM:	jumping profile Hidden Markov Model
kDa:	kilo Dalton
kb:	Kilo – base pairs
LAS:	Lymphadenopathy syndrome
LAV:	Lymphadenopathy virus
LANL:	Los Alamos National Laboratory

LTNPs:	Long–term non-progressors
LPV/r:	Lopinavir/ritonavir
LTR:	Long terminal repeat
MA:	Matrix protein
MEGA:	Molecular evolutionary genetics analysis
MHC:	Major Histocompatibility Complex
mRNA:	Messenger Ribonucleic acid
MSM:	Men having sex with men
MTCT :	Mother-to-child transmission
NC:	Nucleocapsid
nef:	Negative factor gene
NF – kβ:	Nuclear factor kβ
NFV:	Nelfinavir
NNRTIs:	Non-nucleoside/nucleotide reverse transcriptase inhibitors
NRTIs:	Nucleoside/Nucleotide reverse transcriptase inhibitors
NTC:	Non-Template Control
NTD:	N-terminal domain

NVP:	Nevirapine
ORF:	Open reading frame
PBS buffer:	Phosphate buffer saline buffer
PBS:	Primer binding site
PC:	Positive Control
PCR:	Polymerase chain reaction
PI:	Protease inhibitor
PIC:	Pre- Integration complex
PLHA:	People living with HIV/AIDS
PMTCT:	Prevention of mother-to-child transmission
PR:	Protease
Pol:	Polymerase gene
®:	Registered
R:	Ritonavir
RAMs:	Resistance associated mutations
rev:	Regulator of viral expression gene
RIP:	Recombinant Identification Program

RNA:	Ribonucleic acid
rpm:	Revolution per minute
RPV:	Rilpivirine
RTV:	Ritonavir
RT:	Reverse transcriptase
RTIs:	Reverse transcriptase inhibitors
SCUEAL:	Subtype classification using evolutionary algorithms
SDRM:	Surveillance drug resistant mutations
SHAPE:	Selective 2'-hydroxyl acylation analyzed by primer extension
SIV:	Simian immunodeficiency virus
SIVcpz:	SIVs of chimpanzee
SIVgor:	SIVs of Gorilla
SIVsm:	SIVs of sooty mangabeys
SU:	Surface glycoproteins
TAMs:	Thymidine analogue mutations
tat:	Trans-activator of transcription gene
Taq:	*Thermus aquaticus*

TDF:	Tenofovir
tHIVDR:	Transmitted HIV drug resistance
TM:	Transmembrane
TPV/r:	Tipranavir/ritonavir
TRIM:	Tripartite motif
TAE:	Tris-Acetate-EDTA
UNAIDS:	Joint United Nations programme on HIV/AIDS
URFs:	Unique recombinant forms
U3:	Unique 3’ region
U5:	Unique 5’ region
USA:	United States of America
US-FDA:	United States Food and Drug Administrations
V1 to V5:	Variable regions 1 to 5
V3:	Third Variable
Vif:	Virion infectivity factor gene
Vpr:	Viral protein R gene
vpu:	Viral protein U gene

WHO:	World Health Organization

List of tables

List of figures

CHAPTER 1 Introduction and literature review

1.1 Introduction

Human Immunodeficiency Virus (HIV) is the causative agent of Acquired Immunodeficiency Syndrome (AIDS) in humans (Weiss, 1993; Douek *et al*., 2009). The HIV/AIDS pandemic has continued to be a major global health problem ever since the first case was identified in 1981, (Gottlieb *et al*., 1981; Barré-Sinoussi *et al*., 1983). By end of 2016, sub-Saharan Africa remained the worst hit accounting for almost two thirds (25.6 million) of the world's 36.7 million people living with the virus [(WHO, 2017: *http://www.who.int/mediacentre/factsheets/fs360/en/*)]. The West and Central African sub-region accounted for 16.6% of the global tally. Approximately 310, 000 infected individuals succumbed to AIDS-related illnesses, and 370, 000 of the world's 1.8 million new HIV cases were reported in the same region in 2016 [(UNAIDS, 2017- *http://www.unaids.org/sites/default/files/media_asset/20170720 _Data_book_2017_en.pdf*)]. In Cameroon, approximately 560,000 (470, 000 to 650, 000) adults aged 15 to 49 years were living with HIV; a prevalence rate of about 3.8% (3.1% - 4.5%) [(UNAIDS, 2017: HIV and AIDS estimates, 2016 - *http://www.unaids.org/en/regionscountries/countries/cameroon/)*]. In children of 0 – 14 years from Cameroon, HIV infections was about 46,000 (35,000 – 57,000) [(UNAIDS, 2017: HIV and AIDS estimates, 2016 -*http://www.unaids.org/en/ regionscountries/countries/cameroon/)*]. Since the introduction of combination antiretroviral therapy (cART), there has been a decline of new HIV infections and related deaths globally. In 2016, 19.5 million people globally were able to access cART; up from 15.8 million in June 2015 - a 19.0% increase [(Penazzato *et al*., 2015; UNAIDS, 2017: Fact Sheet, 2017- http://www.unaids.org/sites/default/files/media

_asset/UNAIDS_FactSheet_en.pdf – accessed on 20th October, 2017]. According to World Bank data report based on UNAIDS estimates, the cART coverage in Cameroon was about 30.0% in 2016 [(World Bank - https://data.worldbank.org/indicator/SH.HIV.ARTC.ZS?locations=CM]. Many HIV positive patients are now experiencing a healthy and longer life expectancy. Unfortunately, there is still continuous spread of HIV and more so from expectant mothers to their unborn children (Sama *et al*., 2017). One of the contributory factors for continued spread of the virus is its diverse nature. As will be discussed in section 1.2.4, certain subtypes of HIV spread faster as they easily adapt in their new host. Cameroon is known to harbour majority of diverse HIV groups and subtypes, and this could be the reason why new HIV infections are being experienced (Agyingi *et al*., 2014; Faria *et al*., 2014). The upscale of cART coverage with non-adherence to the prescribed regimen has also led to HIV drug resistance (HIVDR) further fuelling HIV spread (Gordon *et al*., 2015; Fonsah *et al*., 2017). The emergence and increase of HIVDR globally is a serious threat for the overall success of comparting the HIV pandemic and this requires continuous surveillance (Aghokeng *et al*., 2013). Many countries in sub-Saharan Africa have reported approximately 10.0% HIVDR in therapy naïve patients [(http://nicd.ac.za/assets/files/NICD%20Communicable%20 - Diseases%20Communique_Mar2016_final.pdf)]. In Cameroon, HIVDR rates for therapy naïve population are at 3.8% and 3.6% for adults and children respectively (Billong *et al*., 2013). The World Health Organization (WHO) is currently supporting surveillance programmes for HIVDR by developing an HIVDR database. The database focuses on (1) supporting countries and genotyping laboratories to produce high-quality reports through quality assurance of epidemiological and sequencing

data, linking to the most recent Stanford HIVDR interpretation algorithms [(*https://hivdb.stanford.edu/)* – accessed on 3rd February 2017], to provide standardized HIVDR interpretations, (2) support dissemination of data for global reporting and, (3) provide countries with a long-term secure repository for their HIVDR data [(*hiv-aids@who.int*)]. The HIVDR surveillance program play a key role in managing the HIV pandemic, but in a resource-limited country like Cameroon, implementing such programs is expensive and rarely available.

1.2 Literature review

This section covers the origin and diversity of HIV and how it crossed into humans from presumably the non-human primates until it became a global pandemic. The cross-species transmission of HIV into humans has been attributed to its ability to adapt, replicate and mutate very fast, giving rise to various subtypes that are currently circulating globally. This section also describes the HIV type 1 (HIV-1) structure, genome organization and its life cycle as they play a vital role in drug resistance.

1.2.1 Origin of HIV

Through molecular clock and phylogenetic analyses, HIV is estimated to have originated from sub-Saharan Africa around 1931 (Zhu *et al*., 1998; Salemi *et al*., 2001). It is believed that the virus resulted from multiple zoonotic cross-species transmission events from Simian Immunodeficiency Virus (SIV) infected non-human primates found in the West-Central Africa region (Hahn *et al*., 2000; Taylor *et al*., 2008). The earliest documented reports that predate the origin of HIV-1 were of an

infected adult male plasma and tissue samples taken in 1959 from Kinshasa - the capital city of the Democratic Republic of Congo DRC (Zhu *et al*., 1998; Worobey *et al*., 2008). The pandemic lineage of HIV-1 group M has been associated with SIV variant (SIVcpz) found in chimpanzees (sub-group; *Pan troglodytes troglodytes)* from Southern Cameroon (Keele *et al*., 2006; Van *et al*., 2007). The non-pandemic variants of HIV-1 groups N (non-M/non-O), O (Outlier) and P (which are associated with SIV variants in gorillas) are confined within Central and West Africa (Plantier *et al*., 2009; Vallari *et al*., 2011). Another SIV (SIVsmm) found in wild sooty mangabey monkeys (*Cercocebus atys atys*), also from West Africa is most closely related to two strains of HIV-2 (A and B) and probably transmitted HIV-2 (Apetrei *et al*., 2004). HIV belongs to a group of retroviruses (*Retroviridae* family) and genera *Lentivirus* - which literally means “a slow virus’. As for all viruses of the *Retroviridae* family, they attack the immune system allowing opportunistic infections and rare malignancies to thrive, and if not treated, infection can lead to early death (Okome-Nkoumou *et al*., 2014; Germany Advisory Committee Blood, 2016). The first reported outbreak of HIV/AIDS in Africa was from a heterosexual population of the Democratic Republic of Congo (DRC). From here, HIV is believed to have spread to other parts of the region (Piot *et al*., 1984).

1.2.2 HIV diversity

Based on genetic similarities, HIV has been phylogenetically classified into two distinct Types: HIV type 1 (HIV-1) and HIV type 2 (HIV-2). The nature of HIV is that it mutates very rapidly and is characterized by extensive heterogeneity that reflects a dynamic evolutionary process that is continuously generating new subtypes and

several new recombinant forms (Roberts *et al*. 1988; Lemey *et al*., 2006). The observed extensive genetic diversity results from many probable events, including a rapid replication cycle, function of point mutations, recombination of different strains, proofreading inability of RT that leads to high error rate and selective immune pressures from the host (Roberts *et al*., 1998; Lessells *et al.,* 2012; Tongo M. and Burgers W., 2014). Genetic differences between HIV-1 and HIV-2 are approximately 45.0 – 50.0%. HIV-1 has been phylogenetically classified into four distinct groups; Group M (Main/Major), Group O (outlier), Group N (Non-M/Non-O) and Group P (Plantier *et al.,* 2009; Sharp *et al.,* 2011). The genetic difference among these HIV-1 groups is approximately 35.0% (Korber *et al*., 2001). HIV-1 group M has been subdivided further into nine distinct non-recombinant subtypes/clades denoted with letters, A-D, F-H, J-K as indicated in Figure 1.1, with over eighty Circulating Recombinant Forms (CRFs) and several Unique Recombinant Forms (URFs) (*https://www.hiv.lanl.gov/content/sequence/HIV/HIVTools.html*).

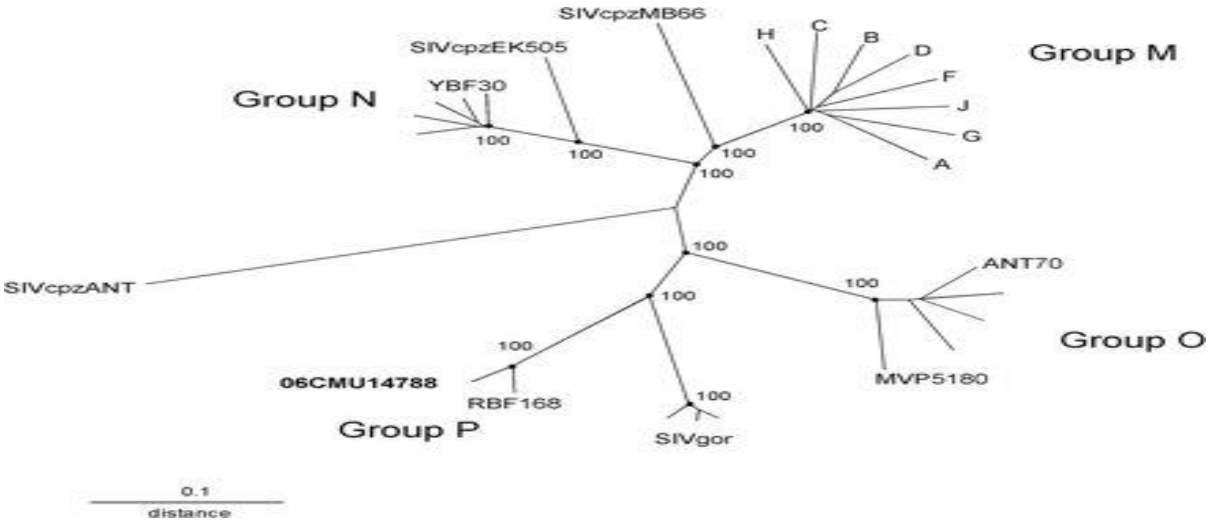


Figure 1.1 A phylogenetic inference of HIV- 1 group M, N, O and the recently discovered group P rooted with SIV variants of non-human primates. HIV-1 group M subtypes are denoted with letters; A-D, F-H and J (Plantier *et al*., 2009; Vallari *et al*., 2011).

The nine HIV-1 group M subtypes are equidistant from one another, with the *pol* gene remaining conserved, while there is an average inter-subtype genetic variability of 15.0% for the *gag* gene and 25.0% for the *env* gene (Robertson *et al.*, 2000; Spira *et al*., 2003). HIV-2 has been classified into eight non-recombinant subtypes; A-H with one recombinant form, namely HIV2_CRF01_AB (Clavel *et al*., 1986; (*https://www.hiv.lanl.gov/content/sequence/HIV/HIVTools.html)*).

1.2.3 HIV infections in Humans

The first clinical cases of HIV were diagnosed in men having sex with men (MSM) in Los Angeles, United States of America (USA) in 1981 (Gottlieb *et al*., 1981). These men developed a rare lung infection (Pneumonia) caused by a fungi called *Pneumocystis carinii*. Other groups of men in New York and California were also

found with an aggressive skin cancer known as “Kaposi’s Sarcoma” caused by Human Herpes Virus 8 (HHV-8) (Hymes *et al*., 1981). These diseases found to be common in people with impaired immunity, was subsequently defined as Gay Related Immune Deficiency (GRID) (CDC, 1982). Up until this point, the causative agent was unknown and efforts for its discovery began. In 1983, Luc Montagnier and Françoise Barré-Sinoussi from the Pasteur Institute in France isolated the virus and named it “Lymphoadenopathy Associated Virus (LAV)” (Barre-Sinoussi *et al*., 1983). The virus was found to infect cells of immune system and eventually leading to death. Soon after, cases of infection were detected in IDUs, Haitians, and heterosexuals (Pape *et al*., 1983; Bloom, 1984). This confirmed that the disease was also transmitted through blood contact. In June 1983, HIV was found in children (Oleske *et al*., 1983) and later confirmed that they may have acquired it directly from their mothers before, during or shortly after birth (CDC, 1983). Broder and Gallo (1984)] at the National Cancer Institute in USA, isolated a virus they called “Human T cell Lymphotropic virus (HTLV III). The LAV and HLTV III were found to be similar (Marx *et al*., 1984). Unfortunately, the HTLV III was co-infected with Human T cell Leukemia virus (HTLV) and therefore, Gallo was not honoured with the Nobel price (Gallo *et al*., 1983). In 1986, HIV-2 was also isolated from the West African region (Hughes A. and Corrah T., 1990; Clavel F. *et al*., 1986). In May 1986, the international Committee on Taxonomy of viruses officially renamed the virus “Human Immunodeficiency Virus” as it is known to date (Case *et al*., 1986).

1.2.4 HIV global distribution

The pandemic HIV-1 group M subtypes is spread globally (Hemelaar *et al*., 2011) and the greatest diversity of its subtypes and recombinants are found in Central-West Africa, as shown in Figure 1.2.

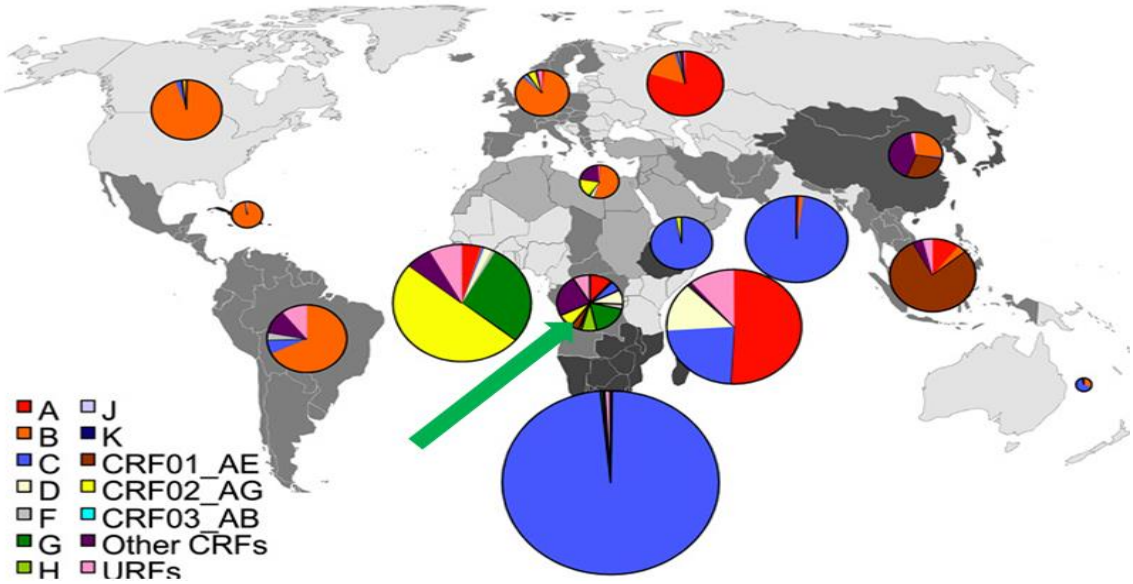


Figure 1.2 Geographical distribution of the pandemic HIV-1 group M subtypes. Each different colour of the superimposed pie chart represents a different HIV-1 group M subtype found within that region. Cameroon, which is part of Central Africa sub-region harbours majority of HIV-1 group M subtype: A,C,D,G,H and J; Circulating recombinant forms like CRF01_AE and CRF02_AG, and other rare subtypes as indicated by the arrow (Source: Hemelaar *et al*., 2011).

The lesser virulent groups of HIV-1; namely group O, N and P, are also confined to Central-West Africa; especially Cameroon and the surrounding countries (Vallari *et al*., 2011; Ayouba *et al*., 2000). Group M, subtype C accounts for almost 50.0% of global infections and is common in India and South Africa. Subtype A accounts for

12.0% and predominates the East Africa axis, while subtype B (10.0%) is common in the Americas and Europe. Subtype G and CRFs, which are common in Central-West Africa, accounts for 6.0% and 5.0% of global infections, respectively. Paradoxically, regions like Central-West Africa with the greatest diversity, experience less HIV pandemic. Other subtypes like F, H, J, K and recombinants account for approximately 13.0% (Hemelaar *et al*., 2011; Lau & Wong, 2013). Group O, which is second in the global spread after the group M pandemic, was discovered in Cameroon in 1998. Globally, group N has been reported in less than 20 individuals; while, group P - that was discovered in 2009, has been reported in less than 2 individuals. Both HIV-1 and HIV-2 and majority of HIV-1 groups, have been found in Cameroon apart from one group N isolate (Plantier *et al*., 2009; Sharp *et al*., 2011).

1.2.5 HIV-1 structure and genomic organization

HIV is an enveloped virus measuring approximately 100-120nm in diameter, spherical in shape with an associated matrix. The matrix is covered by a bi-lipid membrane layer where viral spikes of glycoprotein 120/Surface Protein (gp 120, SU) are attached. The matrix encloses the capsid (p24 CA) which comprises of approximately 2000 gag protein copies. The viral genome is enclosed by the capsid and consists of two identical copies of unspliced positive single stranded RNA that are non-covalently linked, weighing about 9.7 kilo base pair (kb) (Korber *et al*., 2001). Figure 1.3a shows a mature HIV structure.

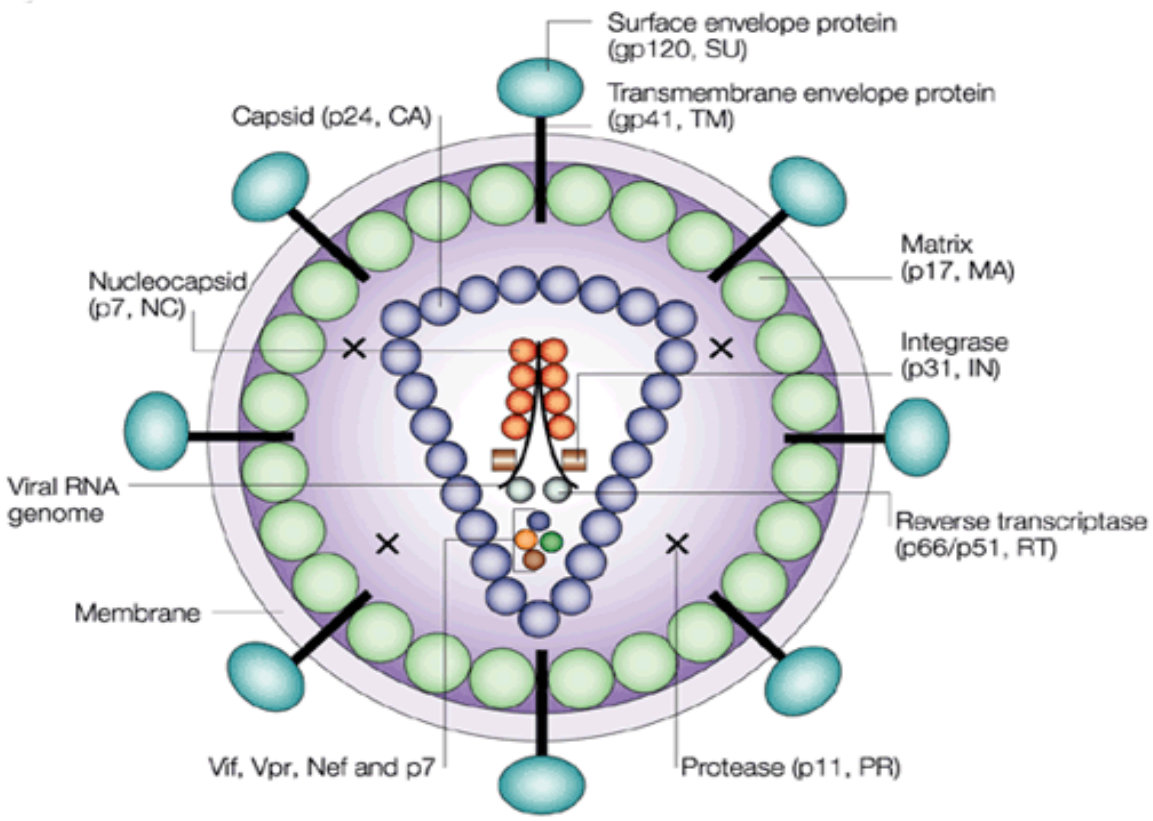


Figure 1.3a. A schematic illustration of a single mature RNA virion showing; Structural proteins that consist of (i) Env glycoprotein- surface glycoprotein (gp120 SU) and the attached transmembrane glycoprotein (gp41 TM). (ii) Gag polyprotein- matrix protein (p17), Capsid protein (p24) and Nucleocapsid protein (p9). The Pol proteins- protease (p10), Reverse transcriptase (p66/p51) and Integrase (p31). The viral RNA indicated at the center consists of two copies of unspliced positive single stranded RNA tightly attached to the Nucleocapsid protein (p9). Source: Nature Review: Available at: *http://www.nature.com/nri/journal/v2/n4/fig_tab/nri776_F1.htm).*

A characteristic of retroviruses is that the genetic information is stored in the RNA genome (Figure 1.3b) and it encodes for nine genes namely: group specific antigens (*Gag*), Polymerase (*Pol*), Envelope (*Env*), trans-activator of transcription (*tat*), regulator of expression of viral proteins(*rev*), Negative factor (*Nef*), Virion infectivity

factor (*vif*), Viral protein r (*vpr*) and Viral protein unique (*vpu*). These nine genes are found along three overlapping Open Reading Frames (ORFs) and are grouped into three categories of proteins: structural proteins (Gag, Pol, and Env), regulatory proteins (Tat and Rev) and accessory proteins (Nef, Vif, Vpu and Vpr). At both 5' and 3' ends, the RNA genome is flanked by two identical nucleotide sequence repeats known as "long terminal repeats (LTRs)" of approximately 640 base pairs (bps) in length.

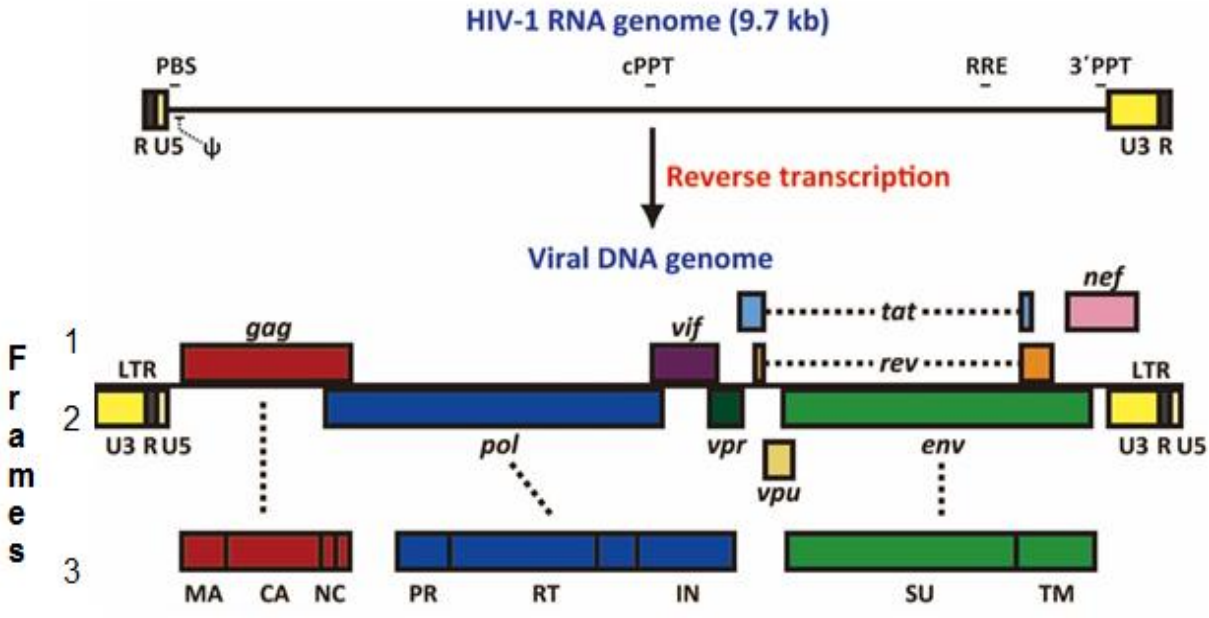


Figure 1.3b. Schematic illustration of HIV-1 genome organization**.** The RNA genome is approximately 9.7 kb flanked with two LTRs at 5'-3' end. The *env and gag genes* codes for structural proteins while, *pol* gene codes for essential viral enzymes. Accessory genes include; the *nef*, *tat* and *rev* and are required for the control of viral gene expression. The *vif*, *vpr* and *vpu* are required for viral infectivity, assembly and release. (Yasutsugu and Youichi, 2011).

1.2.5.1 Polymerase

The Polymerase gene (*Pol*) plays a very important role in the production and functioning of new virions through synthesis of proviral Deoxyribonucleic acid (DNA), viral genomic integration and maturation. The Pol is expressed as a Gag-Pol precursor (Pr160), which during viral maturation process is cleaved to a Pol polypeptide protein. The polypeptide protein is further cleaved into functional viral enzymes - Protease (PR), Reverse Transcriptase (RT), Integrase (IN), RNase H and a transframe protein (TF/p6) (Jack *et a*l., 1988; Oroszlan *et al*., 1990). For the purpose of this study, only the three enzymes- PR, RT and IN will be focused on. The *pol* gene is highly polymorphic, rendering it important in HIV characterization. Furthermore, majority of commonly used cART are designed to target the *pol* genome so as to slow viral progression and transmission (Gea- Banacloche *et al*., 2000; Corado *et al*., 2017).

1.2.5.2 Protease

The PR is a homodimer structure consisting of two non-covalently similar monomer structures of 99-138 amino acid (aa) residues in length. Each monomer has an extended glycine–rich loop that act as an active site for substrate binding (James & Sielecki, 1985). The Protease (PR) has a molecular weight of 11-15 kilo Dalton (kDa) and it is maximally active in acidic environment of ph 4.5-6.5. Furthermore, during the HIV-1 life cycle, the PR enzyme (that is part of the aspartyl protease group) plays a critical role in the budding and maturation process of the new viral particle. The PR cleaves Gag and the Gag-pol polyprotein precursor (Pr160) to produce functional

structural proteins and enzymes necessary for virion development and maturation (Pettit *et al*., 2004). A stable RNA structure and a heptanucleotide slippery sequence (UUUUUUA) mediates ribosomal frameshift at the intersection of the 3' end of Gag and Pol to change the reading frames for a different translation to obtain Gag and Pol functional proteins (de Oliveira *et al*., 2003; Mazauric *et al*., 2009). The structural information on the PR enzyme has led to the development of class of antiretroviral compounds as discussed in section 1.8.2, which are directed at inhibiting the HIV-1 PR activity.

1.2.5.3 Reverse Transcriptase

This is an enzyme encoded by the *pol* gene and consists of two subunits with a molecular weight of 66 and 51 kDa. It is found in all retroviruses and is necessary for the synthesis of new virion particles during a process known as reverse transcription. Three major biochemical activities characterize the nature of the Reverse Transcriptase (RT) enzyme and it includes RNA-DNA-dependent, DNA-dependent polymerase activity and ribonuclease H activity. The RT enzyme performs error-prone reverse transcription as it does not have 3'-5' exonucleolytic proof reading activity. These errors lead to numerous viral genomic point mutations generating new copies of HIV genomes with a diverse genetic variation (Oelrichs et al., 2000). The RT polymerase activity generates a complementary DNA (cDNA) from the single stranded genomic RNA found within the virion particle. During this time of cDNA synthesis, the RNase H removes the original RNA template that acted as the primer. The synthesized cDNA is a long and complex protein molecule that the RT enzyme further converts into a double helix form, which is the normal structure state of the

DNA (Götte *et al*., 1999). The RT enzyme performs error-prone reverse transcription as it does not have 3'-5' exonucleolytic proof reading activity. These errors lead to numerous viral genomic point mutations generating new copies of HIV genomes with a diverse genetic variation (Oelrichs *et al*., 2000). Some of the molecular assays such as the polymerase chain reaction (PCR) and DNA analyses, use RT from Avian Myeloblastosis Virus (AMV) and Moloney Murine Leukemia Virus Reverse (M-MLV) as the basis for cDNA synthesis. The AMV and M-MLV are also retroviruses like the HIV (Yasukawa *et al*., 2008). In HIV treatment, some of the commonly used cART, have been designed to target the RT region to stop further replication and transmission (Sarafianos *et al*. 2009, Men´endez-Arias, 2013). These classes of drugs will be discussed under cART drug mechanism action in section 1.2.10.

1.2.5.4 Integrase

The Integrase (IN) is a 32 kDa protein structure consisting of 288 aas. The IN is synthesized by the viral protease enzyme during the maturation process (Chiu & Davies, 2004). It plays a critical role in viral development by mediating integration of synthesized HIV-1 proviral viral DNA into host cell genome. (Mushahwar and Isa, 2007; Ombretta *et al*., 2012). During this integration, the Integrase (IN) displays three major activities of (1) two nucleotide trimming at 3' end of the linearized viral DNA, (2) cleaves the double stranded host cell DNA through endonuclease activity to allow integration of the trimmed viral DNA, and (3) providing a single covalent bonding at the integrated site through ligase activity (Bushman *et al*., 1990). The integration site is thought to be influenced by the accessibility of the chromosomal DNA found within the chromatin. Integration of the viral DNA into host cell DNA genome allows efficient

expression of viral genes. Current development of some antiretroviral class of drugs has been directed towards inhibiting this integration process, thereby limiting further production of new viral copies, as discussed in section 1.8.4.

1.2.5.5 Other structural genes

***gag* gene**: The *gag* gene codes for the Gag protein that consist of Matrix (MA, p17), viral capsid (CA, p24), [Spacer peptide1 (SPI)], P1), Spacer peptide 2 (SP2), Nucleocapsid (NC), P7) and P6 proteins. The MA provides viral integrity while the NC prevents nucleases from digesting the RNA. The *gag* gene plays a very important role in viral assembly (Gheysen *et al*., 1989).

env* gene:* The *env* gene forms glycoproteins that consist of a surface glycoprotein (gp120 SU), transmembrane glycoprotein (gp 41 TM), fatty lipid bilayer and host cell protein. The gp120 surface glycoprotein and gp41 are essential for viral attachment and fusion before the viral capsid enters the host cell for production of more viruses.

1.2.5.6 Regulatory, accessory and secondary protein structures

Regulatory proteins, Tat and Rev, are important players in gene expression and viral propagation as they target host cells through modulation of transcription and post-transcription. Accessory genes – the *vif, vpr, vpu* and *nef* are necessary for viral virulence and improve the efficiency of the viral replication process.

1.2.6 HIV-1 Replication cycle

HIV is present in body fluids of infected individuals as both free viral particles and virus within living cells. Furthermore, during HIV-1 replication cycle, the viral genome

is integrated into the host-cell genome (integrated provirus) that becomes a viral reservoir. The viral replication cycle is divided into seven stages as illustrated in Figure 1.5, starting with a mature HIV-1 particle recognizing and attaching onto the target cell, usually CD4+ T-cells. This attachment is mediated by interaction between the virus' extracellular domain of gp120 and the host cell receptors (Dash *et al*., 2008). Viral attachment is followed by additional interactions with chemokine receptors like the CCR5, CXCR4 or in some instances the CCR3 chemokine co-receptor (Chan and Kim, 1998; Clapham and McKnight, 2002) (step 1, Figure 1.5). The binding of gp120 to CD4+ T-cells results in a conformational change resulting in the fusion of both cell's membranes. This fusion leads to formation of viral synapses necessary for efficient entry of the viral genome and its enzymes into the host cell cytoplasm (Arthos *et al*., 2008) (step 2, Figure 1.5). In the cytoplasm, the viral single stranded RNA is converted to double stranded DNA by the viral transcribed RT enzyme (Zheng *et al*., 2005) (step 3, Figure 1.5). This stage of reverse transcription is extremely prone to errors as the enzyme does not have a 3' to 5' proof reading capability. These errors result in RAMs and allow for virus immune escape causing drug resistance and participating to viral diversity. The newly synthesized viral DNA (provirus) is imported to the nucleus by the IN enzyme and is covalently integrated into the host DNA (step 4, Figure 1.5). The IN inhibitors (INIs) interfere with integration of the provirus into the host cell DNA, blocking viral replication as will be described under mechanism of cART action in section 1.8.4. The provirus has the ability to lay dormant resulting in a latent infection; a barrier towards cure (Siliciano *et al*., 2003; Zheng *et al*., 2005). The DNA provirus in activated host cells is transcribed producing mRNA transcripts, which are capped at the 5' end and polyadenylated at

the 3' end (step 5, Figure 1.5). Three splicing events takes place on the mRNA transcripts resulting to completely spliced, singly spliced and unspliced mRNAs. The completely spliced mRNAs are translated into regulatory proteins, like the Tat and Rev. The Tat protein initiates transcription by attaching to the trans-activating response element (TAR) at the 5' LTR with association of other cellular proteins like cyclin T and cyclin-dependent protein kinase-9 for new virus production (Wei *et al*., 1998). The singly spliced mRNAs, which encode for Vpr, Vpu and Env viral proteins and unspliced mRNAs encoding for Gag (pr55), *Gag-Pol* (pr160) viral proteins, are exported to the cytoplasm by Rev. The *Env* precursor polyprotein (gp160) is synthesized in the endoplasmic reticulum (ER) from the singly spliced *env* mRNA and modified post-translationally in both the ER and Golgi apparatus (Wyatt & Sodroski, 1998). The gp160 is then cleaved by cellular proteases to produce glycoproteins gp120 and gp41 before transported to the plasma membrane for viral assembly (step 6, Figure 1.5). In the ribosomes, the Gag polyprotein (pr55 Gag) is synthesized using unspliced mRNAs as the message and with small amounts of Gag-Pol precursor proteins that have undergone a translational frameshift. They are incorporated onto the plasma membrane of the immature viral particle (Klimas *et al*., 2008). *Gag* molecules attract two copies of unspliced viral genome, encapsidating them onto the the plasma membrane of the already forming immature viral particle. The immature viral particle buds out of the host cell plasma membrane where cellular PR cleaves it, forming an infectious mature virion (step 7, Figure 1.5).

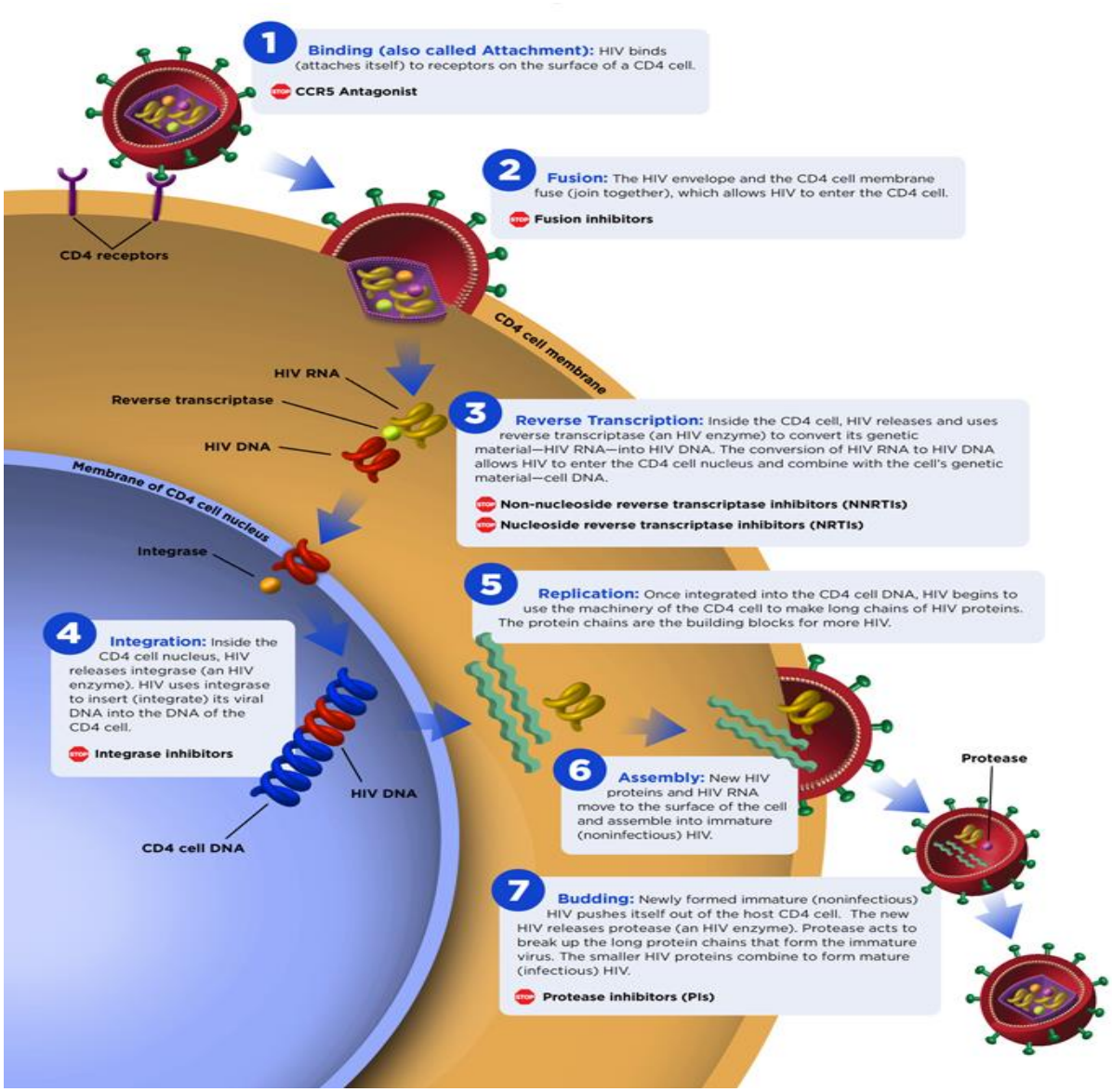


Figure 1.4 Illustration of HIV replication cycle showing HIV particle binding, fusion and entry into the target cell (CD4 cell). After entry, reverse transcription takes place followed by integration of the viral genome onto host genome. Replication, assembly and budding of the mature viral particle completes the life cycle, and the new virus particle is ready to infect other target cells. Current ART regimen commonly used target the reverse transcription (NRTI and NNRTI), Integration (IN inhibitors) and budding (PI) phases. (*https://aidsinfo.nih.gov/education-materials/glossary/1596/life-cycle*).

1.2.7 HIV-1 immunopathogenesis

HIV-1 immunopathogenesis involves a number of viral and host factors that interact in a complex manner resulting in HIV entry and the start of disease progression (Stevenson, 2003; Forsman & Weiss, 2008). The HIV-1 attaches to receptors on the surface of CD4+ T-cells and macrophages gaining entry onto the cell. The viral entry attachment can also be through the chemokine co-receptors (CCR5 or CXCR4). This attack on host immune cells results in a continuously activated immune system that eventually becomes compromised (Cunningham *et al*., 2010). Eventually, opportunistic infections and malignancies thrive, leading to AIDS (Davaro and Thirumalai, 2007). Clinically, HIV-infected patients, progress from an initial HIV infection to HIV/AIDS at different rates – the rates depends on various mechanisms of interaction between the patients' immunological factors and the viral regulatory and accessory proteins. These patients can be classified as rapid progressors (approximately 10.0%) developing AIDS within the first 2-3 years of infection or, typical progressors developing AIDS within a period of approximately 10 years from initial infection or long-term non-progressors (LTNPs) (<5.0%) who do not show symptoms at any time of the infection with stable CD4+ T-cells counts (Panteleo and Fauci, 1995; Levy, 2007). HIV infection has four distinct developmental phases: (1) primary infection or the seroconversion phase (2) clinically asymptomatic/latency phase (3) symptomatic and (4) chronic phases (Embretson *et al*., 1993; Koup *et al*., 1994). The primary phase lasts for a few weeks after initial infection and may be accompanied by a short flu-like illness, while many remain asymptomatic (Schacker *et al*., 1996). There is large amount of HIV-infected peripheral blood mononuclear

cells (PBMCs) and high viremia observed (Graziosi *et al*., 1993). The body defends itself by producing HIV antibodies and cytotoxic lymphocytes – marking the seroconversion process. HIV diagnosis can be missed before seroconversion is completed. The asymptomatic or latency phase is characterized by low viremia caused by increased specific cytotoxic activity. Patients don't show signs of damaged immune system, although it is believed that HIV continues to actively replicate especially in the lymph nodes. During the symptomatic phase there is progressive immune deterioration characterized by multi-system disease as the patient is vulnerable to opportunistic infections, marking the onset of AIDS (Panteleo and Fauci, 1995). The chronic stage of HIV-1 infection is the last stage and is characterized with very low CD4+ cell count and highly elevated viral load. Various opportunistic infections and malignancies are manifested, and the patient can succumb to death if not treated.

1.2.8 Natural resistance

Some individuals infected with HIV-1, especially the LTNPs, show partial or complete natural resistance against the virus (Hunt *et al*., 2009; Kumar, 2013). Resistance to HIV-1 infection has been attributed to natural host restriction factors like Tetherin, Apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3G (APOBEC3G) and Tripartite motif proteins (TRIM5α), antagonistic to viral proteins like Vif, Vpu, Nef and Envelope proteins. This antagonism limits viral replication, viral assembly, budding and infectivity. The APOBEC3G for example, inhibits provirus integration into host cell genome in Vif-free HIV-1. When viral particles without Vif protein bud, they encapsidate the inhibitory APOBEC3G proteins leading to excessive cytidine to

uridine hypermutation. This usually occurs in the newly infected cells and in the negative sense reverse transcripts (Sheehy *et al*. 2002; Malim, 2006*)*. In the positive cDNA strand a guanosine to adenosine mutations is registered and all this leads to functional imbalance between Vif and APOBEC3G proteins. Usually Vif and APOBEC3G functions are antagonistic to each other and any imbalance will negatively affect the viral fitness in infecting new cells (Arriaga *et al*., 2006). The TRIM5α motif is a protein factor that inhibits retrovirus replication (Stremlau *et al*., 2004). Therefore, a mutation in the SPRY domain of TRIM5α and other associated proteins has been shown to improve antiretroviral activity (Nakayama *et al*., 2005).

Natural resistance to HIV-1 has also been associated with three variants of CCR5 co-receptors (CCR5-Δ32, SDF1-3'A and CCR2-64I) (Nkenfou *et al*., 2013). The CCR5 is a chemokine receptor in the host cell involved in the binding of viral envelope proteins during viral entry. If a mutation occurs on the CCR5 protein it greatly reduces chances of viral attachment and entry into the host cell. Individuals who express homozygote state for CCR5-Δ32 appear to have natural resistance against HIV strains using the CCR5 co-receptors for entry into the host cell (Galvani *et al*. 2005; Hütter *et al*., 2009). The CCR5-Δ32 allele is high (5.0-15.0%) in Northern Europe, while mostly absent in Africans and East Asians (de Silva., 2004). Some individuals with Human Leucocyte group antigens like Antigen (HLA)-A2/6802 have reduced risk to HIV-1 infections (MacDonald *et al*., 2001). HLA-A2/6802 contains conserved immunogenic epitopes that are similar to HIV-1 peptide epitopes for T cell recognition. This means that both HLA-A2/6802 and HIV-1 compete for T cell recognition reducing HIV-1 susceptibility. High levels of cystatin is also thought to be

interfering with HIV replication as was described in Kenyan sex workers who tested HIV-1 negative despite being exposed to the virus multiple times. In persistently HIV-1 infected individuals, cysteine raises intracellular glutathione (GSH) level inhibiting HIV-1 replication. (Roederer *et al*., 1990).

1.2.9 HIV-1 treatment

At the start of HIV-1 treatment, antiretroviral drugs were first prescribed as monotherapy. Zidovudine (AZT) was the first monotherapy drug to be approved by the USA Food and Drug Administration (FDA) for HIV/AIDS treatment and was introduced in 1987 (Fischl *et al*., 1987). However, HIV-1 infections continued to rise leading to high mortality and morbidity rates. This meant that monotherapy was not effective and a search for alternative ways in ART management began. This led to the introduction of a cocktail of ART regimen referred to as “combination antiretroviral therapy (cART)” (Collier *et al*., 1996). A large array consisting of six different classes of anti-retroviral drug has been developed targeting various stages of HIV-1 replication (Figure 1.4) (Larder and Kemp, 1989; Palmisano and Vella, 2011). The six antiretroviral classes include viral entry inhibitors (CCR5 antagonists), fusion inhibitors, Nucleoside Reverse Transcriptase inhibitors (NRTIs), non-Nucleoside Reverse Transcriptase inhibitors (NNRTIs), Protease inhibitors (PIs) and Integrase inhibitors (INIs) also known as Integrase strand transfer inhibitors (INSTIs) (Pau and George, 2014) and are indicated in Table 1.1. These antiretroviral drugs have been approved by the FDA and are initiated with three drugs regimen/cART as recommended by WHO– 2 NRTIs + 1 NNRTI/PI as first line of treatment. A second line follows with also three drug combinations from the same or different classes in

case the patient does not respond well to the first line cART. Although cART does not eradicate HIV-1 infection completely, it has successfully improved recovery of the immune system, reducing HIV-1 transmission and protecting the patient against opportunistic infections (Murphy *et al*., 2001; Seth *et al*., 2016). However, research is ongoing looking into developing a single pill that has all three drug combined to increase patient adherence. Raltegravir (RAL, MK-0518) is the first INIs to be approved by the FDA in 2007, followed by Dolutegravir (DTG), and Elvitegravir (EVG, GS-9137) in 2013 and 2014 respectively.

Table 1.1 FDA approved classes of drugs and date of approval

Protease Inhibitors*	**Reverse Transcriptase inhibitors**		**Entry[a] and Infusion[b] inhibitors**	**Integrase inhibitors**
Amprenvir (APV) (October 20, 2003)	NRTIs	NNRTIs	CCR5-Antagonist Maraviroc[a] (August 6, 2007)	Raltegravir (October 12, 2007)
Atazanavir (ATV) (June 20, 2003)	Abacavir (ABC) (December 17, 1998)	Delavirdine (DLV)	Enfuvirtide, T-20[b] (March 13, 2003)	Elvitegravir (September 24, 2014)
Darunavir (DRV) (June 23, 2006)	Didanosine (ddI) (October 9, 1991 and October 31, 2000 for enteric coated)	Etravirine (ETR) (January 18, 2008)		Dolutegravir (August 13, 2013)
Fosamprenavir (FPV) (October 20, 2003)	Emtricitabine (FTC) (July 2, 2003)	Efavirenz (EFV) (September 17, 1998)		

Protease Inhibitors*	Reverse Transcriptase inhibitors		Entry[a] and Infusion[b] inhibitors	Integrase inhibitors
Iopinavir (LPV) (September 15, 2000)	Lamivudine (3TC) (November 17, 1995)	Nevirapine, (NVP) Viramune; (June 21, 1996) Viramune XR; (March 25, 2011)		
Indinavir (IDV) (March 13, 1996)	Tenofovir (TDF) (October 26, 2001)	Rilpivirine (RPV) (May 20, 2011)		
Nelfinavir (NFV) (March 14, 1997)	Zidovudine (AZT) (March 19, 1987)			
Ritonavir (RTV) (March 1, 1996)	Stavudine (d4T) (June 24, 1994)			
Saquinavir (SQV) (December 6, 1995)				
Tipranavir (TPV) (June 22, 2005)				

*PIs is usually prescribed with boosted RTV

Source: FDA-Approved HIV Medicines: Last Reviewed: (https://aidsinfo.nih.gov/understanding-hiv-aids/fact-sheets/21/58/fda-approved-hiv-medicines).

1.2.10 Mechanism of cART action

1.2.10.1 Viral entry/Fusion inhibitors

As the name suggests, entry inhibitors act in preventing viral entry into the host cell by blocking host cell receptors. Currently approved viral entry inhibitor drugs include Maraviroc (MVC): a CCR5 receptor antagonist, and a fusion inhibitor - enfuvirtide (ENF or T-20), which blocks fusion of the virus with the host cell membrane by binding to gp41 protein therefore preventing the conformational change necessary for membrane fusion (Lieberman-Blum *et al*., 2008). There have been challenges with developing viral entry drug inhibitors due to the evolving variable regions within the *env* gene.

1.2.10.2 Protease inhibitors

The protease inhibitors (PIs) are structure-based drugs designed to bind to the active site of the Protease enzyme inhibiting viral maturation. Alternatively, the PI is designed to mimic the transition state that occurs during peptide cleavage blocking the Protease activity (Lv *et al*., 2015). A number of RAMs, both major and minor, found in the *pol* PR genome region have been associated with reduced susceptibility to PI therapy. The major mutations in the *pol* PR genome have been associated with conformational changes that end up improving substrate selection by the viral PR enzyme. Likewise, minor mutations in combination with secondary mutations act as compensatory mutations that can improve viral fitness. Currently, there are about ten approved PIs by the FDA (Wensing *et al*, 2010) and they include first generation PIs - Indinavir (IDU), Ritonavir (RTV) and Saquinavir (SQV) which unfortunately were

found to have a high pill burden and low bioavailability. This necessitated the introduction of second generation of PIs boosted with ritonavir (PI/r) that have high genetic barrier to RAMS. This PI/r includes Atazanavir (ATV), Darunavir (DRV), Fosamprenavir (FPV), Tipranavir (TPV), Amprenavir and Lopinavir.

1.2.10.3 Reverse Transcriptase inhibitors

The Reverse transcriptase inhibitors (RTIs) specifically target and bind to the RT enzyme, thereby preventing viral replication. Depending on the mode of action, these drugs can be further classified into 2 classes; namely, Nucleoside Reverse Transcriptase Inhibitors (NRTIs) and Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs). NRTI analogues lack a free 3' hydroxyl group that terminates viral transcription of the newly synthesized DNA strand once they are incorporated by the viral RT enzyme. The NRTIs analogues include Abacavir (ABC), which is a guanosine analogue, Didanosine (ddI) - an analogue for Adenosine, Emtricitabine (FTC) and Lamivudine (3TC), which are cytidine analogues, Tenofovir disoproxil fumarate (TDF) - an adenosine analogue and Zidovudine (AZT), a thymidine analogue. The NNRTIs bind to a hydrophobic region near the active site of the RT enzyme, preventing the enzyme from synthesizing new DNA copies. A single mutation induced by any of the NNRTIs currently used, can confer cross-resistance against all NNRTI drug classes (Clavel and Hance, 2004). This problem has been countered by designing new NNRTI drugs like Etravirine (ETR), which are effective irrespective of the presence of NNRTI induced mutations. NNRTIs include Efavirenz (EFV), Etravirine (ETR), Nevirapine (NVP) and Rilpivirine (RPV).

1.2.10.4 Integrase inhibitors

The Integrase inhibitors (INIs) are a class of cART agents that inhibit HIV-1 replication by interfering with the IN enzyme activity (Hazuda *et al*., 2004; Evering *et al*., 2007). IN is a 288 aa enzyme that catalyzes the 3' end and strand transfer of the viral DNA during integration into the host cell DNA (Pommier *et al*., 2005; McColl and Chen, 2010). The INIs interfere with IN enzyme activity through strand transfer reaction by two mechanisms: (1) by interacting with two essential magnesium metal ion cofactors at the active site of the enzyme and (2) by binding to a specific complex between the enzyme and viral DNA (Grobler *et al*., 2002). This mechanism of action by INIs is specifically meant to block integration of viral DNA into the host cell DNA, thereby preventing viral replication. The FDA-approved INIs include Raltegravir (RAL, MK-0518, the first INIs to be approved by FDA in 2007), followed by Dolutegravir (DTG), and Elvitegravir (EVG, GS-9137) approved in 2013 and 2014, respectively (Steigbigel *et al*., 2008; Pandey, 2014; Kandel and Walmsley, 2015). These INIs seem to be less prone to the development of resistance (You *et al*., 2016).

1.2.11 HIV drug resistance (HIVDR)

Development of HIV drug resistance (HIVDR) has been a major challenge to the successful cART management globally (Hamers *et al*., 2012). HIVDR has been associated with the upscale of cART coverage, insufficient patient adherence and transmitted drug resistance (Gupta *et al*., 2012; WHO, 2017 - HIV drug resistance report, 2017: http://apps.who.int/iris/bitstream/10665/255896/1/9789241512831-eng.pdf?ua=1). This is further worsened, especially in resource-limited settings like

Africa, by minimal drug resistance testing, lack of monitoring of viral loads, improper cART prescriptions, failures in patient follow up, frequent drug stock outs and poor patient compliance. The emergence of HIV drug resistance results from RAMs that affect the ability of cART to effectively block HIV replication. RAMs can either be primary or secondary. Primary resistance also referred to as transmitted resistance occurs in the event whereby a new HIV-1 infection is caused by a drug resistant strain, whereas secondary or acquired resistance is seen when drug resistance develops while a patient is on cART (Grant *et al*., 2002; Bennett *et al*., 2009). RAMs are based on HIV-1 subtype B; although they occur similarly in all subtypes and some may occur more frequently than others (Westby *et al*., 2007). Furthermore, some of the RAMS have been found to be polymorphisms that naturally occur in non-subtype B leading to inconsistencies in reporting drug resistance rates especially in areas with high HIV-1 prevalence (Ross *et al*., 2007). Based on data available on HIV-1 subtype B, several RAMs are currently documented and interpreted by public database agencies. These agencies include Stanford HIVdb drug resistance interpretation algorithm (*https://hivdb.stanford.edu/hivdb/by-mutations/*); Los Alamos National Laboratories HIV Sequence database (www.hiv.lanl.gov, 2017); International AIDS Society (IAS) - USA Mutations Associated With Drug Resistance (Kantor *et al*., 2014; Wensing et al., 2017) and French National Agency for AIDS Research drug resistant algorithm (*http://www.anrs.fr/)*. There are some discrepancies in interpreting various RAMs. A standardized list of RAMs based on the HIV-1 *pol* gene *(*RT, PR and IN) has been assembled by WHO to act as a reference and is continuously updated (Shafer *et al*., 2007).

1.2.12 Drug resistance testing

Drug resistance testing assays currently performed include genotypic, phenotypic and tropism assays. These assays assist in monitoring the efficacy of cART used in the control of HIV replication or when switching to a new line of treatment. Genotypic assays in use such as GenoSurePRIme® offered by Monogram Bioscience provide predictions of resistance to all NRTIs, NNRTIs, PIs and INIs. This assay involves using genetic material of HIV-1 from patients' plasma sample to determine a pattern of HIV-1 mutations (Shafer, 2002; MacArthur, 2009). The mutations result from changes in the HIV-1 genome at specific aa positions and are represented in a code. For example, a code such as "K65R", represents a mutation where the first letter (in this case "K") represents Lysine present in the wild type, the number ("65") indicates the aa position in the gene of interest and the last letter represents the amino acid present in the mutant aa – in this case "R", which is Arginine aa. Genotypic assays are cheaper and have shorter turn-around time when compared to phenotypic assays; although it is not straightforward to correlate a single mutation and resistance to a specific drug (Van Laethem *et al*., 1999). RAMs are interpreted using different algorithms such as the Stanford drug resistance database (HIVDB), which is an online program that interprets genotypic resistance tests from the submitted gene sequences (Rhee S-Y *et al*., 2015). The HIVDB results are calculated as a penalty score and by adding these scores, the HIVDB algorithm is able to estimate the level of susceptibility, which can be interpreted as susceptible or low/intermediate/ high level of resistance. Phenotypic assays such as PhenoSense® GT and PhenoSense® Integrase offered by Monogram Biosciences are done in the laboratory as *in vitro*

assays. These assays measure viral responses against all antiretroviral drugs to determine the appropriate cART regimen for the individual patient (Hellmann *et al*., 1999). Tropism assays such as Monogram's Trofile® assays are used in the measurement of HIV-1 coreceptor tropism especially when CCR5 inhibitors are meant to be used (Aberg *et al*., 2014).

1.2.13 HIV infection in Cameroon

The first HIV-1 case in Cameroon was diagnosed 1985 (Garcia-Calleja *et al*., 1992; Mbanya *et al*., 2008). From there on, the HIV-1 pandemic has been progressively increasing with adult prevalence rates rising from 2.0% in 1990 to a 5.4% peak in 2000. However, in 2015 the prevalence rate reduced to an average of 4.5% (females at 5.6% and males at 2.9%) (Fonsah *et al*., 2017). This prevalence rate is highest in the West and Central African region. Countries like Côte d'Ivoire, Togo and Senegal have 3.7%, 2.4% and 0.5% prevalence rates respectively (*http://data.worldbank.org/indicator/SH.DYN.AIDS.ZS?locations=CM*). According to estimates by UNAIDS 2015, HIV/AIDS caused between 30,000 to 36,000 deaths and 290, 000 to 350,000 orphans aged 0 to 17 months [(UNAIDS, epidemic update December 2015 - www.unaids.org – accessed on 15th April 2016]. From 1990 - 2015, the country had about 44 000 new HIV infections with a 0.36% incidence rate for adults of 15 - 49 years [(AIDSinfo / UNAIDS, 2015: Global factsheet; Epidemiological slides: http://aidsinfo.unaids.org/)]. Cameroon harbours both HIV-1 and HIV-2, and the majority of HIV-1 groups, subtypes and CRFs are also found here (Faria *et al*., 2014; Agyingi *et al*., 2014). The HIV-1 group M subtype CRF02_AG is described as the most prevalent strain in Cameroon causing over 40.0% of HIV-1

infections (Aghokeng *et al*., 2013). Over 90.0% of new HIV-1 infections have been attributed to rampant risk sexual behaviour, such as having multiple partners and non-use of condoms, which are commonly practised in major cities and towns including Yaoundé and Douala. Mother-to-child transmission (also known as perinatal or vertical transmission) has contributed to approximately 6.0% of new infections; while blood transfusion, drug use and sexual contact have together contributed to the remaining 4.0% (UNAIDS Global report, 2016). The high prevalence of HIV-1 in Cameroon necessitates an upscale distribution of cART, started in 2003, to ensure ease of access of therapy to expectant mothers and children under the prevention of Mother-to-child transmission programs (PMTCT) (Billong *et al*., 2013). According to a recent report from UNAIDS, the Center for Disease Control (CDC) together with the Cameroonian government established 585 PMCTC clinics across various regions of the country [(UNAIDS/WHO, December 2015 – (www.unaids.org) – accessed on 4th May, 2016]. This has seen substantial progress in the implementation of the simplified WHO approach as the cART services have been decentralized (Aghokeng *et al*., 2013). In 2015, approximately 27.0% of the population had access to cART in Cameroon, an increase from 24.0% in 2014 (World Bank-(*http://data.worldbank.org/indicator/SH.HIV.ARTC.ZS?locations=CM*). Despite the huge success of cART in Cameroon, it has led to an increased prevalence of HIV drug-resistant strains. This has been attributed to insufficient patient adherence or incomplete suppression of viral replication during treatment, leading to the emergence of drug-resistant viruses (Gordon *et al*., 2015). In the therapy naïve population, HIV-1 drug resistance rates are approximately at 3.8% for adults and 3.6% in children and infants. HIV diversity, drug resistance and surveillance studies

are important tools for the implementation of cART, with the aim of managing the HIV pandemic. It is often difficult to study HIV resistance RAMs in resource limited settings like Cameroon owing to limited number of centers performing HIV-1 a resistance testing. Furthermore, diagnostic resistance testing remains costly to implement; justifying the limited amount of cART data in naïve and treated children and adults in Cameroon. The government of Cameroon has established resistance testing laboratory center at The Chantal Biya Hospital in the main city of Yaoundé. Although recommendable, more similar centers are required for wide coverage especially in the rural areas.

1.2.14 Aim and specific objective of the study

In this study, our aim and objectives were to analyze the diversity of HIV-1 and resistance patterns from infants/children and adults from the rural and urban surroundings of Yaoundé, Cameroon. HIV-1 diversity is thought to have an impact on the specificity and sensitivity of diagnostic assays, HIV-1 treatment, vaccine design and a possible cause for the emergence of drug resistance (Geretti *et al.,* 2009; Martinez *et al.,* 2009). It has also been suggested that different subtypes may be transmitted at different rates with different rates of disease progression (Baeten *et al.,* 2007; Kiwanuka *et al.,* 2008 and 2009). Despite the up-scale of cART in Cameroon since 2003 and other government initiatives like PMTCT programs, there are still many reported cases of new HIV-1 infections especially in pediatrics. This has called for continual surveillance of HIV-1 subtypes that could be responsible for the spread of the virus.

The specific objectives were;

1. To characterize and identify new HIV-1 circulating recombinant forms

2. Analyse drug resistance mutations from naïve and treated, infants/children and adults in Yaoundé, Cameroon.

CHAPTER 2 Material and Methods

2.1 Introduction

Cameroon is located in Central-West Africa, bordered to the South by the Republic of Congo, Gabon & Equatorial Guinea, to the West by Nigeria and to the Northeast by Chad and East by the Central African Republic. It is a politically stable country with approximately 24 million inhabitants (World Development Indicators (WDI), 2015). (http://knoema.com/WBWDIGDF2015Aug/world-development-indicators-wdi-september-2015?tsId=1555540). Yaoundé is the capital city with approximately 2.5 million inhabitants. The country has a complex cultural, social and geographical diversity, with French and English as the official languages (Fonjong, 2001). Our study focussed on Cameroon because it harbours a high viral diversity of HIV-1 subtypes and drug resistance rates in both Central and West Africa (Aghoken *et al*., 2013). This study consists of two unrelated groups of therapy naïve HIV-1 positive patients; (1) infants/children and (2) adults. From here on, infants/children will be assigned as 'Cohort I' and adults as 'Cohort II'. This chapter will focus on cohort I.

2.2 Ethical statement

Ethical clearance was obtained from the Cameroon National Ethical committee with reference number **049/CNE/SE/2013** (Appendix I) and was reviewed in 2014, 2015, 2016 and 2017. The study protocol has also been approved by the Health Research Ethics Committee at Stellenbosch University with the project reference number **N14/10/130**(Appendix II). The study was conducted according to the ethical guidelines and principles of WHO 2015, South African Guidelines for Good Clinical

Practice and the South African Medical Research Council (MRC) Ethical Guidelines for Research.

2.3 Study Design

This was a cross sectional study conducted between 2013 and 2017. The study participants enrolled at the Chantal Biya Foundation Hospital, the Centre for the Study and Control of Communicable Diseases (CSCCD) research institute - based at the Faculty of Medicine and Biomedical Sciences (FMBS) of the University of Yaoundé I. Other collection sites included District hospitals and other peripheral clinics within Yaoundé, Cameroon as shown in Figure 2.1. A summary of sample processing and analyses are described in Figure 2.2.

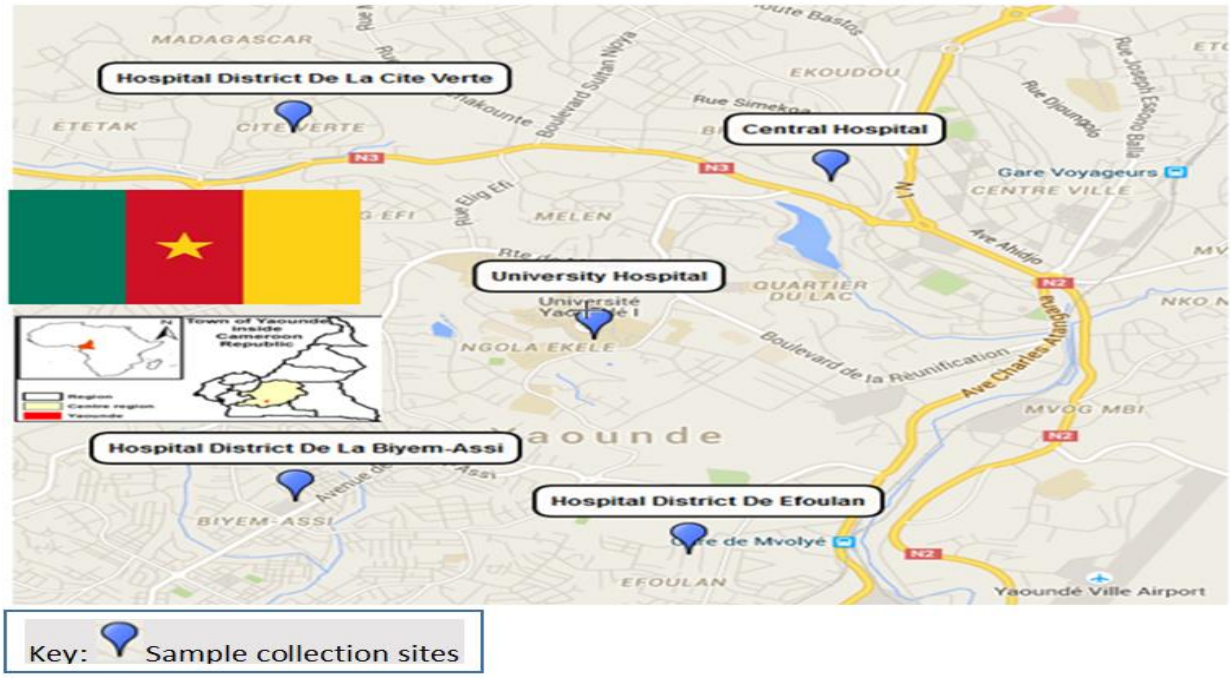


Figure 2.1 A map showing the different health districts around Yaoundé, Cameroon**.** The samples were obtained from the 5 different health districts, surrounding the CSCCD research institute and the University of Yaoundé I, Cameroon. The map was created with scribble maps (www.scribblemaps.com) (Ikomey *et al*., in press).

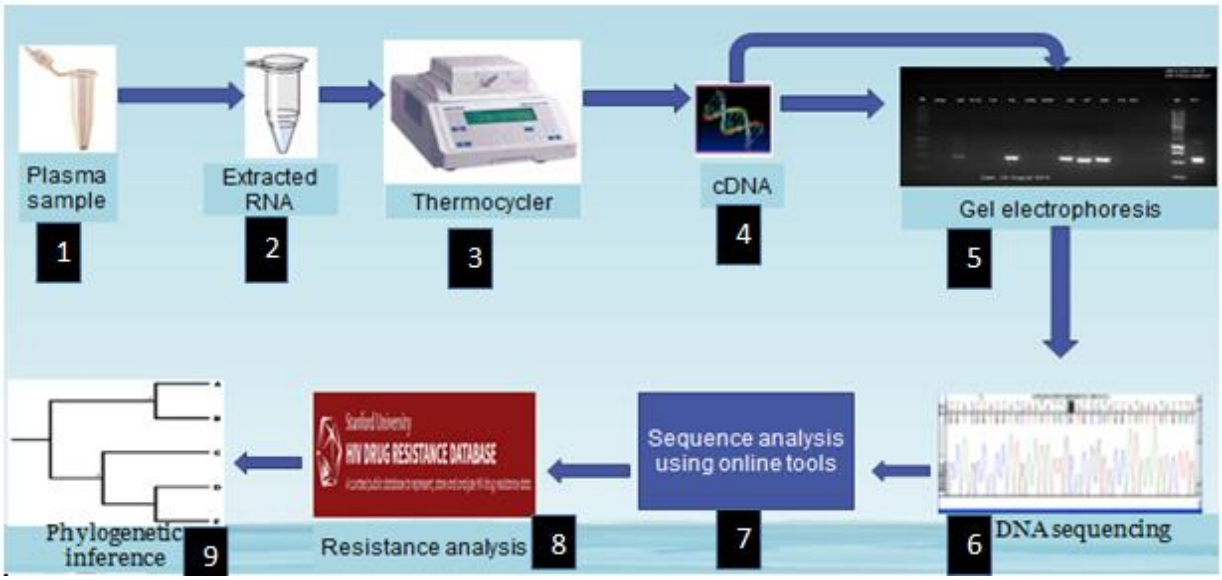


Figure 2.2 An illustration of procedures and molecular techniques used for sample processing and analyses in the study cohorts. Plasma samples were collected in their respective cryovial tubes (step 1) Viral RNA is extracted (step 2). The RNA sample is amplified using PCR (thermocycler) (step 3) to obtain a high yield of amplicons and converted into cDNA (step 4) in a reverse transcription process. Positive amplicons are visualized in agarose gel photo (step 5) and sequenced with consensus sequences in a chromatogram window (step 6). A quick analysis of HIV-1 subtypes is done using various online programs like REGA and RIP (step 7). HIVDR analysis is performed using Stanford drug resistance data base (step 8) and a Neighbor-Joining (NJ)phylogenetic tree drawn using MEGA version 5 (step 9) to cluster closely related HIV subtypes.

2.4 Study participants

They consisted of 55 therapy naïve HIV-1 positive infants/children aged 2 weeks – 72 months old. The infants in this study were under 28 days of age (according to WHO, 2017 - http://www.who.int/topics/infant_newborn/en/) while those above this age and up-to 18 years were defined as children All protocols for infant/children HIV screening were followed according to established consolidated 2015 WHO guidelines. The guidelines include: Patient consent (proxy consent for children),

confidentiality, counselling, correct test results and connection/linkage to prevention, treatment and care services.

2.5 Equipment, chemicals, reagents, commercial assays and enzymes

A list of equipment (Table 2.1), chemical reagents, commercial assays kits and enzymes (Table.2.2), miscellaneous items (Table 2. 3), software packages and HIV-1 subtype online programs (Table 2. 4) for this study are listed as follows:

Table 2.1 Equipment used for sample assays and analysis

Equipment	Application/Methods	Supplier	Source
Automated FACScount machine	CD4 estimation	Becton Dickinson, Bioscience	California, USA
GeneAmp PCR System 9700 thermal cycler	PCR reactions	Applied BioSystems	California, USA
Beckman Coulter Allegra ™ 6R centrifuge	Sample preparation	Beckman Coulter	California, USA
Microfuge® 16	Sample preparation	Beckman Coulter	California, USA
Vortex Mixer VM 300	Sample mixing	Gemmy Industrial Corp	Taipei, Taiwan
UVItec gel documentation system	Gel visualization and imaging	UVItec Limited	Cambridge, United Kingdom
Nanodrop™ ND 1000	Spectrophotometric measurement of DNA or RNA	Nanodrop Technologies Inc	Delaware, USA
ABI 3130xl genetic analyser	Sequencing	Gene Codes Corporation	California, USA Ann Arbor, Michigan, USA

Table 2.2 Chemical reagents, commercial assay kits and enzymes

Products	Supplier	Catalogue number
HIV Determine	Abbott Laboratories, USA	7D2649
ImmunoComb®	BiSpot, Orgenics	06G00063 v01
CD4 FACScount reagent	Becton Dickinson, Belgium	339010
ELISA kits, Integral Enzygnost®	Siemens, France	DP400
qRT-PCR AMplicor 1.5 commercial viral kit	Roche diagnostics, USA	236700791
QiaAmp viral RNA Mini Kit	Qiagen, Germany	151017169
SuperScript® III Taq High Fidelity	Invitrogen, Life technologies	12574-030 and 12574-035
GoTaq® G2 Flexi DNA Polymerase	Promega	M7805
Nuclease free water	QIAGEN, Germany	145045078
SeaKem® LE Agarose	WhiteSci/Lonza, USA	50004
GRGreen Nucleic Acid Gel Stain	Excellgen Ltd., South Africa	IV-1071
6x Blue Orange Loading Dye	BioLabs, USA	B7022S
GeneRuler™ 1kb DNA ladder	BioLabs, USA	0461505
QIAquick PCR Purification Kit	QIAGEN, Germany	28 106
BigDye™ Terminator cycle sequence ready Kit	Applied BioSystems, USA	4 337 035
5x Sequencing Buffer	Applied BioSystems, USA	4 305 603
BigDye XTerminator Purification Kit	Applied BioSystems, USA	4 374 408

Table 2.3 Software packages

Software package	References and/or licensed companies
Sequencher version 5	Gene Codes Corporation, USA
MAFT version 7	http://mafft.cbrc.jp/alignment/server/ (accessed on 15/04/2017)
BioEdit version 5.0.9	Hall, 2003. Ibis Biosciences, USA
MEGA version 5.0	Tamura *et al.,* 2011
Geneious	Biomatters Ltd, New Zealand

Table 2.4 Databases used for sequence analysis

Web Address	Databases
Los Alamos National Laboratory HIV Sequence Database	http://www.hiv.lanl.gov (accessed10/11/2016)
Stanford University- HIV Drug Resistance Database version 8.4	https://hivdb.stanford.edu/hivdb/by-mutations/ (accessed on 22/03/2017)
Los Alamos National Laboratory HIV Sequence Database- Quality Control	http://www.hiv.lanl.gov/content/sequence/QC/index.html (accessed on 22/11/2016)
Context-based Modeling for Expeditious Typing (COMET)	https://comet.lih.lu/ (accessed on 06/10/2016)
REGA HIV-1 Subtyping Tool - Version 3.0	http://dbpartners.stanford.edu:8080/RegaSubtyping/stanford-hiv/typingtool/- (accessed on 08/11/2016
Subtype classification using evolutionary algorithms (SCUEAL)	http://www.datamonkey.org/dataupload_scueal.php (accessed on 03/02/2017
jumping profile Hidden Markov Models (jpHMM)	http://jphmm.gobics.de (accessed on 12/02/2017
Recombinant Identification Program version 3 (RIP 3.0)	http://www.hiv.lanl.gov/content/sequence/RIP/RIP.html (accessed on 18/02/2017

Table 2.5 Miscellaneous items

Item	Supplying Company
Laboratory wear	Labotec, South Africa
1x TAE buffer	BioLabs, USA
Glassware	Labotec, South Africa
Laboratory liquids (Ethanol, Methanol)	Sigma-Aldrich, USA
Parafilm	Sigma-Aldrich, USA
Pipettes	Thermo Fisher Scientific, South Africa
Pipette tips	Thermo Fisher Scientific, South Africa
Sterile filters and filter papers	Thermo Fisher Scientific, South Africa
BD vacutainer® tubes	BD Biosciences, USA
Plastic material	Thermo Fisher Scientific, South Africa

2.6 Sample collection, preparation and storage

Study participants who consented / proxy-consented (for infants / children), completed a case record form. Briefly, 3-5 mL of whole blood from 55 therapy naïve infants / children, were drawn into Ethylene diamine tetra-acetic acid (EDTA) Becton, Dickinson (BD) Vacutainer® collection tubes (BD Biosciences, USA) using standard phlebotomy collection procedures. The blood samples were anonymized and given a unique study number. Plasma samples were separated by centrifuging the anticoagulated EDTA whole blood at 2000 revolution per minute (rpm) for 10 minutes at 18-25˚C room temperature using a Microfuge® 16 centrifuge (Beckman Inc, USA). The collected plasma sample was aliquoted into two separate 5 mL cryogenic vials

(Rotest, China) with one tube of plasma kept at -20°C for serological testing and viral load screening at the CSCCD laboratory. The other plasma sample tubes were safely packaged under -20°C conditions and shipped to the Division of Medical Virology, Tygerberg Academic Hospital - Stellenbosch University, South Africa for diversity studies and drug resistance testing.

2.7 HIV testing

HIV Serology status were confirmed based on the Cameroon national algorithms which involves the use of two HIV rapid diagnostic test and a confirmation with an antigen p24 antibody detection ELISA kits. The Alere Determine™ HIV–1/2, (Abbott Laboratories, USA) and the SD Bioline HIV 1/2 version 3.0 (Standard Diagnostics, Inc.) were used with an ELISA antigen antibody kit (Integral Enzynost ELISA kits, Siemens, France). All tests were carried out following the manufacturer's guidelines.

2.8 Determination of Cluster of differentiation 4 Thymus-cell counts and percentages

The CD4 cell counts was done using standard methods following manufacturer's instructions.

Principle: The FACS count system consists of a pair of reagent tubes containing fluorochrome-conjugated monoclonal antibodies and reference beads for calculation of absolute lymphocyte values. The system is designed to use EDTA anticoagulated whole blood. When whole blood is added into FACS count reagent tube, the antibodies bind specifically to lymphocyte cell surface antigens forming a cluster/population of cells. These cluster cells fluoresce different colors and the BD

FACS count machine detects two colors and measures the relative size of the cell. The cluster CD4+ and CD8+ T-cells fluorescence yellow, while CD3+ cells fluoresce red colour. Five percent (5%) formaldehyde fixative is added to the sample/reagent mixture for sample stability and as a biohazard protection. Using an internal software algorithm (FACS count Software v 1.0, 4/93), the BD FACS count instrument enumerates the absolute values of the CD4+, CD8+ and CD3+ cell count.

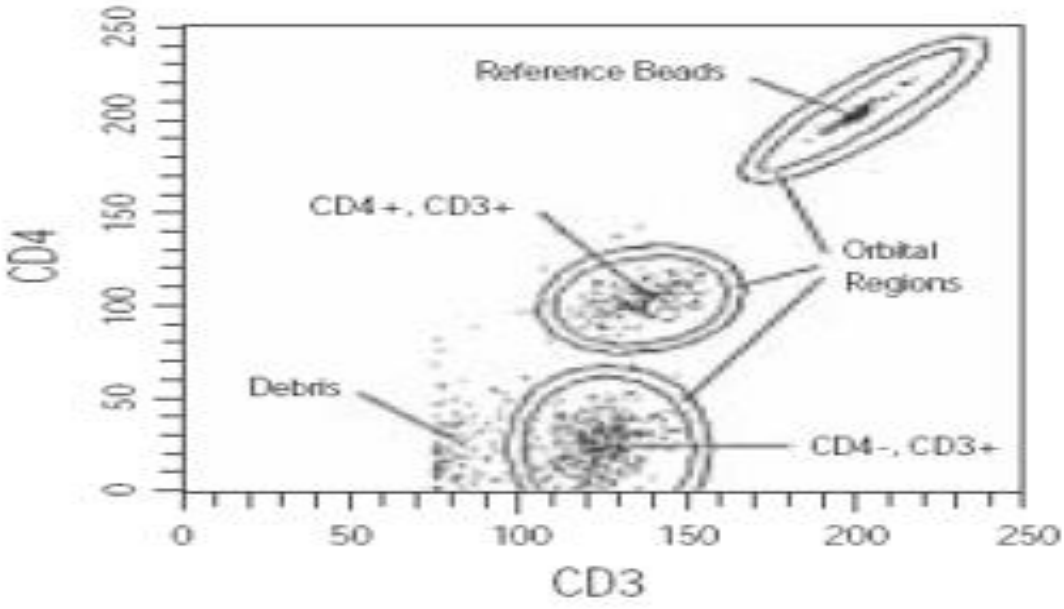


Figure 2.3 A graphic representation of the Becton Dickinson FACS count software algorithm evaluating cluster of CD4+ T-cells with reference beads using an elliptical orbit region that localizes on the center of each population. Sufficient number of CD4+ T- cells (up to 30 000) are collected to ensure statistically accurate results. [(http://resources.psmile.org/resources/equipment/specific-equipment/facscount/Equ2.302%20FACSCount%20System%20Overview.pdf/view)]

Method: Fifty-five microLitre (50µL) of Whole blood collected in EDTA tubes were used for CD4 absolute and percentages counts. This was based on the principle of immunophenotyping using an automated Fluorescence Activated Cell Sorting (FACS) Count Analyzer using the Becton Dickinson (BD) FACSCount trio CD4/CD8/CD3 reagent kit (BD Biosciences, USA). Samples were analysed based on the manufacturers' guidelines including the quality controls as illustrated in Figure 2.3 and previous studies (Ikomey *et al*., 2016).

2.9 RNA extraction

HIV RNA was extracted from plasma samples using the QIAamp Viral RNA Mini Kit (Qiagen, Germany) (Figure 2.3).

Principle: The QIAamp Viral RNA Mini procedure extracts all RNA molecules larger than 200 nucleotides. The sample is first lysed under highly denaturing conditions to inactivate RNases and to ensure isolation of intact viral RNA. Buffering conditions are adjusted to provide optimum binding of the RNA to the QIAamp membrane and the sample is loaded onto the QIAamp Mini spin column. The RNA binds to the membrane and contaminants are efficiently washed away in two steps using two different wash buffers. High-quality RNA is eluted in a special RNase-free buffer, ready for direct use or safe storage. The purified RNA is free of protein, nucleases and other contaminants and inhibitors.

Method: Working under a biosafety cabinet (HR60-IIA2, United Scientific (PTY) Ltd, South Africa), we added 140µL of equilibrated plasma into 560µL of prepared viral lysis buffer (Buffer AVL) containing carrier RNA into a 1.5mL microcentrifuge tube

(ThermoFisher, USA). Complete sample lysis was achieved by pulse-vortexing the mixture using the Vortex Mixer VM 300 (Gemmy Industrial Corp, Taiwan), followed by a 5 minutes incubation period at room temperature. The lysed sample was pipetted into the QIAamp mini column so that the viral RNA could bind onto the QIAamp mini column membrane. This was followed by two wash rounds with prepared wash buffer 1 (buffer AW1) and wash buffer 2 (buffer AW2) to remove contaminants that would otherwise lower the purity of the RNA elute. Pure RNA was eluted using RNase free Viral Elution buffer (buffer AVE). Centrifugation was done using the Microfuge® 16 (Beckman Inc, California, USA) and all protocols strictly observed the manufacturer's instructions to avoid any contamination.

2.10 Viral load screening

The quantitative qRT-PCR Amplicor1.5 commercial viral load kit was used to determine HIV-1 viral load Test (Roche Diagnostic Systems, USA). The assay was performed with a prepared master mix containing oligonucleotide primers specific to regions of the HIV-1 *gag* gene with upstream sense primers of Sk145 (1359 – 1388 in HXB2 gene map location) - 5'-ACTGGGGGGACATCAACCACCCATGCAAT-3′ and the downstream antisense of SKCC1B (1486-1513 in HXB2 gene map location) - 5'-TACTAGTAGTTCCTGCTATGTCACTTCC-3'. The limit of detection of the assay was 67log10 or less than 40 copies /mL.

2.11 RNA concentration and quality estimation

This was done to detect the presence of other proteins or contaminants/co-purified contaminants. It also helped us to determine the right amount of RNA sample to use

for our downstream procedures. We used 1µL of RNA sample to determine RNA concentration and quality spectrophotometrically in ng/µL using the NanoDrop® ND-1000 V3.1.0 (Nanodrop Technologies Inc, USA). As per the manufacturer's instructions, RNA absorbance was measured at a wavelength of 260nm to determine the RNA concentration. RNA purity was determined by a ratio of sample absorbance at 260nm and 280nm with accepted RNA purity at approximate ratio of 2.0. Any RNA sample of a ratio 1.8 to 2.2 was generally accepted as of good quality and hence forth used for our next procedure.

2.12 Location of genome of interest

For HIV-1 molecular characterization and detection of RAMs, we used PR, RT and IN fragments of the HIV-1 *pol* gene. The *pol* gene location is in relation to HXB2 gene map below Figure 2.4.

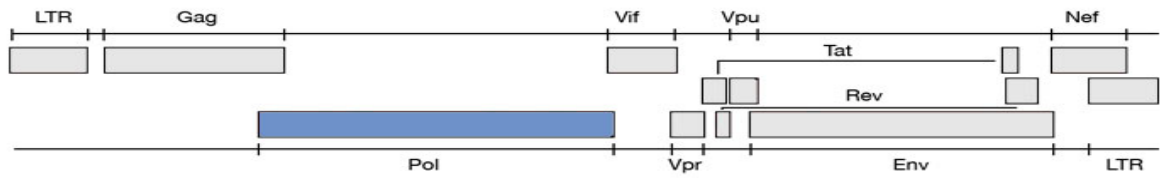


Figure 2.4 HXB2 gene map showing the *pol* genome (colored blue) – available at: http://www.bioafrica.net/proteomics/POLprot.html.The partial *pol* gene (PR) at HXB2 location 2136 – 2650 for PR; 2530 – 3334 for RT and 4470 – 4806 for IN were amplified and sequenced for genotypic typing and drug resistance testing.

2.13 Primers for PCR amplification

Primers for PR and RT amplification were used as previously described (*Plantier et al., 2005; Jacobs et al., 2008*). Primers for the PR fragment are indicated in Table 2.6

and primers for the RT fragment are indicated in Table 2.7; Primers for Integrase (IN) as described (Miura *et al*., 1990) are indicated in Table 2.8. Annealing temperatures for all primers are also given.

Table 2.6 PR Fragment amplification primers

Primers	Annealing Temperature	Oligonucleotide sequence (5'-3')	HXB2 Location	Size (bp)
Outer/Reverse 30'prot1	55^0C	GCAAATACTGGAGTATTGTATG GATTTTCAGG	2703-2734	652
Outer/Forward 50'prot1	55^0C	TAATTTTTTAGGGAAGATCTGGC CTTCC	2082-2109	
Inner/Reverse 30'prot2	55^0C	AATGCTTTTATTTTTTCTTCTGTC AATGGC	2621-2650	514
Inner/Forward 50'prot2	55^0C	TCAGAGCAGACCAGAGCCAACA GCCCCA	2136-2163	

Table 2.7 RT fragment amplification primers

Primers	Annealing Temperature	Oligonucleotide sequence (5'-3')	HXB2 Location	Size (bp)
Outer/Reverse Mj3	55^0C	AGTAGGACCTACACCTGT CA	2480 -2499	940
Outer/Forward Mj4	55^0C	CTGTTAGTGCTTTGGTTC CTCT	3399-3420	
Inner/Reverse A35	60^0C	TTGGTTGCACTTTAAATT TTCCCATTAGTCCTATT	2530- 2564	804
Inner/Forward NE135	60^0C	CCTACTAACTTCTGTATG TCATTGACAGTCCAGCT	3300-3334	

Table 2.8 IN fragment amplification primers

Primers	Annealing Temperature	Oligonucleotide sequence (5'-3')	HXB2 Location	Size (bp)
Outer/Forward UNIPOL1	50^0C	TGTCAACATAGTAACA GATTCACAATA	4025-4051	953
Outer/Reverse UNIPOL2	50^0C	ACTACTGCCCCTTCAC CTTTCCA	4956-4978	
Inner/Forward UNIPOL3	50^0C	AGTGGATTCATAGAAG CAGAAGT	4470-4492	336
Inner/Reverse UNIPOL4	50^0C	CCCCTATTCCTCCCCTT CTTTTAAAA	4781-4806	

2.14 Polymerase Chain Reaction

Polymerase chain reaction (PCR) was done in two reactions – Pre-nested and Nested PCR. In the pre-nested PCR, the targeted RNA fragment was reverse transcribed to cDNA. This was followed with amplification of the cDNA to dsDNA amplicons using gene specific primers and SuperScript® III (SSIII) One-Step Reverse Transcriptase-PCR System with Platinum® *Taq* High Fidelity enzyme protocol (Invitrogen, USA). According to the manufacturer's instructions, the SSIII - RT system detects RNA targets from 300bp to 10kb and very little amount of RNA ranging from 1pg to 1µg can be used. The second PCR reaction (Nested PCR) involved amplifying the DNA PCR products from the first PCR assay using gene specific primers and GoTaq® G2 Flexi DNA Polymerase Protocol (Promega, USA). The GoTaq® assay kit amplifies a wide range of PCR fragments, consistently producing high amplicon yields that are highly reliable. The target PCR amplification

products for each fragment were 514bp, 804bp and 336bp for PR, RT and IN respectively.

Principle: A PCR amplification assay operates under optimally set temperature conditions and cycles to obtain a successful yield. Three major concurrent steps consisting of heat denaturation, primer annealing to the template and elongation of the target RNA/DNA fragment are involved. The PCR amplifies exponentially and expected target copy numbers is given by the formula; 2^n, where n = number of PCR cycles and the resultant copies range in billions.

2.14.1 Pre-nested RT- PCR

This PCR reaction targeted HXB2 location 2082 – 2734 (652bp) for the PR fragment while 2480 – 3420 (940bp) and 4025 – 4978 (953bp) were targeted for the RT and IN fragments respectively. To obtain the PCR pre-nested products, the SSIII kit protocol (Invitrogen, USA) was used for initial amplification for all 3 HIV-1 *pol* fragments (PR, RT and IN). Specific outer target primers with their specific annealing temperatures used are indicated above in Table 2.6 for PR, Table 2.7 for RT and Table 2.8 for IN fragments. The RNA was first reverse transcribed to cDNA using following manufacturer's instructions. All PCR amplification steps were done using the GeneAmp PCR System 9700 thermal cycle (Applied Biosystems, USA). Briefly, a master mix of a 50µL reaction volume (shown in Table 2.9) was prepared as follows: 25 µL of 2x reaction mix buffer was added with 1 µL of 10µM specific forward primer and 1µL of 10µM specific reverse primer (as indicated above in Table 2.6, 2.7 and 2.8), 1µL of SSIII enzyme mix, 5 µL of 1pg -1µg total RNA and Nuclease free water

(QIAGEN GmbH, Germany) filling up to the 50 volume in a 0.2 mL thin-walled PCR tube. We used a standard protocol of PCR amplification as indicated in the Table 2.10.

Table 2.9 A master mix for pre-nested PCR in a 50 µL reaction volume reaction

Component	Volume
2x reaction Mix	25 µL
Outer/Forward primer,10 µM as indicated above in Table 5 for PR, Table 6 for RT and Table 7 for IN fragments	1 µL (0.5mg/µL)
Outer/Reverse primer,10µM as indicated above in Table 5 for PR, Table 6 for RT and Table 7 for IN fragments	1 µL (0.5mg/µL)
SuperScript® III RT/Platinum® *Taq* High Fidelity Enzyme Mix	1 µL
Template RNA (1 pg- 1 µg)	5 µL
Nuclease Free water	17 µL
Total	**50 µL**

Table 2.10 Cycling conditions for Pre-nested PCR

Step	Temperature	Time	Hold/Cycle
Initialization and cDNA synthesis	50^0C	30 minutes	1 hold
pre-denaturation	94^0C	2 minutes	
Denaturation	94^0C	15 seconds	
Primer annealing	55^0C for PR and RT 50^0C for IN	30 seconds	40 cycles
Extension	68^0C	1 minute/kb	
Final extension	68^0C	5 minutes	1 hold
Product was Cooled and stored at 4^0C until used			

2.14.2 Nested PCR

This was the second PCR reaction using GoTaq® G2 Flexi DNA Polymerase Protocol (Promega, USA). The PCR products amplified was 2136 -2650 (514bp), 2530 – 3334 (804bp) and 4470 – 4806 (336bp) in HXB2 location for PR, RT and IN respectively. All the nested PCR steps were done using the GeneAmp PCR System 9700 thermal cycle (Applied Biosystems, USA). Briefly, a master mix of a 50µL reaction volume as shown in Table 2.11 was prepared as follows; 10 µL of 5x GoTaq® Flexi buffer1 was added into 4µL of 25mM $MgCl_2$ Solution, 1µL of dinucleotide Triphosphates (dNTPs) mix (10mM each), 1µL of 0.1-1.0µM inner/forward primer and 1µL of 0.1-1.0µM inner/reverse primer (as indicated above in Table 2.6, 2.7 and 2.8), 0.5µL of 5u/µL GoTaq® Flexi DNA Polymerase, 5µL of DNA template and nuclease free water filling the required 50 volume in a 0.2 mL thin-walled PCR tube. A standard protocol for cycling conditions was used as indicated in Table 2.12.

Table 2.11 A master mix for Nested PCR (PR fragment) in a 50 reaction volume

Component	Volume
5X colorless GoTaq® Flexi Buffer[1]	10 µL
MgCl2 Solution, 25mM[1]	4 µL
PCR Nucleotide Mix (dNTPs), 10mM each	1µL
Inner/forward primer (10 µM) as indicated above in Table 5 for PR, Table 6 for RT and Table 7 for IN fragments	1µL(0.5mg/µL)
Inner/Reverse primer (10 µM) as indicated above in Table 5 for PR, Table 6 for RT and Table 7 for IN fragments	1µL(0.5mg/µL)
GoTaq® G2 Flexi DNA Polymerase (5U/µL)	0.5 µL
Template DNA(cDNA)	5µL

Table 2.12 Cycling conditions for nested PCR

Step	Temperature	Time	Hold/Cycle
Initialization	95^0C	2 minutes	1 hold
Denaturation	95^0C	1 minute	30 cycles
Primer annealing	60^0C for PR and RT 50^0C for IN	1 minute	
Extension	72^0C	1 minute/kb of DNA	
Final extension	72^0C	5 minutes	1 hold
Cooling and storage at 4^0C until used			

2.14.3 Visualization of the nested PCR amplicons

Agarose gel electrophoresis was used to separate and identify the nested PCR amplicons.

Principle: Agarose gel electrophoresis separates macromolecules such as DNA by size (length in base pairs) in a matrix of agarose under an electric field with a constant voltage and a predetermined time for optimal results. The macromolecules move from a negative to a positive electrode separating the samples (Sambrook *et al*., 2001; Yilmaz *et al*., 2012).

Method: The gel was prepared according to the method described by Sambrook *et al.,* 2012 as follows; one gram of agarose powder (Seakem LE® agarose: FMC Bioproducts, USA) was dissolved in 100mL of 1x Tris Acetate EDTA (TAE) buffer to obtain a 1% weight per volume (w/v) agarose gel concentration. The melted agarose was cooled to 55^0C and 10µL of GRGreen gel stain (Excellgen Ltd, South Africa) was

added before the gel was cast onto a tray measuring 7 x 10 cm. Three microliter (3µL) of PCR product was mixed well with 5µL of Orange (6X) gel loading dye (2.5% Ficoll and 0.15% Orange (6X) (BioLabs, USA) and loaded into sample wells with positive and negative controls. The GeneRuler^{TM}1kb DNA ladder (Promega, USA) was used as a molecular weight marker. The gel ran at 75 Volts for approximately 45 minutes and was visualized under ultra violet (UV) light with a wavelength of between 254nm-365nm using the UVItec gel documentation system (Cambridge, United Kingdom). Expected positive bands of fragment size 514bp for PR, 804bp for RT and 336bp for IN were identified, photographed and the corresponding PCR products used for purification and sequencing.

2.14.4 Purification of PCR products

The QIAquick PCR purification kit (Qiagen, Germany) was used to purify the PCR products as per the manufacturer's instructions.

Principle: Nucleic acids binds to the QIAquick spin column membrane under buffering conditions (10 mM Tris·Cl, pH 8.5) and impurities are washed down through centrifugation at 8000 revolutions for 10 minutes (Sambrook *et al*., 2001). This is followed by elution of purified RNA under the same buffering conditions and an aliquot of the purified RNA product is obtained and stored at -20^{0}C to prevent any contamination especially through RNase activity.

Method: PCR amplified products (double stranded DNA) were purified from excess primers, enzymes, nucleotides, polymerases and salts used during PCR amplification

and gel separation. The eluted purified DNA was kept at -20^{0}C for the next procedure, to prevent degradation.

2.15 DNA sequencing

DNA sequencing was developed by Fredrick Sanger and colleagues in 1977 (Sanger *et al*., 1977). It determines the order of the four bases - ATCG in the targeted DNA sample through a capillary process.

Principle: The sequencing method is based on selective chain-termination of modified dideoxynucleoside triphosphate (ddNTPs) in an enzymatic reaction. The four dNTPs are enzymatically added by the DNA polymerase forming a long chain of DNA fragment. Once the DNA fragment is formed, the fluorescently labelled ddNTPs terminate the reaction. This is because the ddNTPs lack 3'- OH group which is necessary for the formation of a phosphodiester bond between two nucleotides. The results are interpreted by the genetic analyzer software by calling the bases at each data point depending on the intensity of emitted fluorescence from the ddNTPs.

Method: A master mix of 10µL reaction volume was prepared as follows; 2µL of DNA template was added into 1µL of BigDye^{TM} X Terminator (Applied Biosystems, USA) Enzyme mix, 2µL of 5x sequence buffer, 1µL of each primer (Table 2. 13) and 4µL of nuclease free water to fill-up the required volume in a 0.2mL thin-walled sequencing tube. The reaction cycle was run as described below in Table 2.14. The products were purified using the BigDye^{TM} XTerminator Cycle Sequencing Ready Reaction Kit and analysed using the automated ABI 3130xl genetic analyser (Applied Biosystems, USA).

Table 2.13 Sequencing primers

PR Primers	Annealing Temperature	Oligonucleotide sequence (5' -3')	HXB2 Location	References
JA217-Reverse	60°C	CTTTTATTTTTTCTTCTGTCAATG	2623-2646	Lindström and Albert, 2003
50 PROT 2-Forward	60°C	TCAGAGCAGACCAGCAGCCCCA	2136-2163	Plantier *et al*., 2005
RT Primers				
AK 11-Forward	60°C	GTACCAGTAAAATTAAARCCAG	2571-2592	Lindström and Albert, 2003
*pol 3D-Forward	60°C	AATGTTGCACTTTAAATTTATCC	2529-2552	Loxton *et al.,* 2004
AK 10-Reverse	60°C	TYCCCACTAAYTTCTGTATRT	3316-3336	Plantier *et al*., 2005
NE 135-Reverse	60°C	CCTACTAACTTCTGTATGTCATTGT	3310-3334	Plantier *et al*., 2005
IN primers				
Poli 3	55°C	TAGTGGGATGTGTACTTCTGAAC	5195-5217	Miura *et al*.,1990
Poli 4	55°C	CACACAAAGGRATTGGAGGAAATG	4162-4185	Miura *et al*., 1990

Table 2.14 Sequence PCR cycling parameters

Step	Temperature	Time	Number of cycles
Pre-denaturation	96^0C	1 minute	1
Denaturation	96^0C	10 seconds	25
Primer Annealing	55^0C	7 seconds	
Elongation	60^0C	4 minutes	
Cooling and storage at 4^0C until used			

2.15.1 Quality control of DNA sequences

Quality DNA sequences obtained were checked to exclude contamination, detection of phylogenetic signal and recombinants in HIV-1 using the Los Alamos sequence database (*http://www.hiv.lanl.gov/content/sequence/QC//index.html*).

2.15.2 Multiple sequence alignment

Sequence alignment allows for all homogenous nucleotide sequences to be compared and aligned. The sequence alignment uses a heuristic approach whereby sequences are aligned progressively following a guide tree (Feng & Doolittle, 1987). In this study, sequences were checked for presence of multiple stop codons, frameshifts and hypermutations before aligning them. Multiple alignments of nucleotides were aligned using MAFFT version 7 (*http://mafft.cbrc.jp/alignment/server/*).

2.16 HIV-1 subtyping using online programs

A variety of online HIV-1 programs were used for preliminary HIV-1 subtype and recombinant characterization. These were, Recombinant Identification Program version 3 (RIP 3.0) (*http://www.hiv.lanl.gov/content/sequence/RIP/RIP.html*; Holguin *et al*., 2008), REGA HIV Subtyping Tool version 2 (*http://www.bioafrica.mrc.ac.za/rega-genotype/html/subtypinghiv.html)*; Alcantara *et al*., 2009), Subtype classification using evolutionary algorithms (SCUEAL) (*http://www.datamonkey.org/dataupload_scueal.php)*; Delport *et al*., 2010), Context-based Modeling for Expeditious Typing (COMET) (*http://www.comet.retrovirology.lu*;Struck *et al*., 2014) and the jumping profile Hidden Markov Models (jpHMM) webserver (*http://www.jphmm.gobics.de)*; Schultz *et al*., 2009; Bulla *et al*., 2012).

2.17 Drug resistance analyses

RAMs analyses was performed using the Stanford HIV drug resistance data base (HIVDB) version 8.4 algorithm, available at *http://hivdb.stanford.edu*. The HIVDB was preferred because it mostly targets genotypic data from the *pol* genome (PR, RT and IN). The data collected can then be easily correlated with treatments of individuals through genotypic drug resistance assays. The Stanford HIVDR is easily accessible through HIVDB data base and is continuously updated to reflect emerging RAMs against FDA approved antiretroviral drugs (Kantor *et al*., 2001; Shafer, 2006; Wang *et al*., 2007).

2.18 Phylogenetic inference

Information from DNA sequences is converted into evolutionary trees using molecular phylogenetic analyses (inference). The topology of the tree is based on similarities and/or differences between the DNA sequences. Thus, the more closely they appear the more similar they are. As the HIV-1 evolves rapidly, a phylogenetic tree is better placed to give detailed information, especially on viral origin, diversity and spread, which is important in managing the HIV-1 pandemic. To make the best estimate of the evolutionary history that is based on the inheritance of ancestral characteristics of DNA sequences obtained from HIV-1, a phylogenetic inference was used based on the inheritance of ancestral characteristics. In this study cohort, a Neighbor-Joining (NJ) phylogenetic tree was inferred using MEGA version 5.0 (Tamura *et al*., 2011) as follows – the contigs obtained from chromatograms were assembled using Sequencher version 5 (Gene Codes Corporation, USA) and HXB2 sequence was used as a reference. Sequence contigs were obtained and edited using BioEdit version 5.0.9 (Ibis Biosciences, USA). Multiple sequence alignments (MSA) of sample nucleotide sequences were created using MAFFT version 7 (*http://mafft.cbrc.jp/alignment/server/*). Gaps were removed from the final alignment before performing a phylogenetic inference. A jModelTest was used to determine the "best model" for phylogenetic tree reconstruction. This was done by loading the MAFFT aligned DNA sequences file unto the jModelTest folder and analysis done by computing likelihoods. The "best model" determined for this study was the TIM1+G model (with 012230 partition and 0.2820 gamma distribution) that was selected under the Bayesian Information Criterion (BIC). The MEGA version 5.0 that was used for

phylogenetic analysis does not have the TIM1+ G model so the closest model with similar parameters was Tamura 3 model (T92) was selected. Reliability of the branching order of the phylogenetic trees was tested using bootstrapping analyses of 1000 replicates for all datasets.

CHAPTER 3 Results – Cohort I

This section describes the social-demographics (gender and age) and clinical data of the 55 study participants in this cohort as tabulated in Table 3.1. The clinical data includes the CD4 and viral load measurements that were taken as part of the diagnostic assays. Amplification and sequencing was successful in 29/55 (52.7%), 27/55 (49.1%) and 28/55 (50.9%) for PR, RT and IN fragments as summarized in Table 3.2.

3.1 Social-demographics and clinical data

Fifty-five infants/children were recruited for the study [32 (58.0%) females and 23 (42.0%) males]; with an age range of 9 months to 6 years and a mean age of 2.5 years old. CD4 percentages and absolute counts ranged between 30.0% and 44.0% (median 33.0%, 500-2000 cells/m3, with a mean of 957). HIV-1 plasma RNA viral loads ranged between 4.6 and 5.87 log10 copies per milliliter (median 4.96). The social-demographic information and clinical data are summarized in Table 3.1.

Table 3.1 Summary of infants/children clinical and demographic data

parameters	Frequency (n=55)	Percentage
Gender		
Male	23	42
Female	32	58
Settings		
Rural	41	74
Urban	14	26
Treatment		
Naïve ART Pregnant mothers	34	62
ART exposed pregnant mothers	21	38
	Mean	**Range**
Age (Years)	2.5	0.8 – 6
CD4 Count (absolute)	958	500 – 2000
CD4 (%)	36	15 – 44
Viral Load (log_{10})	4.40	4.6 – 5.9

3.2 Viral RNA amplicons

Viral RNA was extracted from all the 55 plasma samples with RNA concentration in ng/µL ranging between 35-120 ng/µL and purity levels between 1.70 and 2.0. Above 10 ng/µL and purity from 1.8 - 2.0 was required for successful downstream procedures, which was obtained from all of the 55 samples.

3.3 Detection of PCR amplicons (PR, RT and IN)

PCR amplification of the 3 fragments of the *pol* gene (PR, RT and IN) were done on all 55 viral RNA extracted samples. The samples were amplified in 2 runs (first and second) and positive bands were visualized on 1.0% TAE agarose gel picture as shown in examples in Figure 3.1, 3.2 and 3.3 for PR, RT and IN; respectively.

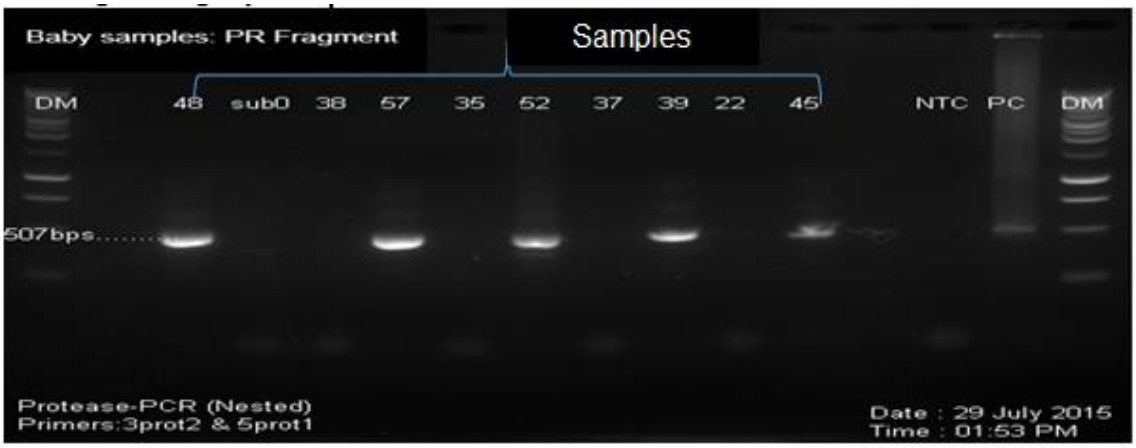


Figure 3.1 Partial *pol* **(**PR) fragment of 507 bp length in 1.0% TAE agarose gel picture. An example of positively amplified samples that include number, 45, 39, 52 and 57. Samples number, SubO, 38, 35, 37 and 22 are examples that could not be amplified with available PR primers. One kilo-base (1kb) DNA ladder at both ends of the gel picture - showing different bands of corresponding DNA concentration in nanograms (ng), was used as DNA marker (DM). MJ4 plasmid positive control (PC) and Non-template control (NTC) were included as controls. The gel was stained with GRGreen Nucleic Acid Gel Stain (Excellgen Ltd., South Africa).

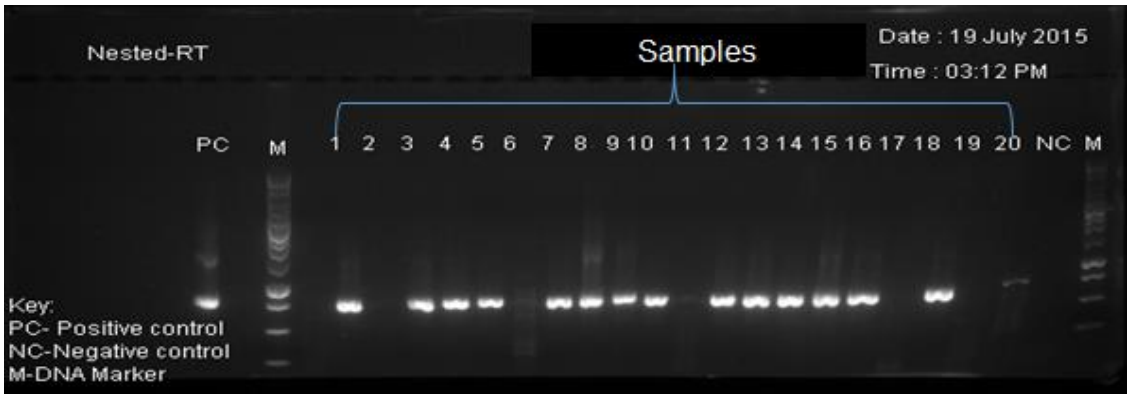


Figure 3.2 Partial *pol* **(**RT) fragment of 804 bp length in 1.0% TAE agarose gel picture. An example of positively amplified samples that include numbers, 1, 3-5, 7-10, 12-16 and 18 is indicated, while examples of samples 2, 6, 11, 17 and 19 would not be amplified with available RT primers. One kilo-base (1kb) DNA ladder at both ends of the gel picture - showing different bands of corresponding DNA concentration in nanograms (ng), was used as DNA marker (DM). MJ4 plasmid positive control (PC) and Non-template control (NTC) were included as controls. The gel was stained with GRGreen Nucleic Acid Gel Stain.

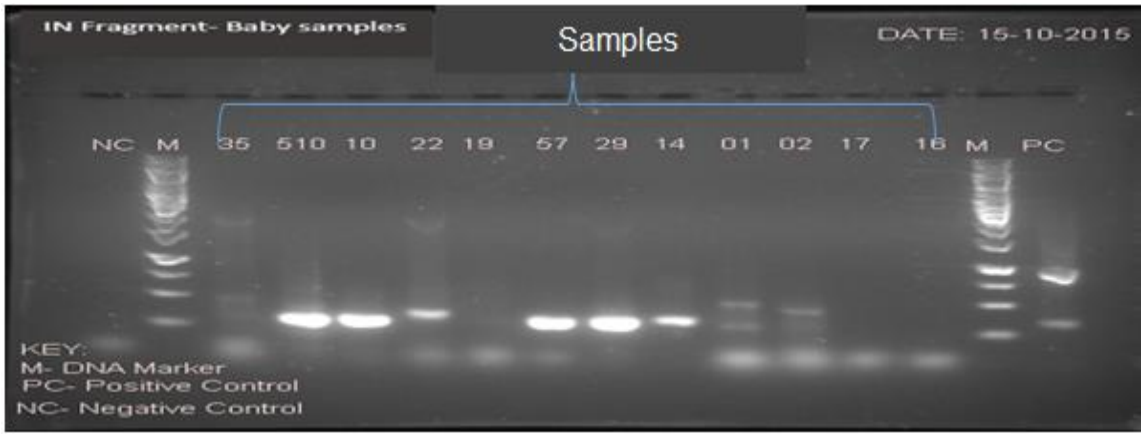


Figure 3.3 Partial *pol* IN fragment of 306 bp length in 1.0% TAE agarose gel picture. An example of positively amplified samples that include numbers, 510, 10, 22, 57,29,14,01 and 02, while an example of samples 35,19,17,16 would not be amplified with available IN primers. One kilo-base (1kb) DNA ladder at both ends of the gel picture - showing different bands of corresponding DNA concentration in nanograms (ng), was used as DNA marker (DM). MJ4 plasmid positive control (PC) and Non-template control (NTC) were included as controls. The gel was stained with GRGreen Nucleic Acid Gel Stain.

From the 55 samples, 29/55 (52.7%) PR amplicons were amplified, while 27/55 (49.1%) and 28/55 (50.9%) amplicons were amplified for RT and IN; respectively (Table 3.2 section A). A total of 37/55 (67.3%) amplicons were amplified for at least one of the *pol* gene fragments (PR or RT or IN) (Table 3.2 section A). Eighteen out of 55 samples (32.7%) amplifications were successful across all the 3 fragments; while, 23/55 (41.8%), 22/55 (40.0%) and 19/55 (34.0%) were amplified for PR/RT, PR/IN and RT/IN, respectively. A summary of all amplifications is provided in Table 3.2.

Table 3.2 Summary of positively amplified *pol* gene fragments

	Fragment	Number of samples out of 55	%age
A	PR	29	52.7%
	RT	27	49.1%
	IN	28	50.9%
	Either PR or RT or IN amplified	37	67.3%
	Two or 3 fragments amplified per sample		
B	Samples amplified in all of the 3 fragments (PR, RT and IN)	18	32.7%
	PR and RT	23	41.8%
	PR and IN	22	40%
	RT and IN	19	34%

3.4 Sequencing

All PCR amplicons for the three *pol* fragments (PR, RT and IN) were sequenced on both strands with forward and reverse primers mentioned under materials and methods in section 2.9 (Table 2.12). Typical sequence clean data chromatogram was

obtained as shown in example of sample CM-014 (Figure 3.4). The Sequencher parameters were customized to a 3' gap placement with a minimum overlap of 20 and a minimum match of 75.0%.

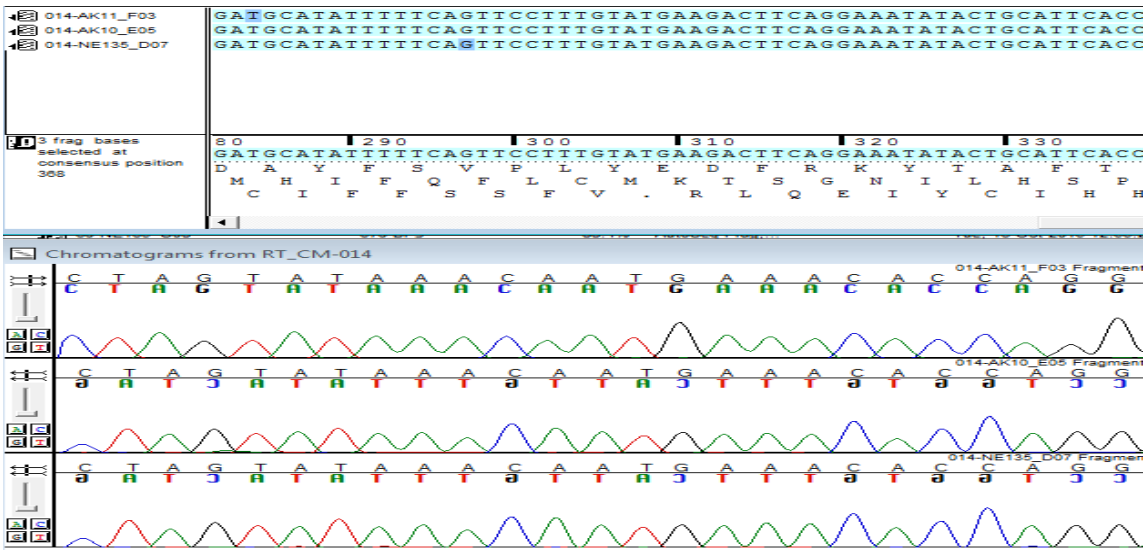


Figure 3.4 An example of a clean chromatogram from sample CM-014 with clearly defined peaks. The peaks in different colour sheds of green, red, blue and black represent adenine, thymidine, cytosine and guanosine nucleotides; respectively. The size of the nucleotide peak corresponds to its concentration.

3.5 Sequence alignment

After a quick screening of HIV-1 subtypes in all amplified *pol* gene products using online programmes, an assembly of overlapping sets of nucleotide sequences of different reads was run using MAFFT v7 (http://mafft.cbrc.jp/alignment/server/).This was only possible in RT fragment of 23 samples that included: CM-03, CM-05, CM-08, CM-09, CM-013, CM-014, CM-16, CM-017, CM-17, CM-19, CM-021, CM-21,CM-24, CM-29, CM-30,CM-31, CM-33, CM-39, CM-44, CM-47,CM-57, CM-177 and CM-510. These 23 samples had >400 nucleotide sequences in length enabling a quality and reliable tree to be drawn. The other two *pol* fragments (PR and IN) and 4 RT

samples had majority of sequences being <300 nucleotides due to lack of appropriate primers for better amplification. HIV reference sequences used to align sample sequences were retrieved from Los Alamos National Library. Similar nucleotides for both reference and sample sequence in the MAFFT window are identified in the same column, and any dissimilarity observed like “g” for “t” or “c” for “t” is assumed to be a mutation that can result in change of HIV subtype (Figure 3.5).

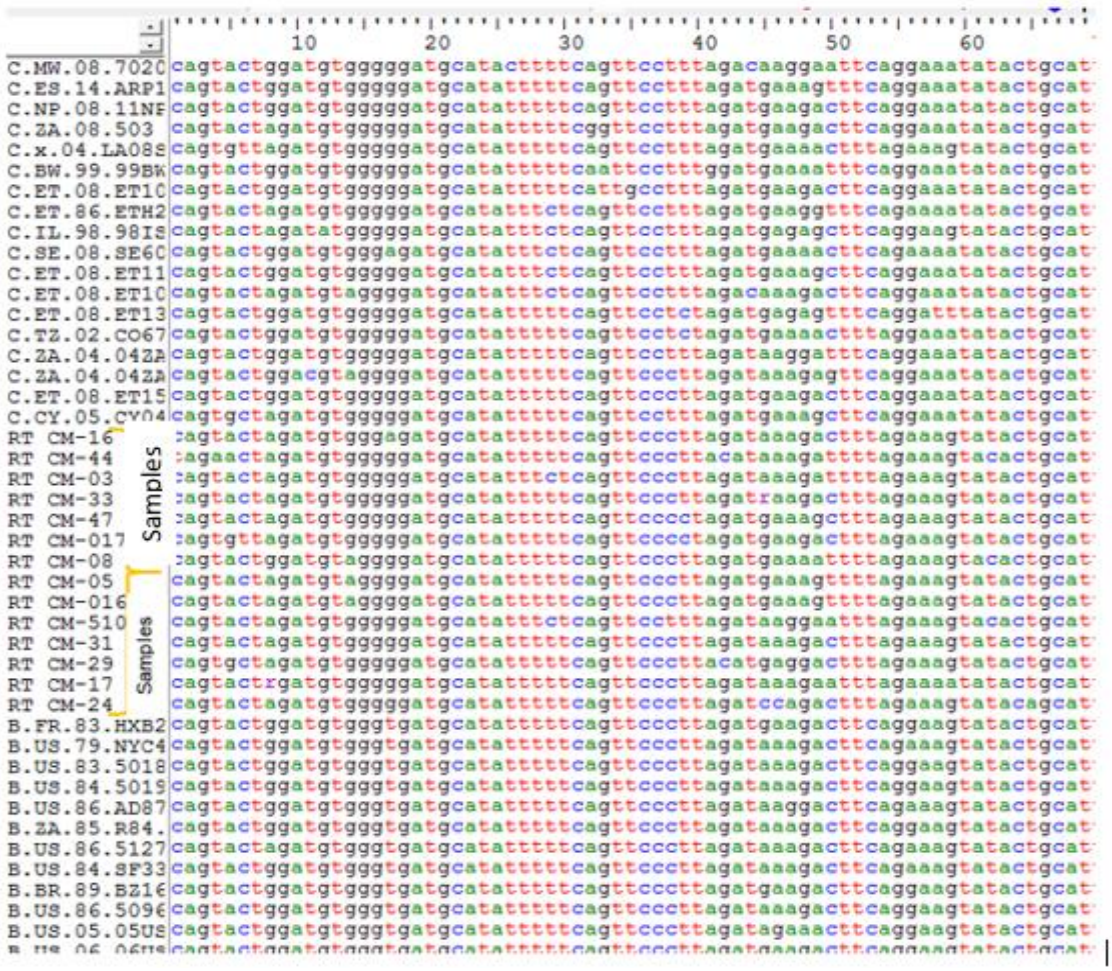


Figure 3.5 Nucleotide sequences of RT fragment from 14 samples (CM-16, CM-44, CM-03, CM-33, CM-47, CM-017, CM-08, CM-05, CM-016, CM-510, CM-31, CM-29, CM-17 and CM-24) aligned with HXB2 reference sequences using MAFT window. In the first column, a “t” nucleotide has replaced a “c” nucleotide, an indicator of a mutation.

3.6 Preliminary HIV-1 subtyping using various online programmes

HIV-1 subtyping analysis was interpreted using various online programs that included, jPHMM, RIP, REGA, SQUEAL, COMET and Stanford University HIV Drug Resistance Database. These online programs provided a quick HIV-1 subtype analyses of sample sequences before confirming through phylogenetic inference. An example below shows RT sequence of sample (CM-21) classified by REGA, jPHMM and RIP online programs (Figure 3.6, 3.7 and 3.8). The location of the sequence in HIV-1 gene map was based on the HXB2 reference sequences. The other subtyping tools – SCUEAL, COMET and Stanford only gives subtypes with no HIV-1 map location.

3.6.1 Jumping profile Hidden Markov Model (jpHMM)

The jpHMM subtyping program characterized 100.0% of the sequenced *pol* fragments (PR-29, RT- 27 and IN-28). The subtype results for all the three fragments for each sample sequence are shown in Table 3.4, 3.5 and 3.6 for PR, RT and IN respectively. An example of subtype inference for patient CM-21 based on RT fragment is subtyped as A1 using jpHMM is shown in Figure 3.6 below.

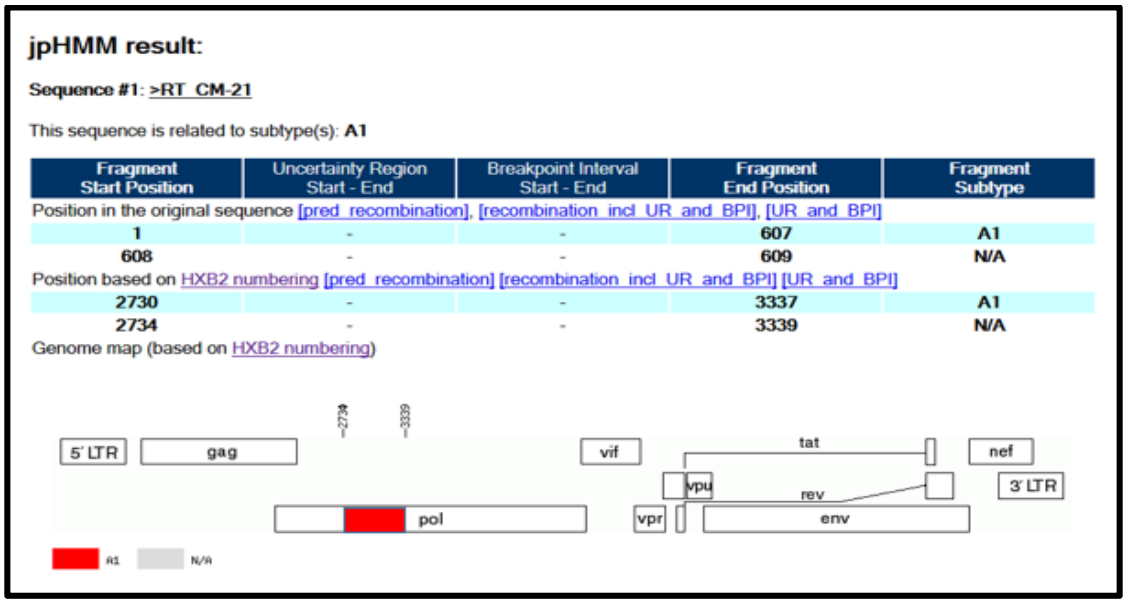


Figure 3.6 An example of sample sequence (CM-21) subtyped as HIV-1 subtype A1 (colored red in the *pol* genome) using jpHMM online subtyping program. The HIV-1 gene map was based on HXB2 reference.

3.6.2 Recombinant Identification Program (RIP)

The RIP subtyping tool characterized 20/29 (69.0%) for the PR and 26/27 (96.3%) for RT fragments but would not infer any subtype in all 28 sample sequences of the IN fragment. There were no recombinants detected. The RIP characterized sample CM-013 as subtype G and D for RT and IN, respectively while on the PR fragment it was unknown. An example of subtype inference for patient CM-21 based on RT fragment is subtyped as A1 using this online program and is shown in Figure 3.7 below.

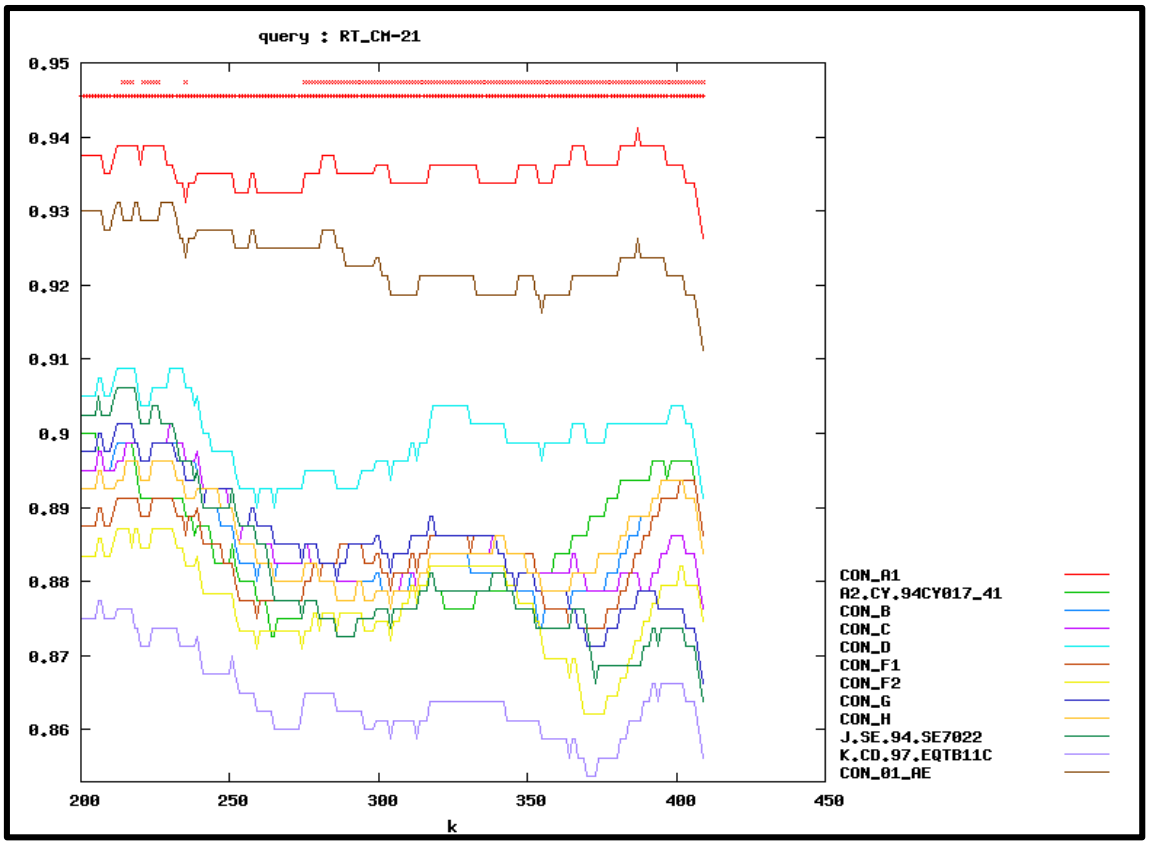


Figure 3.7 Sample CM_21 screened rapidly for inter-subtype recombination using RIP online program. The background contains consensus and subtype reference sequences of HIV-1 subtypes with subtype A1 (in red line) being the closest to the sample sequence.

3.6.3 REGA version 3.0

The REGA version 3.0 subtyping Programme characterized 8/29(27.6%), 13/27(48.0%) and 20/28(71.0%) of sequenced PR, RT and IN samples, respectively. In sample CM-013, it only characterized the IN fragment as subtype A1, while the PR and RT were unknown. An example of patient CM-21 based on RT fragment was inferred as subtype A1 as shown in Figure 3.8 below.

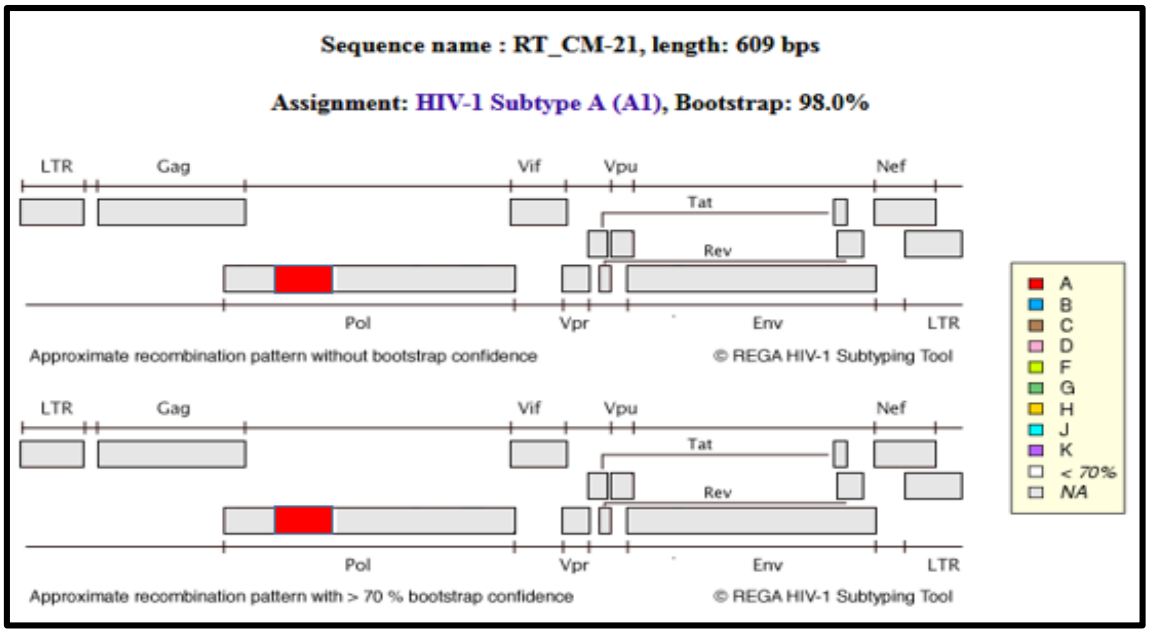


Figure 3.8 An example of sample sequence (CM-21) of 609 bps in length subtyped as HIV-1 subtype A1 (colored red in the *pol* genome) using REGA online subtyping program. The bootstrap is 98.0% which confirms the sample sequence as closely related to A1 subtype. The HIV-1 gene map was based on HXB2 reference.

3.6.4 SCUEAL (DataMonkey Program)

The DataMonkey subtyping program characterized all 27 and 28 sample sequences for RT and IN, respectively; while 26/29 (89.7%) were characterized for PR fragment with two remaining unknown. In sample CM-013, it was characterized as subtype CRF37, G and CRF19 for PR, RT and IN, respectively.

3.6.5 Context-based Modeling for Expeditious Typing Program (COMET)

This program characterized all 29, 27 and 28 sample sequences for PR, RT and IN respectively. In sample CM-013, it was characterized as FI, D and A1 for PR, RT and IN, respectively.

3.6.6 Stanford University HIV Drug Resistance Database

All sample sequences were characterized. Sample CM-013 was characterized as CRF02_AG, G and D for PR, RT and IN, respectively.

Table 3.3 Summary of HIV-1 subtype analyses based on *pol* (PR) fragment using six different online programmes

Sample	HIV subtype analyses					
	jPHMM	**COMET**	**SCUEAL**	**REGA**	**RIP**	**Stanford**
CM-02	G	G	CRF37	Unknown	Unknown	G
CM-03	CRF01_AE /F1	CRF02_AG	CRF02_AG	Unknown	Unknown	CRF02_AG
CM-05	G	CRF02_AG	CPX	G	Unknown	CRF02_AG
CM-08	CRF01_AE	CRF02_AG	CPX	Unknown	Unknown	CRF02_AG
CM-09	F2	F2	F2	Unknown	F2	F
CM-011	G	G	G	G	G	CRF02_2
CM-013	**A1**	**FI**	**CRF37**	Unknown	**G**	**CRF02_AG**
CM-014	C	C	C	C	C	C
CM-016	A1G	CRF02_AG	CPX	Unknown	G	CRF02_AG
CM-16	G	CRF02_AG	G	Unknown	G	CRF02_AG
CM-017	A1	CRF02_AG	G	Unknown	A1	CRF02_AG
CM-17	CRF01_AE	CRF02_AG	CRF37	Unknown	CRF01_AE	CRF02_AG
CM-19	C	C	C	C	C	C
CM-20	G	C	CRF01_AE / G	Unknown	Unknown	CRF02_AG
CM-21	C	C	C	C	C	C
CM-24	G	CRF02_AG	CRF02	Unknown	G	CRF02_AG

Sample	jPHMM	COMET	SCUEAL	REGA	RIP	Stanford
CM-30	A1	A1	CRF22	A(A1)	Unknown	CRF01_AE
CM-31	A1	CRF02_AG	G	Unknown	G	CRF02_AG
CM-32	A1	A1	AE	Unknown	CRF01_AE	K
CM-33	A1	CRF02_AG	CRF37	Unknown	G	CRF02_AG
CM-39	CRF01_AE	01AE	AE	Unknown	CRF01_AE	K
CM-44	A1	CRF02_AG	A4 / G.	Unknown	G	CRF02_AG
CM-45	C	C	C	C	C	C
CM-47	01_AE / G	CRF02_AG	CPX	Unknown	Unknown	CRF02_AG
CM-48	F1	CRF02_AG	A2 / G	Unknown	Unknown	CRF02_AG
CM-52	G	CRF02_AG	G	Unknown	G	CRF02_AG
CM-57	A1	A1	CRF22	A(A1)	CRF01_AE	CRF01_AE
CM-177	CRF01_AE	A1	NONE	Unknown	Unknown	CRF01_AE

*CM-013 is an example of a sequence that had the PR and IN fragments subtyped as subtype A1 while in RT fragment it was subtyped as B/G by jPHMM subtyping program

*Unknown: the relevant subtyping program was not able to characterize the sequence

Table 3.4 Summary of HIV-1 subtype analyses based on *pol* (RT) fragment using six different online programmes

Sample	jPHMM	COMET	SCUEAL	REGA	RIP	Stanford
CM-02	G	G	G	G	G	G
CM-03	CRF01_AE	CRF02_AG	CRF25	Unknown	G	CRF02_AG
CM-05	G	G	CRF02_AG	G	G	CRF02_AG
CM-08	G	CRF02_AG	G	Unknown	G	CRF02_AG
CM-09	F2	F2	F2	Unknown	F2	F

Sample	jPHMM	COMET	SCUEAL	REGA	RIP	Stanford
CM-014	C	C	C	C	C	C
CM-016	G	G	CRF02_AG	Unknown	G	CRF02_AG
CM-16	G	CRF02_AG	G	Unknown	G	G
CM-017	G	G	CRF02_AG	Unknown	G	CRF02_AG
CM-17	CRF01_AE	CRF02_AG	CRF01_AE	Unknown	G	CRF02_AG
CM-019	**C**	**CRF45-CPX**	**G / K**	**Unknown**	**CPX**	**K**
CM-19	C	C	C	C	C	C
CM-021	G	G	G	G	G	G
CM-21	A1	A1	CRF22	A (A1)	A1	CRF01_AE
CM-24	G	G	CRF02_AG	Unknown	G	CRF02_AG
CM-29	A1 / G	CRF02_AG	A1 / G.	Unknown	G	CRF02_AG
CM-30	A1	A1	CRF22	A (A1)	A1	CRF01_AE
CM-31	G	G	CRF02_AG	Unknown	G	CRF02_AG
CM-32	A1	A1	CRF22	A(A1)	A1	CRF01_AE
CM-33	G	G	CRF02_AG	G	G	CRF02_AG
CM-39	A1	A1	CRF22	A (A1)	A1	CRF01_AE
CM-44	G	CRF02_AG	A4 / G.	Unknown	G	CRF02_AG
CM-47	G	G	G	G	G	G
CM-57	A1	A1	CRF01_AE	A (A1)	A1	CRF01_AE
CM-177	F2	F2	F2	Unknown	F2	F
CM-510	G	CRF02_AG	CRF02_AG	Unknown	G	CRF02_AG

***CM-019** is an example of sequence clustered in different subtypes by respective subtyping programs

Table 3.5 Summary of HIV-1 subtype analyses based on *pol* (IN) fragment using five different online programmes.

Sample	HIV subtype analyses				
	jPHMM	COMET	SCUEAL	REGA	Stanford
CM-02	G	G	G	G	CRF02_AG
CM-03	G	G	G	G	G
CM-05	G	G	CRF36	G	CRF02_AG
CM-09	F2	F2	F1	Unknown	F2
CM-013	**A1**	**A1**	**CRF19**	**A1**	**D**
CM-016	G	G	CRF36	G	CRF02_AG
CM-017	G	G	G	G	CRF02_AG
CM-17	G	G	CRF02	Unknown	CRF02_AG
CM-19	C	C	C	C	C
CM-020	B	B	B	B	B
CM-20	G	G	CRF02	Unknown	CRF02_AG
CM-021	G	G	G	G	CRF02_AG
CM-24	A1	A1	AE	Unknown	CRF06_cpx
CM-28	G	G	G	G	G
CM-29	G	G	G	G	CRF02_AG
CM-30	C	C	CRF22	Unknown	CRF06_cpx
CM-31	G	G	G	G	CRF02_AG
CM-32	A1	A1	CRF22	Unknown	CRF06_cpx
CM-33	G	G	G	G	CRF02_AG
CM-39	A1	A1	CRF22	Unknown	A
CM-43	G	G	Unknown	G	CRF02_AG
CM-44	G	G	G	G	CRF02_AG

Sample	jPHMM	COMET	SCUEAL	REGA	Stanford
CM-47	G	G	G	G	CRF02_AG
CM-48	G	G	G	G	CRF02_AG
CM-57	A1	A1	CRF22	Unknown	A
CM-177	F1	F1	F2	Unknown	F2
CM-220	A1	A1	CRF22	Unknown	CRF11_cpx
CM-SubO	**Unknown**	**O**	**O**	**O**	**O**

***CM-013 subtyped as A1 by jPHMM, COMET and REGA programs while SCUEAL and Stanford subtyped as CRF19 and D respectively.

*CM-SubO was the only sequence subtyped as HIV-1 group O based on the IN fragment.

3.7 Group O sample

One sample from the 28 analyzed IN fragments was identified as HIV-1 group O, as indicated by REGA and RIP subtyping tools (Figure 3.9 and Figure 3.10). Other subtyping programs that could identify this subtype include COMET, SCUEAL, REGA and Stanford (CM-SubO in Table 3.5). The other 2 subtyping programs (jpHMM and RIP) would not assign the sequence to any HIV-1 subtype as the nucleotide sequence was not long enough.

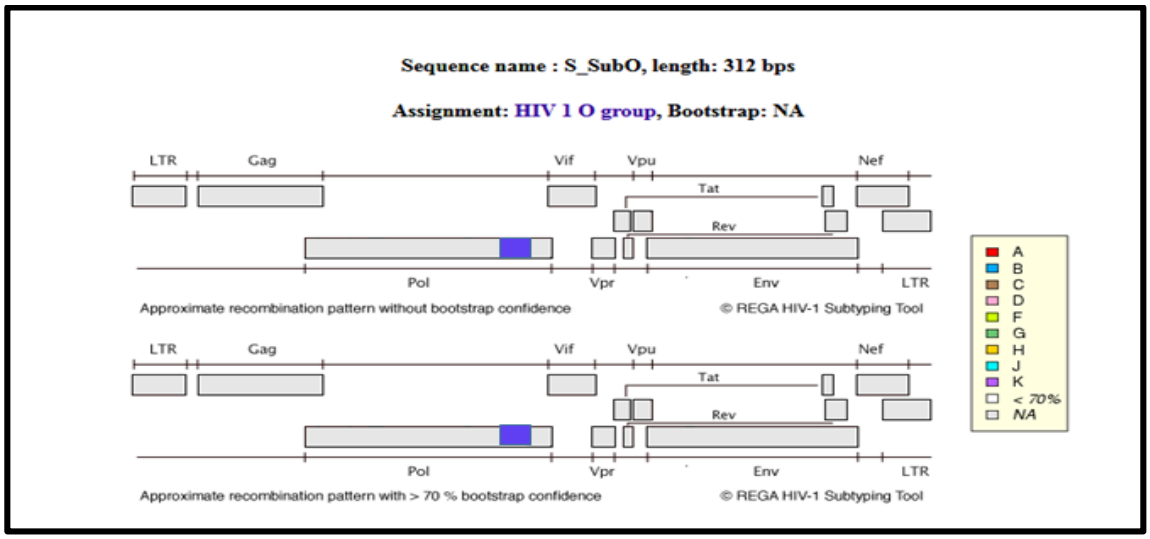


Figure 3.9 A sample sequence (Sample_SubO) assigned based on 312bp with a bootstrap of less than 70.0%. The sample sequence clustered to HIV-1 group O in the *pol* (IN) fragment (light blue in color) using REGA subtyping program.

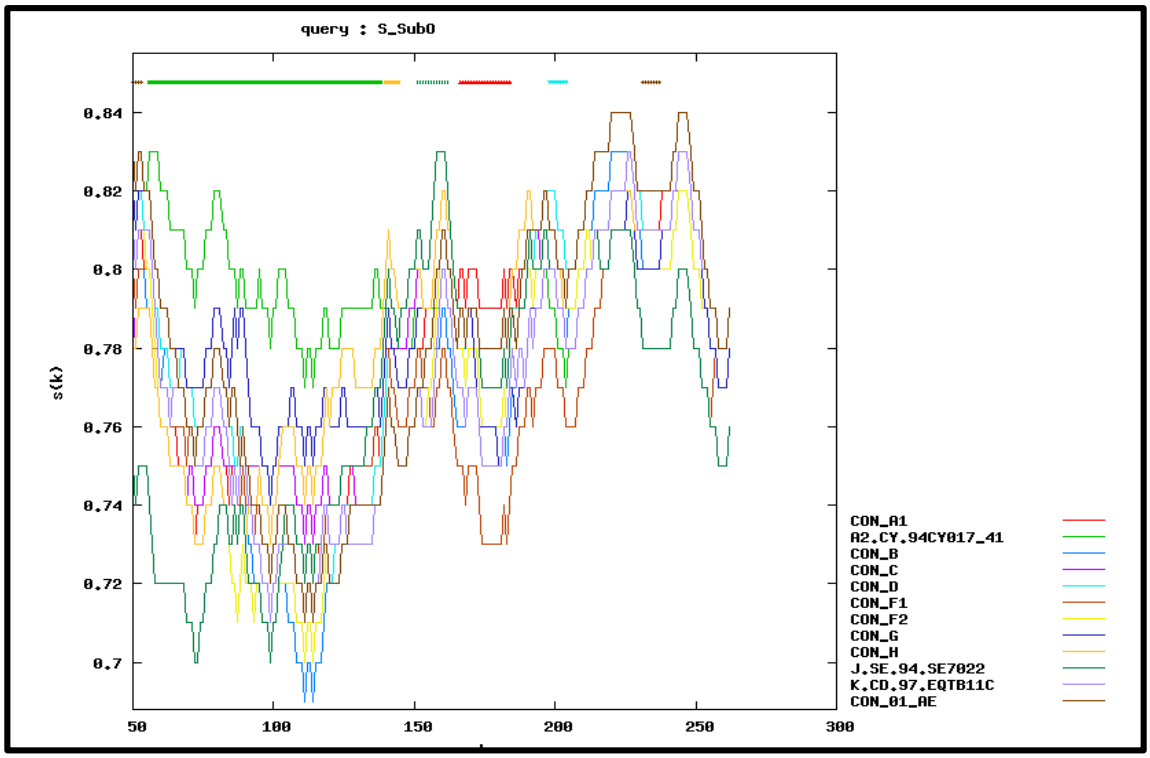


Figure 3.10 HIV-1 group O screened using RIP subtyping program. The background contains consensus and subtype reference sequences of HIV-1 subtypes.

3.8 Phylogenetic analyses

A phylogenetic evaluation for all 23 *pol* RT samples using neighbor-joining tree was reconstructed using MEGA version 5 with 1000 bootstrap replications (Figure 3.11). The majority of the sample sequences clustered with HIV-1 subtype G which accounted for 56.2% (n = 18), Subtype A 19.0% (n = 6), Subtype C and F2, 9.4% each (n = 3) and 2 subtype D (6.0%). The four subtypes observed are well illustrated in specific subtrees as indicated in Figure 3.12, 3.13, 3.14 and 3.15. Four of the sequenced samples could not be assigned to any subtype with confidence.

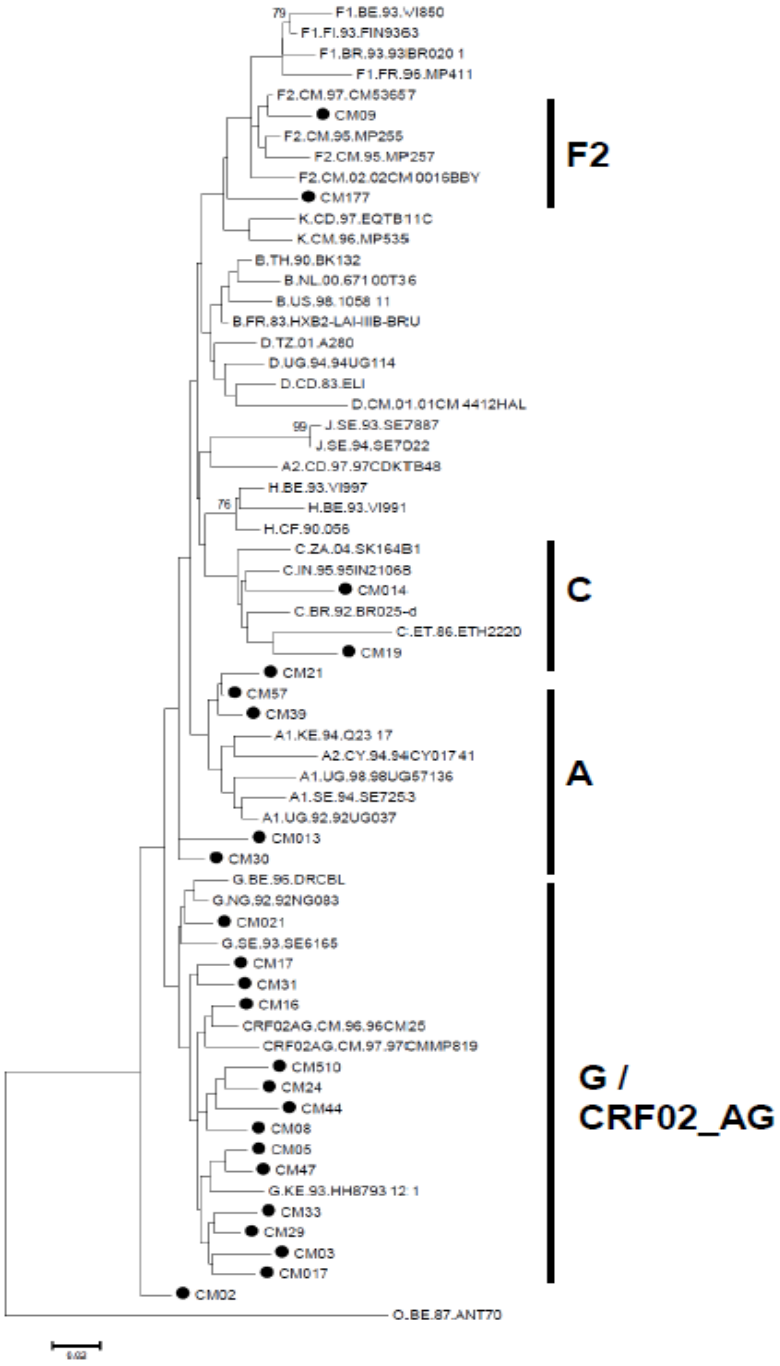


Figure 3.11 Phylogenetic relationship of the *pol* (RT) gene fragment of 450 bp in 23 samples (in bolder circles), clustered with HXB2 reference strains from los alamos database**.** The tree was drawn with Mega v 5 using Kimura 2 parameter and bootstrap performed in 1000 replicates. The tree was rooted to an outlier (group O). Four confirmed HIV-1 group M subtypes were characterized and include; G/CRF02_AG at 62.5% (n = 15), A (20.8%; n = 5), C (8.3%; n = 2) and F2 (8.3%; n = 2).

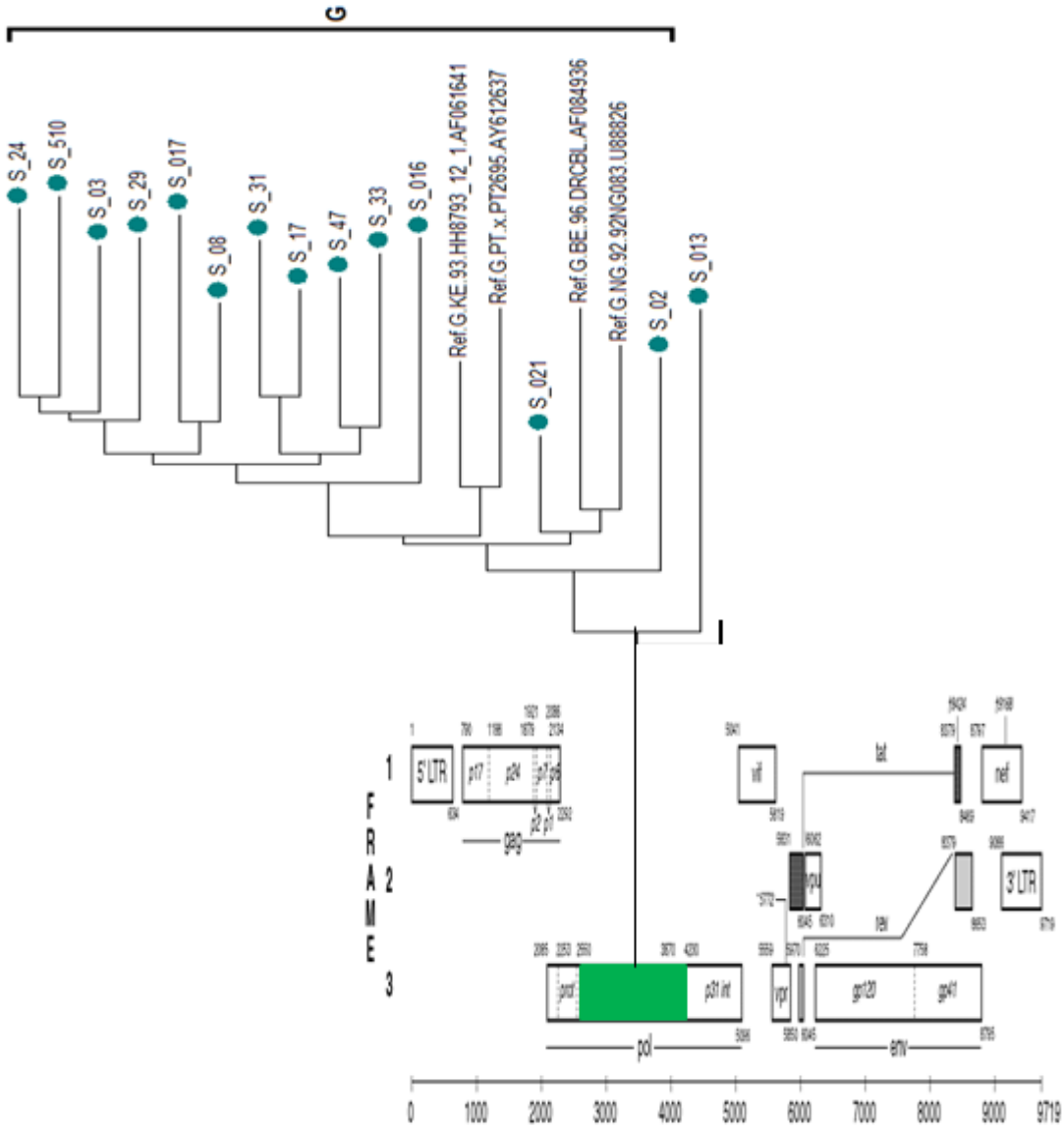


Figure 3.12 A phylogenetic subtree indicating HIV-1 group M subtype G samples (02,03,08,013,016,17,017,021,24,29,31,33,47 and 510) clustered with HXB2 references in *pol* (RT fragment).

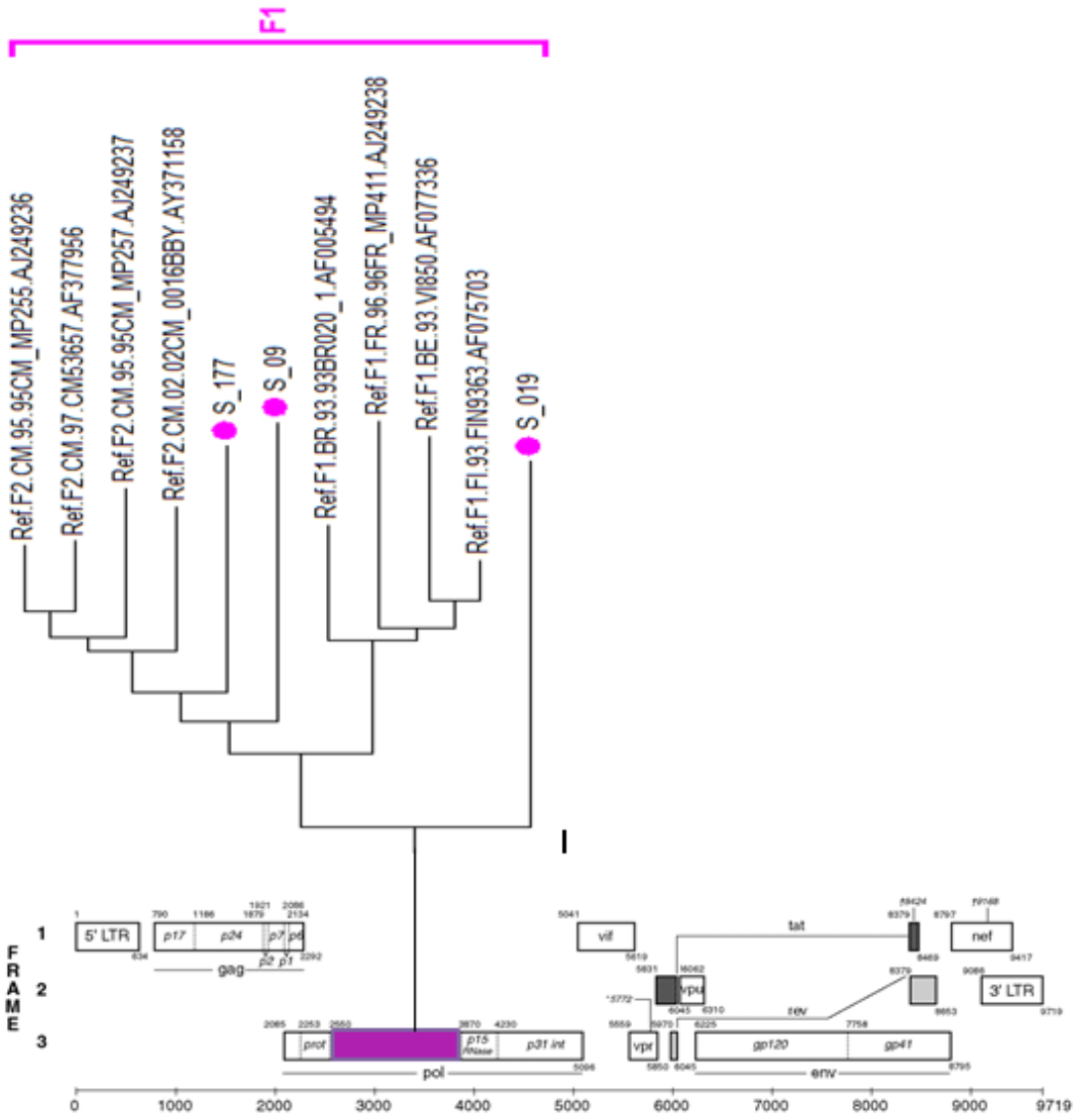


Figure 3.13 A phylogenetic subtree indicating HIV-1 group M subtype F1 samples (09, 019 and 177) clustered with HXB2 references in *pol* (RT fragment).

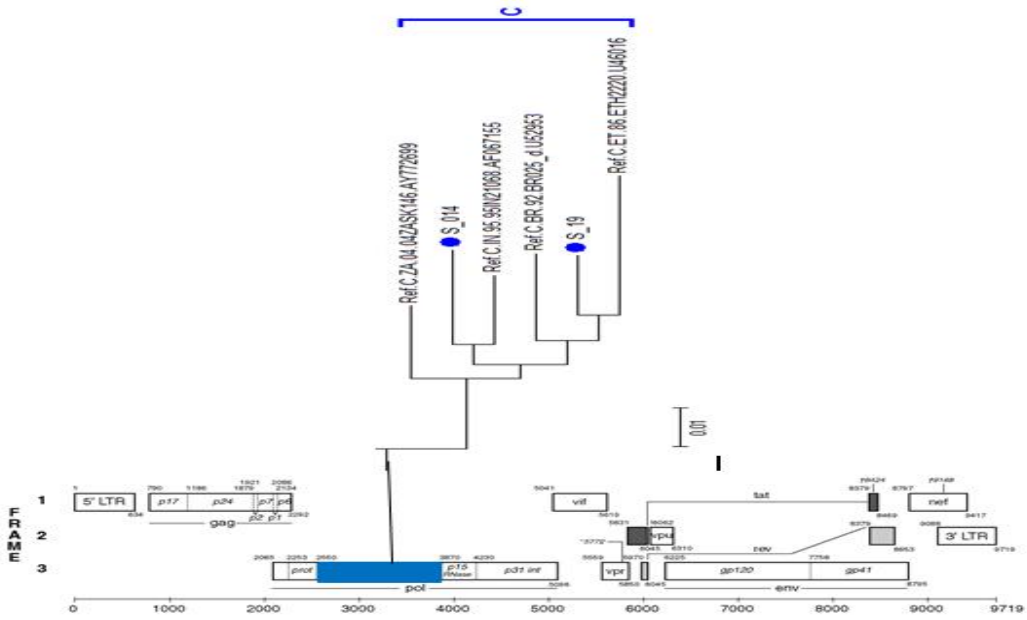


Figure 3.14 A phylogenetic subtree indicating HIV-1 group M subtype C samples (014 and 19) clustered with HXB2 references in *pol* (RT fragment).

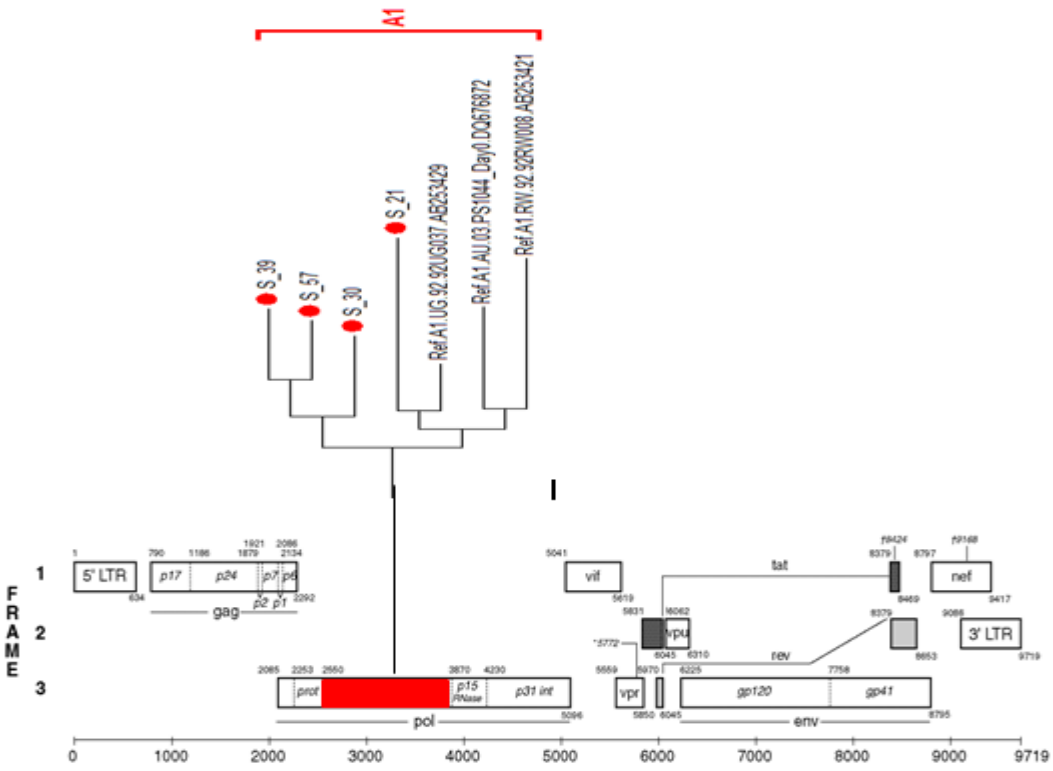


Figure 3.15 A phylogenetic subtree indicating HIV-1 group M subtype A1 samples (21, 30, 39 and 57) clustered with HXB2 references in *pol* (RT fragment).

3.9 HIV-1 drug resistance associated mutations in the *pol* (PR, RT and IN) genome

The 3 fragments of the *pol* gene (PR, RT and IN) were genotypically analysed for presence of various RAMs. The *pol* genome is a conserved region that is found with mutations that act against the commonly used cART (PIs, RTIs – NRTIs/NNRTIs and INIs). In the overall, 29, 27 and 28 sample sequences were screened for RAMs in the PR, RT and IN fragments respectively based on HIV Stanford Drug resistance data base (Table 3.6 and 3.7).

3.9.1 Protease RAMs

There was no major RAM observed in all 29 PR sequences analyzed while a total of 33 minor PIs RAMs were identified in 25 out of 29 (86.0%). These minor PIs were grouped into 8 types according to the Stanford HIV DR Database and they include; L10I, K20I, L10V, K20M, V11I, V11IV, G48GR and K20IK (Table 3.6). The L10I RAM is a polymorphic accessory mutation that increases the replication of viruses while K20I is a PI-selected accessory mutation that reduces NFV susceptibility. The V11I is a non-polymorphic PI-selected mutation associated with reduced DRV and FPV susceptibility when it occurs in combination with other PI-resistance mutations. The role of other minor PI mutations - L10V, K20M, V11IV, K201K and G48GR in drug resistance is under investigation. A summary of number of PR RAMs and their equivalent percentages is indicated in Figure 3.16.

Table 3.6 Minor PI RAMs as observed in the PR region of the HIV-1 genome using Stanford HIV Drug Resistance database

Sample ID	DRMs	
	Major resistance	Minor resistance
CM-02	NONE	L10I,K20I
CM-03	NONE	K20I
CM-05	NONE	K20I
CM-08	NONE	K20I
CM-09	NONE	L10V
CM-011	NONE	K20M
CM-013	NONE	V11I,K20I
CM-014	NONE	NONE
CM-016	NONE	K20I
CM-16	NONE	K20I
CM-017	NONE	V11IV,K20I
CM-17	NONE	K20I
CM-19	NONE	K20I
CM-20	NONE	K20I
CM-21	NONE	NONE
CM-24	NONE	K20I
CM-29	NONE	L10I,K20I
CM-30	NONE	NONE
CM-31	NONE	K20I
CM-32	NONE	L10I
CM-33	NONE	L10V,V11I,K20I
CM-39	NONE	L10I

Sample ID	Major resistance	Minor resistance
CM-44	NONE	K20I
CM-45	NONE	NONE
CM-47	NONE	K20I
CM-48	NONE	K20I
CM-52	NONE	K20I, G48GR
CM-57	NONE	K201K
CM-177	NONE	V11I,K20I
Total number of samples= 29	Total RAMs observed = 33	

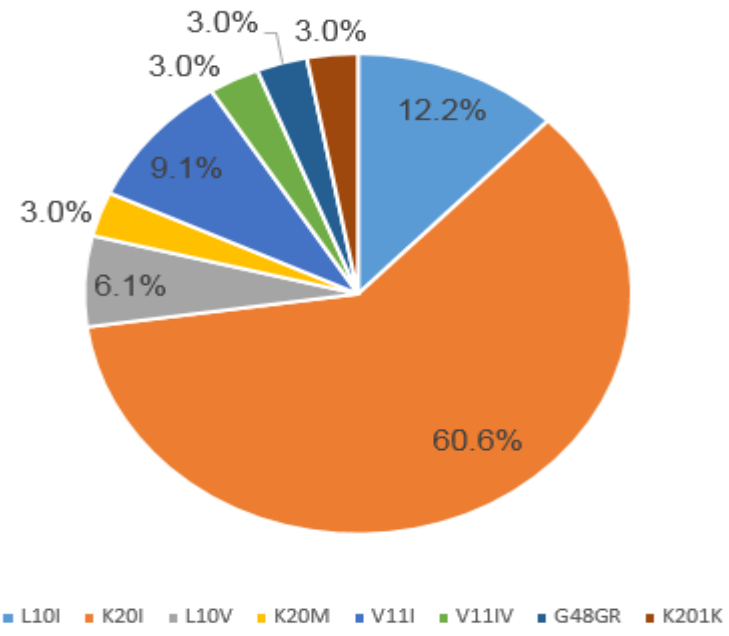


Figure 3.16 Summary of the percentages of minor PI RAMs observed in 25 out of 29 sequences analysed. Some samples had more than one RAM as observed in sample CM-02, CM-013 and CM-017 in Table 3.6 above.

3.9.2 NRTIs and NNRTIs RAMs

Both NRTI and NNRTI RAMS were observed in this study cohort with 27 RT sequences analyzed (Table 3.7). The NRTI mutations (only observed in one patient) included K65R, T215I and K219E; while the NNRTI resistance associated mutations were observed in 11 out of the 27 sequences (40.7%). The NNRTI mutations noticed were; V106M, Y181C and Y188H with each observed at 6.7% in the study population (Figure 3.17). The other NNRTI RAMs included the V90I at 20.0% and V179E at 13.3%. A number of other polymorphisms that were not necessarily associated to any drug resistance were also observed, and they include; E138A (13.3%), V179D (6.7%), V108I (6.7%), V106I (6.7%) L100IL (6.7%) and V90IV (6.7%) (Figure 3.17). A summary of drug resistance to both PIs, and NRTI/NNRTI RAMs in this study cohort is below (Figure 3.18).

Table 3.7 Resistance Associated Mutations against RT genomic region using Stanford HIV Drug Resistance database

Sample	Major DRMs	
	NRTI	**NNRTI**
CM-02	NONE	V179D
CM-03	NONE	NONE
CM-05	NONE	V90I,V179E
CM-08	NONE	NONE
CM-09	NONE	NONE
CM-013	NONE	NONE
CM-014	**K65R,T215I, K219E**	**V106M,Y181C, Y188H**

CM-016	NONE	V90I,V179E
CM-16	NONE	NONE
CM-017	NONE	V90I
CM-17	NONE	NONE
CM-019	NONE	NONE
CM-19	NONE	V108I
CM-021	NONE	V106I
CM-21	NONE	E138A
CM-24	NONE	NONE
CM-29	NONE	NONE
CM-30	NONE	NONE
CM-31	NONE	NONE
CM-32	NONE	NONE
CM-33	NONE	L100IL
CM-39	NONE	NONE
CM-44	NONE	V90IV
CM-47	NONE	NONE
CM-57	NONE	NONE
CM-177	NONE	E138A
CM-510	NONE	NONE

*CM-014 is the only patient with multiple drug resistance mutations against nearly all NRTIs and NNRTIs available approved by FDA.

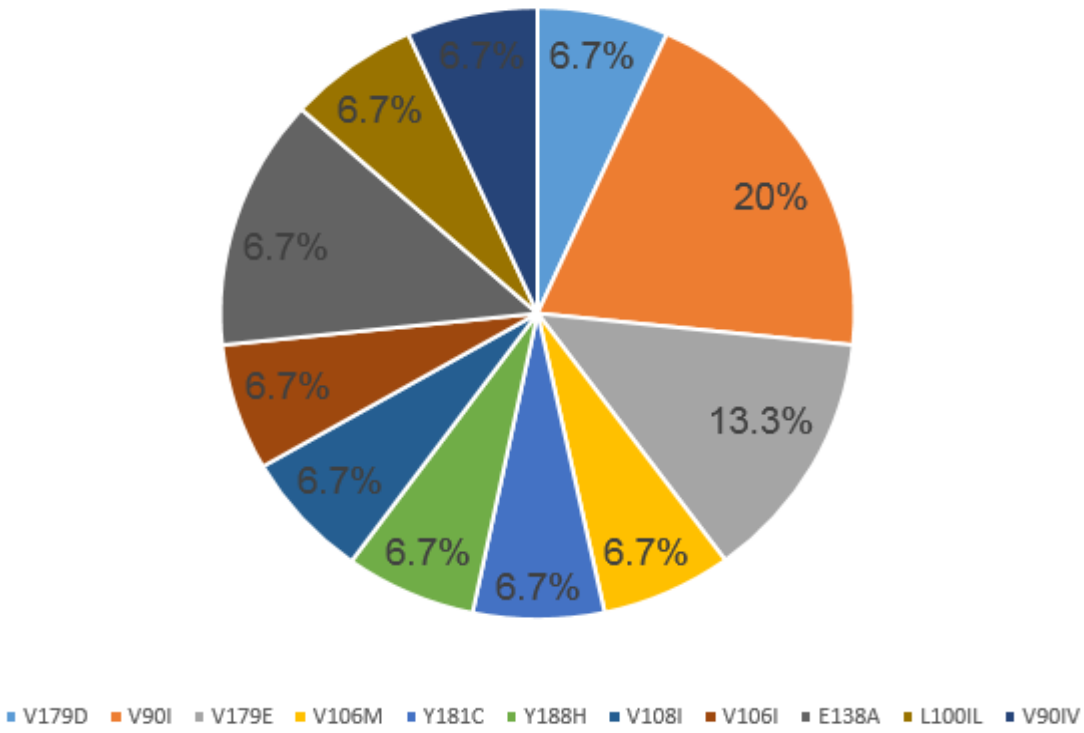


Figure 3.17 Summary of the percentages of NNRTI observed in 11 out of 27 RT sequences with E138A and V179E mutations with high frequency of occurrence. The V179E is a non-polymorphic mutation that is occasionally selected by NVP and EFV inhibitors while other mutations are not necessarily associated to any drug resistance.

Infant/Children: Level of drug resistance against NRTIs/NNRTIs

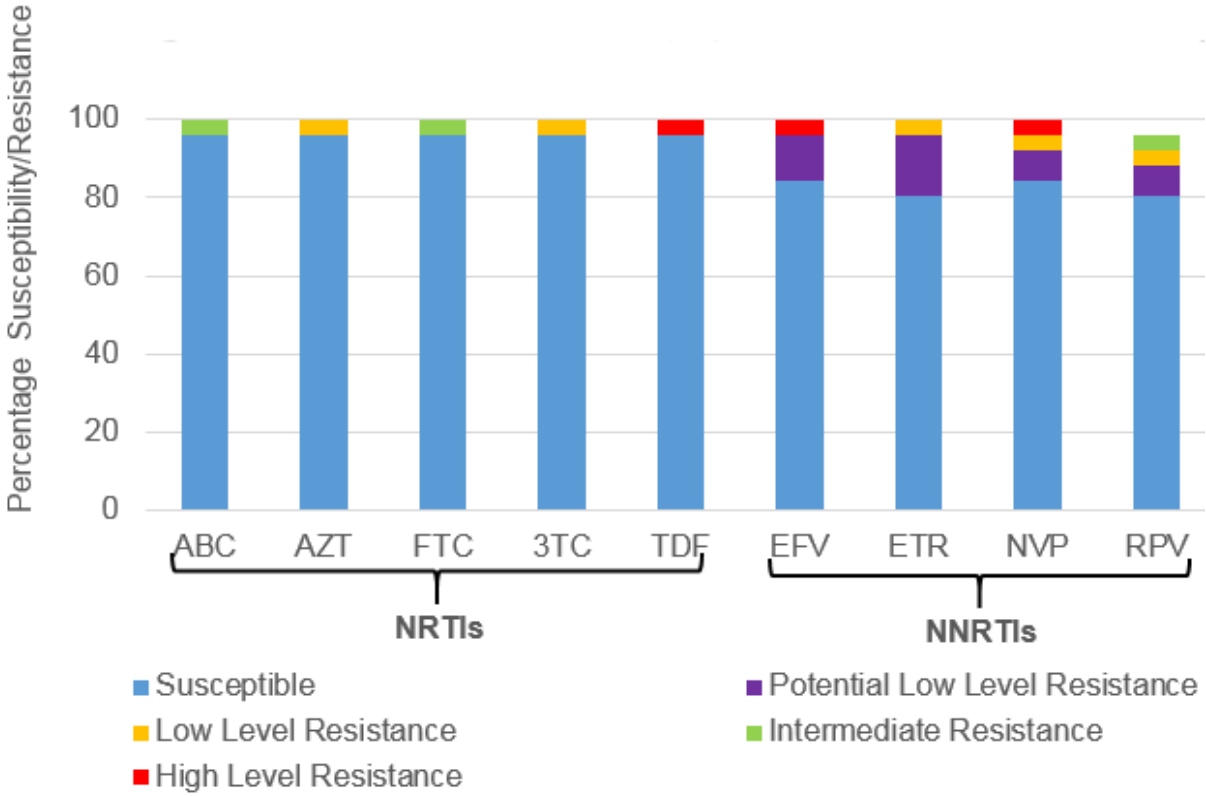


Figure 3.18 Observed drug resistance against RTIs (NRTI and NNRTIs). Majority of study participants (over 95.0%) were susceptible to all NRTIs while 3.7% had low level resistance to 3TC and intermediate to high level resistance to ABC, FTC and TDF. Approximately 55.0% were susceptible to NNRTIs while 40.7% of study participants had low level to potential low level resistance to EFV, ETR, NVP and RPV. Less than 3% had intermediate to high level resistance in RPV and NVP respectively.

3.9.3 Integrase RAMS

One sample (CM-020) was observed with a major resistance mutation E92EG which indicates intermediate resistance to Elvitegravir (EVG), low-level resistance to

Raltegravir (RAL) and susceptible to Dolutegravir (DTG). The other sample CM-33 showed accessory resistance mutation E157Q which is a polymorphic mutation with potential low-level resistance to Elvitegravir (EVG) but susceptible to Dolutegravir (DTG).

CHAPTER 4 Discussion

4.1 Overview of the study

This study focused on characterizing HIV diversity and drug resistance patterns from a cohort of therapy naïve infected infants from Yaoundé, Cameroon. The majority of infants (74.0%) were from the rural areas, while 26.0% were from the urban area. The urban areas (in this case - are patients from referral hospitals within Yaoundé city that include the mother and child center of the Chantal Biya Foundation (FCB) and the CSCCD research institute - both associated with Yaoundé 1 University Teaching Hospital. These two health institutions receive a high number of HIV positive patients due to better infrastructure and easy accessibility, especially - FCB of Cameroon -which receives support from the First lady, Mrs. Chantal Biya. Patients from rural areas visits district hospitals serving within those regions and include l'Hospital de district de la Cité Verte, l'Hospital de district d'Efoulan and l'Hospital de district de Biyemi-Assi. For the purpose of this study, the HIV-1 *pol* PR, RT and IN were examined to determine HIV diversity and drug RAMs that are currently circulating within this study cohort. The *pol* PR, RT and IN genome is a conserved region good for diversity studies and is also characterized with a variety of RAMs that inhibit the effectiveness of the commonly used cART regimens. Given that there are new HIV infections reported in this region, continuous HIV characterization and drug resistance patterns are very important in managing this pandemic.

4.2 HIV diversity

In this study cohort, the partial *pol* genomic regions (PR, RT and IN) were analysed for HIV-1 subtype inference. Majority (25/32, 75.7%) of sequenced samples were characterized by most of the online programs with concurring subtypes. Some of the sequences analysed were classified into the following subtypes and CRFs; CR02_AG, G, A, C, G, F1/F2 and CR02_AG. The major difference was observed in sample CM-013 that showed multiple subtypes in all the three fragments by each subtyping program.

4.3 Observed HIV subtypes using various online subtyping programs

The online subtyping programs – jpHMM (https://comet.lih.lu/), REGA (http://dbpartners.stanford.edu:8080/RegaSubtyping/stanford-hiv/typingtool**/**) SCUEAL (http://www.datamonkey.org/dataupload_scueal.php), RIP (http://www.hiv.lanl.gov/content/sequence/RIP/RIP.html). COMET (https://comet.lih.lu/)) and Stanford (https://hivdb.stanford.edu/hivdb/by-mutations/)-, were able to infer various subtypes in the majority of patient samples. In jpHMM, SCUEAL, COMET and Stanford subtyping programs all sequenced samples from the three *pol* (PR, RT and IN) fragments were inferred with a subtype. In RIP subtyping program, 69.0% and 96.5% of sequenced samples for PR and RT, respectively were inferred with a subtype, but in the *pol* (IN) fragment, none of all the sequenced samples were characterized. This was attributed to the short sequences (less than 300bp) obtained in the amplification of the IN fragment across all sample cohort. The REGA subtyping program was successful in 27.6%, 48.0% and 71% for PR, RT and

IN fragments, respectively, while SCUEAL was successful in all RT and IN fragments and 89.7% for PR fragment across all amplified and sequenced samples. There was uniformity in the majority of subtype inference, for example - all subtyping programs inferred patient CM-014 as subtype C. In a study done from Portuguese HIV Drug Resistance Database in comparing various online subtyping programs, pure subtype B and C had more than 95% sensitivity and specificity than other subtypes (Pineda-Pena AC.*et al*., 2013). The major discrepancy in subtype inference for all subtyping programs used for this study was observed in sample sequence CM-013, where the jpHMM subtyping tool inferred the PR and IN fragments as subtype A1, while the RT fragment was clustered as B/G. The RIP subtyping tool inferred similar sample sequence (CM-013) as subtype G and D for RT and IN, respectively while on the PR fragment it was unknown. The REGA subtyping tool clustered the sample sequence as subtype A1 based on the IN fragment while the PR and RT were unknown. The SCUEAL inferred the same sample sequence as subtype CRF37, G and CRF19 for PR, RT and IN respectively while COMET inferred it as FI, D and A1 for PR, RT and IN respectively. Finally, the Stanford subtyping tool inferred the same sample sequence as CRF02_AG, G and D for PR, RT and IN respectively. Cases like this of sample sequence CM-013 makes it difficult to completely rely on the subtyping programs because of instances of inability to infer specific subtype to a particular sample sequence or ambiguity in subtype inference for the same sample sequence. Online subtyping programs fail to infer subtype either because the sample sequence is a recombinant and cannot be assigned to any specific subtype or the length of the sample size is small for the algorithm to cluster it. To solve this problem, the use of phylogenetic trees to confirm the subtype is very crucial and this can be done either

through distance based analysis like for example use of NJ phylogenetic tree or by character based analysis like for example the Maximum likelihood (ML) phylogenetic tree. In patient CM-013 described above, the NJ phylogenetic tree was used and it was confirmed to be subtype A.

4.4 Neighbor – Joining phylogenetic inference

A NJ phylogenetic tree confirms subtypes observed using online programs. It gives information on similarities and differences of nucleotide sequences from their original ancestor, which are displayed in a phylogenetic tree. The confidence of the topology is based on the bootstrap value, which should be over 75.0% in sequences of the same clade. In this cohort, two HIV-1 groups were observed - group M and group O. In group M, four distinct subtypes were detected and they included; G/CRF02_AG, A, C and F2. The observed mosaic of HIV-1 subtypes in this study cohort, corroborates with other previous findings confirming the wide presence of HIV diversity in this region (Ndembi *et al*., 2008; Vallari *et al*., 2011; Hemelaar *et al*., 2011). The most predominant strain was G/CRF02_AG (n = 15; 62.5%), remaining the most circulating recombinant form in this study cohort. This high prevalence of G/CRF02_AG corresponds with other studies in Cameroon (Ndembi *et al*., 2008; Carr *et al*., 2010; Fokam *et al*., 2011; Ragupathy *et al*., 2011; Vallari *et al*., 2011). The CRF02_AG is a complex mosaic of HIV with alternating subtype A and G sequences (Agyingi *et al*., 2014). Other subtypes were less predominant, with A (n = 5, 20.8%), C (n = 2, 8.3%), and F2 (n = 2, 8.3%). Of interest was the presence of HIV-1 group O and HIV-1 group M subtype C. Group O in patient SubO was characterized based on an IN sequence, while subtype C in patient CM-014 was observed in all *pol* (PR, RT

and IN) fragments. HIV-1 group M subtype C is uncommon in this region and comparing the small sample size that was available for this study and the partial amplification of the *pol* fragment, it could be high. This raises concern over other transmissions and the geographical distribution of HIV, which could be due to the cosmopolitan nature of the city of Yaoundé and the large number of migrants from neighboring cities within Central Africa.

4.4.1 Group O sequence

HIV-1 group O represents 1.0% of all global infections (Vessiere *et al*., 2010; Yamaguchi *et al*., 2004). It is commonly confined to west and central Africa and more common in Cameroon (Leoz *et al.,* 2015). It could be possible that the presence of group O is underestimated due to inadequate screening procedures. In molecular characterization, HIV group O can easily remain undetected due to the complexity of designing the right molecular assays and primers for RNA amplification and sequencing. In the single case that was identified as group O in this study, it was clustered using only two subtyping program tools - REGA and RIP based on the *po*l (IN) fragment. Other subtyping tools like COMET, SCUEAL, REGA, jpHMM, RIP and Stanford would not cluster the sequence to any HIV subtype possibly because the nucleotide sequence was not long enough (only 300bp was available). HIV near full-length genome (NFL) analysis is necessary to confirm whether it is a pure HIV-1 group O. The other *pol* fragments (PR and RT) of the same sample were also unsuccessfully amplified due to primer failure and therefore missing more information on sample sequence characterization. There were challenges with PCR amplification of the same sample due to lack of appropriate primers or even designing them. The

small amount of sample was a huge limitation in performing further molecular analysis – like checking whether it is a recombinant. Group O is the most diverse of HIV-1 subtypes after group M. The routine serological methods for testing group O subtype are not effective as majority indicate negative results. Molecular characterization is more reliable in diagnosing this subtype and this is important in controlling the spread of this subtype which recent studies show that it is becoming a major concern in HIV control (Villabona-Arenas *et al*., 2015).

4.5 HIV-1 immunologic and virological markers

Most resource-limited countries, such as Cameroon, rely on WHO guidelines of clinical symptoms and viral load to monitor patients failing first or second-line cART. This is exactly the opposite in developed countries that have adopted routine viral load screening as a gold standard in diagnosing immunologic of virologic failure. According to the 2013 WHO guidelines that have been adopted by Cameroon, it regards immunologic failure in children perfectly adhering to treatment as those with CD4 count less than 200 cells/mm^3 or CD4 percentage of less than 10.0% in children less than 5 years. In children more than 5 years, CD4 count less than 100 cells/mm^3 is regarded as immunological failure. In virological failure under the same 2013 guidelines, it refers to those children following adequate treatment but with viral loads above 1000 RNA copies/ml after at least 24 weeks under cART. In this study Cohort, most of the children had CD4 percentage and absolute values within the therapeutic normal range (> 30.0% and >500 copies/ml). The value confers their immune-competent status as outlined in the WHO and CDC treatment guide for HIV immunological staging [(WHO, 2015: Country report:

(http://www.who.int/countries/cmr/en/)]. On the other hand, the majority had high viral load approaching 1000 RNA copies/ml mark but cannot be termed as virological failure, since they are not on treatment according to the 2013 WHO guidelines. The risk of such high viral load may not be well understood, especially when they start treatment. The CD4 values and as also indicated in other studies, suggests that the use of CD4 percentage and / or absolute values should be re-evaluated when decisions are made to start cART in limited resource settings (Takow *et al*., 2013). There was no correlation between CD4 cell percentage counts and viral-loads in this study.

4.6 cART for infants/children in Cameroon

First line treatment for infants under 3 years in Cameroon includes two NRTIs: ABC or AZT plus 3TC with LPV/r as a PI which are combined in the first-line cART regimen with a NNRTI: EFV or NVP. The PI include LPV/r. In the alternative first-line therapy, only EFV is replaced with NVP. In the second-line therapy, treatment is based on NNRTI used in first-line therapy and the preferred regimen is ABC or TDF plus 3TC plus LPV/r. In children over 3 years to 10 years, the first line therapy includes ABC plus 3TC plus EFV and alternative being AZT or TDF plus 3TC or FTC plus EFV or NVP. In the second line of treatment, AZT or ABC plus 3TC plus LPV/r with a replacement of TDF plus 3TC plus LPV/r has been adopted. Drug resistance on the above currently used class of drugs in this study cohort from Cameroon is described in the section 4.5 below in relation to corresponding observed resistance associated mutations.

4.7 Resistance associated mutations and polymorphisms in Cameroon

RAMs observed in this study cohort include 25/29 (86.0%) patients for minor PIs, one patient (1/27, 3.7%) for NRTIs and 11/27 (40.7%) patients for NNRTIs while 2/28 (7.1%) patients had RAMs against INIs. Some of NRTI RAMs observed in this study cohort like K65R, T215I and K219E (which is an accessory Thymidine analogue mutation) and NNRTI RAMs like V179D, V106M and Y181C have also been observed in other studies done in Cameroon (Laurent *et al*., 2006; Fokam *et al.,* 2013). A Single NRTI mutation like K65R is responsible for high-level resistance to majority of drugs like TDF, ddI and ABC while K219E mutation is associated with reduced susceptibility to drugs like AZT and possibly d4T (Larder & Kemp, 1989; Kuritzkes *et al*., 2000). Due to high diversity as observed in this study, it may difficult to identify all mutations that are relevant as described in subsequent sub-sections below.

4.7.1 Protease mutations

There were no major RAMs observed in all 29 PR sequences analyzed, while a total of 33 RAMs against minor PIs were identified in 25 out of 29 (86.0%) analysed samples. These minor PIs were grouped into 8 types and they include; L10I, K20I, L10V, K20M, V11I, V11IV, G48GR and K20IK. The L10I RAM was observed in 4 / 33 PIs RAMs (12.2%), and is a polymorphic accessory mutation that increase the replication of viruses with other PI- RAMs, while K20I was observed in 20/33 PI RAMs (60.6%) and is a consensus amino acid in subtype G and CRF02_AG. In subtype C, the K20I is a PI-selected accessory mutation that reduces NFV

susceptibility. The V11I (n = 3; 9.1%) is a non-polymorphic PI-selected mutation associated with reduced DRV and FPV susceptibility when it occurs in combination with other PI-resistance mutations. Other minor mutations observed in the PR fragment included; L10V (n = 2; 6.1%), K20M (n = 1; 3.0%); V11IV (n = 1; 3.0%) G48R (n = 1; 3.5%) and K201K (n = 1; 3.0%) - and their role in drug resistance has not yet been fully studied.

4.7.2 Reverse transcriptase mutations

The Reverse transcriptase mutations are grouped into two - NRTI or NNRTI RAMs and both cause drug resistance. The NNRTI cause resistance to ABC, AZT, FTC, 3TC and TDF while NNRTI RAMs can cause resistance to EFV, ETR, NVP and RPV In the 27 RT sequences analyzed, one patient (CM-014) had multiple major NRT and NNRTI resistance associated mutations. The NRTI mutations observed for this patient were K65R, T215I and K219E. The K65R mutation indicates high-level resistance to TDF, DDI and D4T, but low to intermediate-level resistance to 3TC and FTC. The T215I mutation indicates intermediated to high-level resistance to AZT and D4T, with low-level resistance to ABC, DDI and TDF. The K219E mutation, with accessory TAMs, is associated with reduced susceptibility to AZT and possibly to D4T. These mutations indicate that this patient is highly resistant to all current classes of NRTI and NNRTI drugs. The NNRTI resistance associated mutations were observed in 11 out of the 27 sequences (40.7%). These included the mutations V106M, Y181C and Y188H with each observed at 6.7% in the study population. The V106M mutation indicates high-level resistance to NVP and EFV, Y181C indicates reduced susceptibility to NVP, ETR, RPV and EFV, while Y188H indicates reduced

susceptibility to NVP and EFV. The other NNRTI RAMs included the V90I at 20% and is a polymorphic accessory mutation with minimal detectable reduction in NNRTI susceptibility. The V179E was at 13.3% and is a non-polymorphic mutation that is occasionally selected by NVP and EFV inhibitors. Some of these mutations like Y188H and Y181C have previously described in other studies conducted in Cameroon (Laurent *et al*., 2006; Aghoken *et al*., 2013). A number of other polymorphisms that were not necessarily associated to any drug resistance were also observed, and they include; E138A (13.3%), V179D (6.7%), V108I (6.7%), V106I (6.7%) L100IL (6.7%) and V90IV (6.7%).

4.7.3 Integrase mutations

There was one major resistance associated mutations (E92EG) observed in patient CM-020. This mutation showed intermediate resistance to Elvitegravir (EVG) and low-level resistance to raltegravir (RAL), but was susceptible to Dolutegravir (DTG). Another patient CM-33 had the E157Q mutation, which is a polymorphic mutation with potential low-level resistance to Elvitegravir (EVG), but susceptible to Dolutegravir (DTG). This patient had NNRTI mutation – L100IL and PI mutations – L10V, V11I and K20I. The E157Q mutation has been observed in CRF01_AE subtypes, which are common in West Africa region and in a study done in Cambodia showing resistance to RAL and EVG drugs (Maiga *et al*., 2009; Nouhin *et al.,* 2011). Since there are no INIs currently prescribed in Cameroon and many other resource-limited settings as per the 2013 WHO guidelines, it is likely that these mutations may have a negative impact in future if the same drugs are used in managing these

patients. The INIs have shown to be useful for salvage therapy and apparently less prone to resistance (Geretti *et al*., 2012; Dow & Bartlett, 2014).

4.8 Impact of this study cohort in HIV management

Despite concerted efforts by the Cameroon government and other international agencies, such as UNAIDS, in stopping HIV transmissions from mother-to-child, a significant number of children are still born with HIV, as observed in this study. Furthermore, one child had high level of drug resistance to almost all classes of cART. This is a big challenge especially in resource-limited countries like Cameroon, where few alternative cART regimens are available. This data can therefore be used in strengthening HIV prevention programs, like PMTCT, especially in the rural areas where few facilities are available. High prevalence of HIV diversity in Cameroon as observed in this study and other previous studies is a challenge in providing standard laboratory operating procedures as most assays are based on subtype B. The continuous presence of these diverse subtypes is a call to policy makers and international community to channel more resources in research on subtype association with new HIV infections being reported in Cameroon and elsewhere. Drug resistance in Cameroon is associated with non-adherence to cART or due to transmitted drug resistance associations. HIV resistance is also directly related to the rapid scale up of cART (Chaix *et al*., 2007; Fokam *et al*., 2011; Kityo *et al*., 2016). Access to genotyping and resistance assays can help manage HIV patients failing cART by providing baseline knowledge on the RAMS detected in the patient.

CHAPTER 5 Materials and Methods – Cohort II

5.1 Introduction

Study participants in this cohort were therapy naïve adult males and females of ages between 15 to 50 years old from Cameroon. The settings were the same as for those in cohort I as briefly described below. The CD4 count and viral load estimation was not performed due to lack of sufficient assays to carry out the respective laboratory procedures. Furthermore, only PR and RT were amplified and sequenced. The IN which was part of the *pol* fragment for analysis, was not done due to time constraints and funding for the project.

5.2 Ethical statement

There was one Ethical clearance covering letter for both cohorts as indicated in section 2.2.

5.3 Study design

This was also a cross-sectional study in the same period and settings as Cohort I in section 2.3**.**

5.4 Study participants

These were 125 therapy naïve adult - 43 Males and 83 Females totaling to 125 study participants (n = 125), with their ages ranging between 15 to 50 years old.

5.5 Equipment, chemicals, reagents, commercial assays and enzymes

The list was the same as in Cohort I in section 2.5; Table 2.1, Table 2.2, Table 2.3, Table 2.4, Table 2.5 and Table 2.6.

5.6 Sample collection, preparation and storage

Ten ml of whole blood from all 125 participants was drawn, stored and shipped to the Division of Medical Virology at University of Stellenbosch for further analysis as in cohort I in section 2.6. Due to inadequate laboratory assay kits for CD4 and viral load assays, CD4 and viral loads were not done.

5.7 Viral RNA extraction

Viral RNA was extracted from 125 plasma samples using the same standard operating procedures as in Cohort I section 2.6.3.

5.8 Viral RNA concentration and estimation

Viral RNA concentration and estimation was done so as to obtain a high yield of pure RNA and to remove other protein impurities or contaminants that would affect downstream procedures. Viral estimation was performed as described in Cohort I section 2.6.5.

5.9 Location of genome of interest (*pol* genome)

To characterize the HIV subtypes and drug resistance testing for this study Cohort, the *pol* genome (PR and RT) fragments was the focus as indicated on the HXB2 gene map location in Cohort I section 2.7, Figure 2.4.

5.10 PCR amplification

Both PCR reactions for RNA amplification (Pre-nested and Nested PCR) were run as indicated in Cohort I section 2.8 - the master mix preparation and cycling conditions

were as Table 2.9 and Table 2.10 for pre-nested PCR, while Table 2.11 and Table 2.12 were for nested PCR. The specific primers used for both runs are indicated in Table 2.5 and Table 2.6 for PR and RT, respectively as in Cohort I.

5.11 Detection and purification of PCR amplicons (PR and RT)

PCR amplifications were performed in 2 pol gene (PR and RT) fragments and positive bands visualized as in Cohort I section 2.8.3. Purification of PR and RT amplicons from excess primers, enzymes, nucleotides, polymerases and salts were performed as in Cohort I section 3.1.11. The eluted purified DNA for all the two fragments pol (PR and RT) were kept at -20^{o}C for subsequent downstream procedures.

5.12 Sequencing

All amplified adult samples obtained from the *pol* gene (PR and RT) were subjected to DNA sequencing which is based on the selective chain-determination of Sangers sequencing as performed in Cohort I section 2.9. The sequencing primers and PCR cycling parameters used were as indicated in Cohort I Table 2.12 and Table 2.13, respectively.

5.13 Quality control

To exclude contamination of sample DNA sequences obtained for this cohort study, all DNA sequences were subjected for quality check using los alamos laboratory data base (http://www.hiv.lanl.gov) as described in Cohort I Section 2.10.

5.14 Alignment

Overlapping DNA fragments were assembled after manual bio-editing using Sequencer version 5.0 with default parameter settings of sequence quality above 75.0% as in Cohort I section 2.10. Multiple alignments were created using MAFFT version 7 (http://mafft.cbrc.jp/alignment/server/) with HXB2 sequences used as reference.

5.15 Preliminary HIV-1 subtyping using online programs

Six online subtyping tools for HIV, which included COMET, jpHMM, SCUEAL, RIP, REGA and Stanford were used for preliminary HIV-1 subtypes as performed in Cohort I section 2.11.

5.16 Phylogenetic inference

To evaluate the relationship between the study cohort sequences and the currently available reference sequences, a Neighbor-Joining (NJ) phylogenetic tree was drawn using MEGA v.5 as described in cohort I section 2.13.

5.17 Drug resistance analyses

DNA sequences from the partial *pol* gene (PR and RT) fragments from positively amplified patient samples were subjected to the Stanford HIV drug resistance online programs as described in Cohort I section 2.12. The PI, NRTI and NNRTI RAMs were analysed and the data tabulated as indicated in Chapter 6, section 6.5 below.

CHAPTER 6 Results - Cohort II

6.1 Positively amplified *pol* gene (PR and RT) fragments

Of the 125 samples collected for this study cohort, there were 73/125 (58.4%) amplified amplicons for at least one of the two pol gene fragments. Specifically, 59/125 (47.2%) were positively amplified for PR, while 55 (44.0%) were amplified for the RT fragment. A summary of amplified amplicons for each of the two pol genes (PR and RT) fragments for this cohort is shown in Table 6.1. Examples of positive amplicons on the gel photo are shown in Figure 6.1 and 6.2.

Table 6.1 Summary of positively amplified *pol* gene fragments (n = 125)

Fragment	Number of samples out of 125	%age
Atleast 1 fragment amplified	73	58.4%
PR	59	47.2%
RT	55	44.0%
PR and RT	54	43.2%

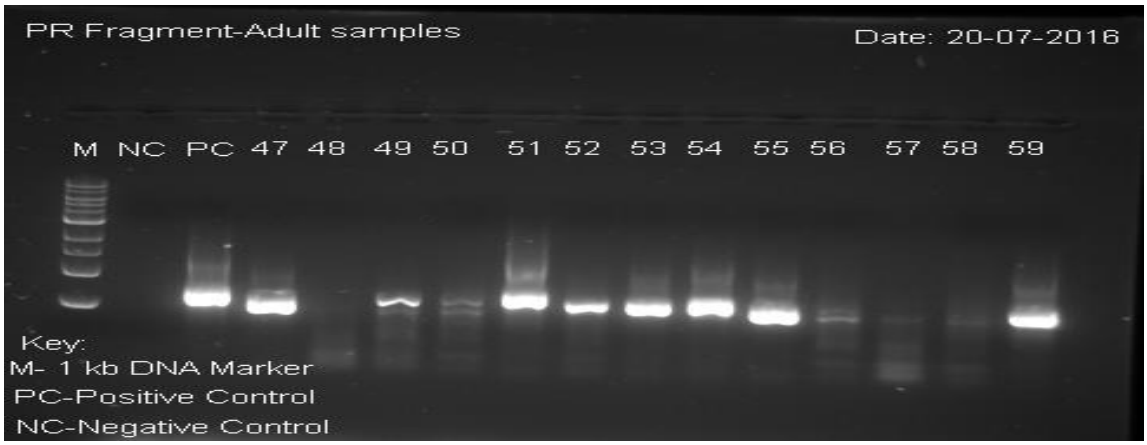


Figure 6.1 Partial *pol* **(**PR) fragment of 514 bp length in 1.0 % TAE agarose gel picture. An example of positively amplified samples that include numbers; 47, 49, 50,51,52,53,54,55,56 and 59 while samples number 48, 57 and 58 are examples that could not be amplified with available PR primers. One kilo-base (1kb) DNA ladder indicated as M showing different bands of corresponding DNA concentration in nanograms (ng), was used as DNA marker (DM). MJ4 plasmid positive control (PC) and Non-template control (NTC) were included as controls. The gel was stained with GRGreen Nucleic Acid Gel Stain.

Figure 6.2 Partial *pol* **(**RT) fragment of 804 bp length in 1.0 % TAE agarose gel picture. An example of positively amplified samples that include numbers; 41-43, 46,53,56,57 and 59 while an example of sample 58 would not be amplified with available RT primers. One kilo-base (1kb) DNA ladder showing different bands of corresponding DNA concentration in nanograms (ng) was used as DNA marker (M). MJ4 plasmid positive control (PC) and Non-template control (NTC) were included as controls. The gel was stained with GRGreen Nucleic Acid Gel Stain.

6.2 Sequencing and alignment

After DNA sequencing of cohort II sample sequences, multiple alignment was done using MAFFT version 7 and an example of aligned sequences with HXB2 references from los alamos data base sequences is given below in Figure 6.3

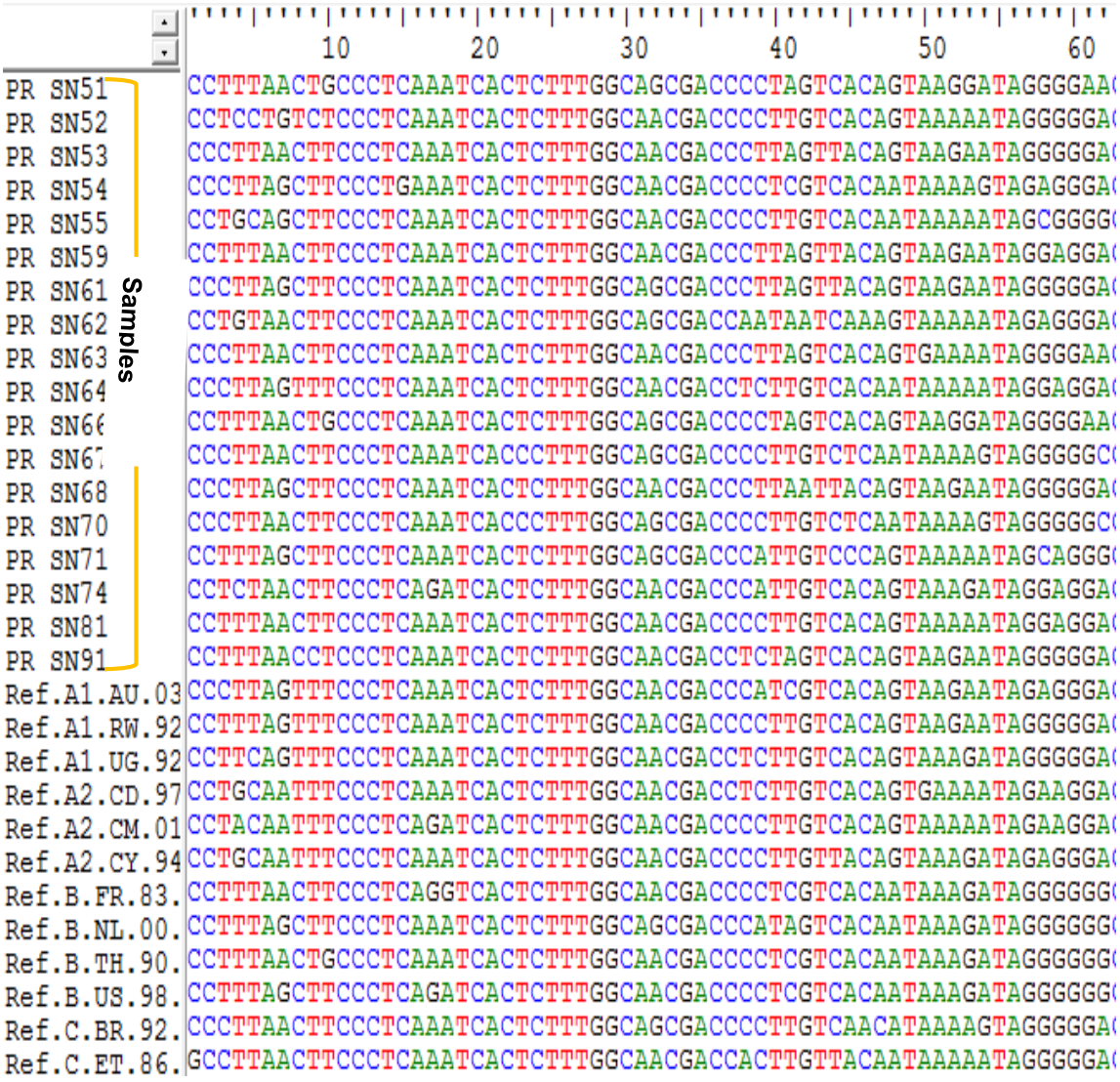


Figure 6.3 An example of Nucleotide sequences of PR fragment from 18 adult patients (PR SN51, SN52, SN53, SN54, SN55, SN59, SN61, SN62, SN63, SN64, SN66, SN67,SN68, SN70,SN71,SN81 and SN91). The sequences are aligned with HXB2 reference sequences using MAFT V7.0 online program.

6.3 Preliminary HIV-1 subtyping using various online programs

As indicated in the methods in section 5.13, the HIV-1 subtyping analysis was run using COMET (https://comet.lih.lu/), jpHMM (http://jphmm.gobics.de), SQUEAL (http://www.datamonkey.org/dataupload_scueal.php), REGA (http://dbpartners.stanford.edu:8080/RegaSubtyping/stanford-hiv/typingtool/-), RIP - (http://www.hiv.lanl.gov/content/sequence/RIP/RIP.html) and Stanford (https://hivdb.stanford.edu/hivdb/by-mutations/) online subtyping programs. An example of a preliminary HIV subtyping using REGA, jpHMM and RIP online subtyping tools is shown below in Figure 6.4, 6.5 and 6.6 respectively. The sample sequence is of the Pol (PR) fragment from patient SN52. Other online programs as indicated in cohort I section 3.17 do only show HIV subtype without HXB2 gene map location. A summary of HIV subtypes based on the six online tools for PR and RT fragments is tabulated in Table 6.2 and 6.3 respectively. Highlighted on Table 6.2 are PR sample sequences that include SN17, SN18, SN21, SN24, SN26, SN39 SN54, SN55, SN132 and SN148 with different subtypes as assigned by each subtyping program. Sample sequences for SN03, SN63 and SN224 on Table 6.3 are based on RT fragment and were assigned different subtypes using the online subtyping programs. Below is a brief description of observed subtypes in this study Cohort.

Table 6.2 PR Fragment for Cohort II

		HIV Subtypes Using six Online Subtyping Programs					
No.	**Sample ID**	**jPHMM**	**COMET**	**SCUEAL**	**REGA**	**RIP**	**Stanford**
1.	SN03	F2	F2	F2	Unknown	Unknown	F2
2.	SN16	G	CRF02_AG	G	Unknown	G	CRF02_AG
3.	SN9	G	G	CRF13	Unknown	G	CRF13_CPX
4.	SN10	A1	CRF02_AG	CRF09	Unknown	Unknown	A
5.	SN17	A1F1	C	CRF37	Unknown	G	G
6.	SN18	A1	F1	G	Unknown	Unknown	CRF02_AG
7.	SN21	01_AEA1G	F1	G	Unknown	Unknown	CRF02_AG
8.	SN24	G	B	Complex	Unknown	G	CRF02_AG
9.	SN26	CF1	A1	CRF02	Unknown	Unknown	CRF02_AG
10.	SN32	01_AEG	CRF02_AG	AE, G reco	Unknown	G	CRF02_AG
11.	SN34	G	A1	G	Unknown	Unkown	CRF02_AG
12.	SN35	AIG	A1	Complex	Unknown	G	CRF02_AG
13.	SN39	A1	F1	CRF37	Unknown	Unknown	CRF02_AG
14.	SN46	G	CRF02_AG	G	Unknown	G	CRF02_AG
15.	SN47	D	D	D	Unknown	D	D
16.	SN50	01_AE	CRF02_AG	CRF37	Unknown	01AE, G	CRF02_AG
17.	SN51	G	G	CRF13	Unknown	G	CRF13_cpx
18.	SN52	A1	CRF01_AE	CRF22	A1	Unknown	CRF01_AE
19.	SN53	G	CRF02_AG	G	Unknown	G	CRF02_AG
20.	SN54	DH	H,B	H	H	H	H
21.	SN55	A1G	D	CRF11	Unknown	Unknown	CRF11_cpx

No.	Sample ID	jPHMM	COMET	SCUEAL	REGA	RIP	Stanford
22.	SN59	01_AE	02_AG	Complex	Unknown	Unknown	CRF02_AG
23.	SN61	01_AEG	CRF02_AG	Complex	Unknown	G	CRF02_AG
24.	**SN62**	**A1F1G**	**F1**	**CRF37**	**Unknown**	**G**	**G**
25.	SN63	01-AE	CRF02_AG	G	Unknown	Unknown	CRF02_AG
26.	SN64	A1	F1	AE, F2 reco	Unknown	Unknown	CRF01_AE
27.	SN66	G	G	CRF13	Unknown	G	CRF13_cpx
28.	SN67	C	C	C	Unknown	C	C
29.	SN68	AEF1	CRF02_AG	Complex	C	Unknown	CRF02_AG
30.	SN70	C	C	C	C	C	C
31.	SN71	J	J	CRF11	J	Unknown	CRF11_cpx
32.	SN74	01_AE	CRF01_AE	CRF22	A1	Unknown	CRF09_cpx
33.	SN81	01_AE	A1	CRF22	A1	Unknown	CRF01_AE
34.	SN90	G	G	G	G	G	G
35.	SN102	CRFO2-AG	G	CRF37	Unknown	Unknown	G
36.	SN103	CRF01_AE /F1	CRF02_AG	CRF02_AG	Unknown	Unknown	CRF02_AG
37.	SN105	G	CRF02_AG	CPX	G	Unknown	CRF02_AG
38.	SN108	CRF01_AE	CRF02_AG	CPX	Unknown	Unknown	CRF02_AG
39.	SN109	F2	F2	F2	Unknown	F2	F
40.	SN111	G	G	G	G	G	CRF02_2
41.	SN116	A1G	CRF02_AG	CPX	Unknown	G	CRF02_AG
42.	SN117	A1	CRF02_AG	G	Unknown	A1	CRF02_AG
43.	SN119	C	C	C	C	C	C

No.	Sample ID	jPHMM	COMET	SCUEAL	REGA	RIP	Stanford
44.	SN120	G	C	CRF01_AE / G	Unknown	Unknown	CRF02_AG
45.	SN121	C	C	C	C	C	C
46.	SN124	G	CRF02_AG	CRF02	Unknown	G	CRF02_AG
47.	SN129	F1G	CRF02_AG	Unknown	Unknown	G	CRF02_AG
48.	SN130	A1	A1	CRF22	A(A1)	Unknown	CRF01_AE
49.	SN131	A1	CRF02_AG	G	Unknown	G	CRF02_AG
50.	SN132	A1	A1	AE	Unknown	CRF01_AE	K
51.	SN133	A1	CRF02_AG	CRF37	Unknown	G	CRF02_AG
52.	SN139	CRF01_AE	01AE	AE	Unknown	CRF01_AE	K
53.	SN144	A1	CRF02_AG	A4 / G.	Unknown	G	CRF02_AG
54.	SN145	C	C	C	C	C	C
55.	SN147	01_AE / G	CRF02_AG	CPX	Unknown	Unknown	CRF02_AG
56.	SN148	F1	CRF02_AG	A2 / G	Unknown	Unknown	CRF02_AG
57.	SN152	G	CRF02_AG	G	Unknown	G	CRF02_AG
58.	SN157	A1	A1	CRF22	A(A1)	CRF01_AE	CRF01_AE
59.	SN217	CRF01_AE	CRF02_AG	CRF37	Unknown	CRF01_AE	CRF02_AG

*****Highlighted are examples of different subtypes assigned to the same sequence based on a different subtyping program.

Table 6.3 RT Fragment for Cohort II

		HIV Subtypes Using six Online Subtyping Tools					
No.	**Sample ID**	**jpHMM**	**COMET**	**SCUEAL**	**REGA**	**RIP**	**Stanford**
1.	SN02	G	G	G	Unknown	G	CRF02_AG
2.	SN03	A1F2	A1F2	AE, F2	A1	F2	F2
3.	SN06	G	G	A4, G	Unknown	G	CRF02_AG
4.	SN9	G	G/02_AG	G	Unknown	G/02-AG	CRF02_AG
5.	SN12	A1	A1	CRF22	A(A1)	A1	CRF01_AE
6.	SN16	G	CRF02_AG	G	Unknown	G	G
7.	SN17	A1	A1	A3	A(A1)	01_AE, A1	A
8.	SN18	A1	CRF02_AG	Complex	Unknown	01AE, G	G
9.	SN20	G	02_AG	AG	G	G	CRF02_AG
10.	SN22	G	02_AG	G	Unknown	G	CRF02_AG
11.	SN26	G	02_AG	G	G	G	CRF02_AG
12.	SN29G	A1	A1	CRF22	A(A1)	A1	A
13.	SN30G	G	02_AG	A1,G	Unknown	G	CRF02_AG
14.	SN31	G	G	CRF36	Unknown	G	CRF02_AG
15.	SN32	G	G/02_AG	G	Unknown	G	CRF02_AG
16.	SN34	G	02_AG	CRF02	Unknown	G	CRF02_AG
17.	SN35	G	02_AG	G	Unknown	Unknown	CRF02_AG
18.	SN36	G	G	G	Unknown	G	CRF02_AG
19.	SN38G	C	C	C	C	C	C
20.	SN39	G	G/02_AG	CRF02	G	G	CRF02_AG
21.	SN41	C	C	C	C	C	C
22.	SN42	C	C	C	C	C	C

No.	Sample ID	jpHMM	COMET	SCUEAL	REGA	RIP	Stanford
23.	SN43	A1	A1,02_AG	Complex	Unknown	G	CRF02_AG
24.	SN51	B	13-cpx	CRF13	Unknown	13-cpx	CRF13_cpx
25.	SN52G	A1G	02_AG	AG	Unknown	G	CRF63_02 A1
26.	SN54	A1	02_AG	G	Unknown	G	CRF02_AG
27.	SN55	C	11_cpx	CRF11	Unknown	11cpx	CRF11_cpx
28.	SN56	A1	A1/26-AU	Complex	Unknown	G	A
29.	SN58G	A1G	02_AG	A,G	Unknown	G	CRF02_AG
30.	SN59	C	13-cpx	CRF13	Unknown	13-cpx/H	CRF13_cpx
31.	**SN63**	**B**	**13_cpx**	**CRF13**	**Unknown**	**H,B,G**	**CRF13_cpx**
32.	SN72G	C	C	C	C	C	C
33.	SN92G	A1	A1-D	CRF22	A(A1)	A1	A
34.	SN-103	CRF01_ AE / G	CRF02_AG	CRF25	Unknown	G	CRF02_AG
35.	SN-105	G	G	CRF02_A G	G	G	CRF02_AG
36.	SN-108	G	CRF02_AG	G	Unknown	G	CRF02_AG
37.	SN117	CRF01_ AE / G	CRF02_AG	CRF01_A E / G	Unknown	G	CRF02_AG
38.	SN119	C	C	C	C	C	C
39.	SN121	A1	A1	CRF22	A (A1)	A1	CRF01_AE
40.	SN124	G	G	CRF02_A G	Unknown	G	CRF02_AG
41.	SN129	A1 / G	CRF02_AG	A1 / G.	Unknown	G	CRF02_AG
42.	SN130	A1	A1	CRF22	A (A1)	A1	CRF01_AE
43.	SN131	G	G	CRF02_A	Unknown	G	CRF02_AG

No.	Sample ID	jpHMM	COMET	SCUEAL	REGA	RIP	Stanford
44.	SN132	A1	A1	CRF22	A(A1)	A1	CRF01_AE
45.	SN133	G	G	CRF02_AG	G	G	CRF02_AG
46.	SN139	A1	A1	CRF22	A (A1)	A1	CRF01_AE
47.	SN144	G	CRF02_AG	A4 / G.	Unknown	G	CRF02_AG
48.	SN147	G	G	G	G	G	G
49.	SN157	A1	A1	CRF01_AE	A (A1)	Unknown	CRF01_AE
50.	SN217	G	G	CRF02_AG	Unknown	G	CRF02_AG
51.	SN219	C	CRF45-CPX	G / K	Unknown	CPX	K
52.	SN221	G	G	G	G	G	G
53.	SN224	BG	02_AG	AG	Unknown	G	CRF02_AG
54.	SN232	G	G/02_AG	G	Unknown	G	CRF02_AG
55.	SN252	A1	A1	A1AE	A(A1)	Unknown	A

*Highlighted example of sequences with different subtypes as per each subtyping program.

6.3.1 REGA subtyping program

The REGA subtyping tool characterized 15/59(25.42%) and 23/55(41.81%) of sequenced PR and RT samples, respectively. One example of assigned subtype based on PR fragment of sample SN52 by this program is shown in Figure 6.4 below.

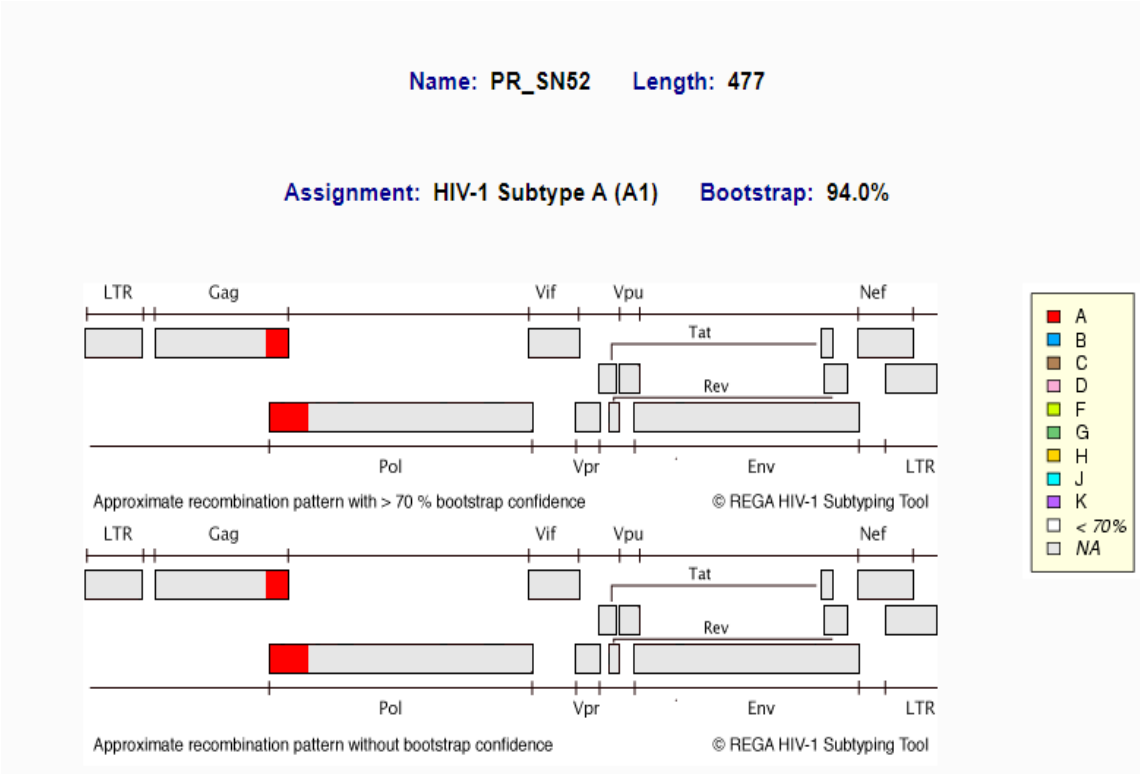


Figure 6.4 An example of sample sequence (PR_SN52) of 477 bp in length subtype assigned is pure HIV-1 subtype A (A1) colored red in the *pol* genome. The subtyping was performed using **REGA** online subtyping program. The bootstrap support is 94.0% which confirms the sample sequence as closely related to A1 subtype. The HIV-1 gene map was based on HXB2 reference.

6.3.2 jumping profile Hidden Markov Model (jpHMM)

The jpHMM subtyping tool characterized 100% of the sequenced pol fragments (PR-59 and RT- 55). An example of subtype inference for patient SN52 based on PR fragment is subtyped as A1 as shown in Figure 6.5 below.

jpHMM result:

Sequence #1: >PR_SN52

This sequence is related to subtype(s): **A1**

Fragment Start Position	Uncertainty Region Start - End	Breakpoint Interval Start - End	Fragment End Position	Fragment Subtype
Position in the original sequence [pred_recombination], [recombination_incl_UR_and_BPI], [UR_and_BPI]				
1	1 - 477	-	477	A1
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2150	2150 - 2638	-	2638	A1

Genome map (based on HXB2 numbering)

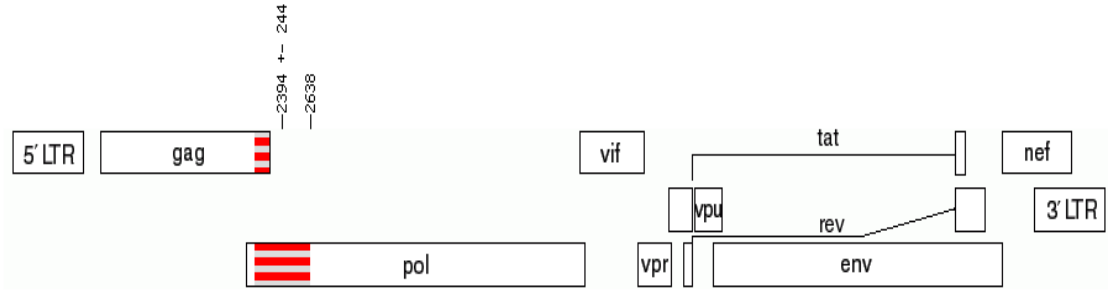


Figure 6.5 An example of sample sequence (PR_SN52) with assigned subtype HIV-1 subtype A1 (colored red in the *pol* genome) using jpHMM online subtyping program. The HIV-1 gene map was also based on HXB2 reference

6.3.3 Recombinant Identification Program (RIP)

The RIP subtyping tool characterized 37/59 (62.71%) for the PR and 26/27 (94.54%) for RT fragments. An example of subtype inference for patient SN52 based on PR fragment was assigned subtype A1 using and is shown in Figure 6.6 below.

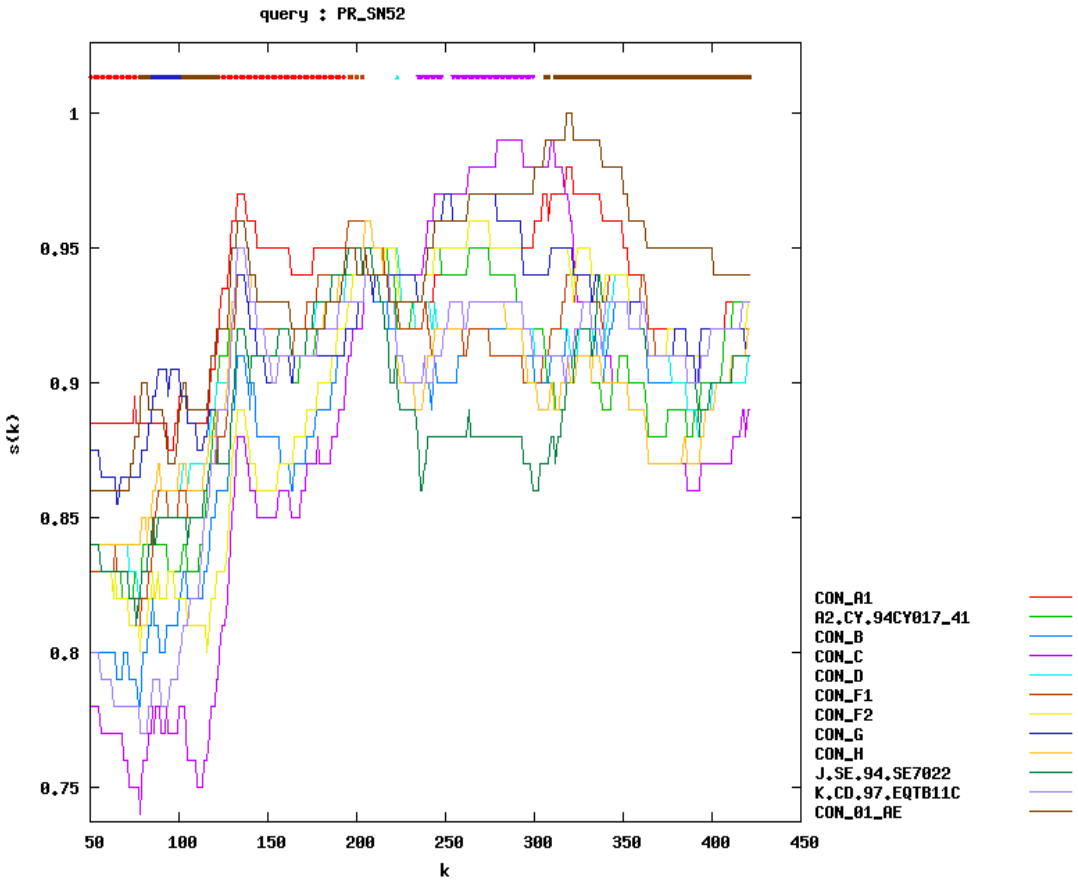


Figure 6.6 Sample PR_SN52 screened rapidly for intersubtype recombination using RIP online program. The background contains consensus and subtype reference sequences of HIV-1 subtypes with subtype A1 (in red line) being the closest to the sample sequence.

6.3.4 SCUEAL (DataMonkey program)

The DataMonkey subtyping program characterized all 59 and 55 sample sequences for PR and RT fragments, respectively.

6.3.5 Context-based Modeling for Expeditious Typing (COMET)

The COMET characterized all 59 and 55 sample sequences for PR and RT, respectively.

6.3.6 Stanford HIV drug database Subtyping program

All sample sequences that included 59 for PR and 55 RT fragments were assigned subtypes accordingly.

6.3.7 Phylogenetic analysis

A phylogenetic evaluation for 53 *pol* RT samples using neighbor-joining tree was constructed using MEGA V5 with a 1000 bootstrap replicates (Figure 6.7). Majority of the sample sequences clustered with HIV-1 subtype G/CRF02_AG which accounted for 58.5% (n = 31/53), CRF22_01A1 10/53 (18.9%), C 4/53(7.5%) F2 2/53 (3.8%) CRF11_cpx 3/53 (5.7%) D and other complex forms 1/53(1.9% each). Two of the sequenced RT samples could not be assigned to any subtype with confidence.

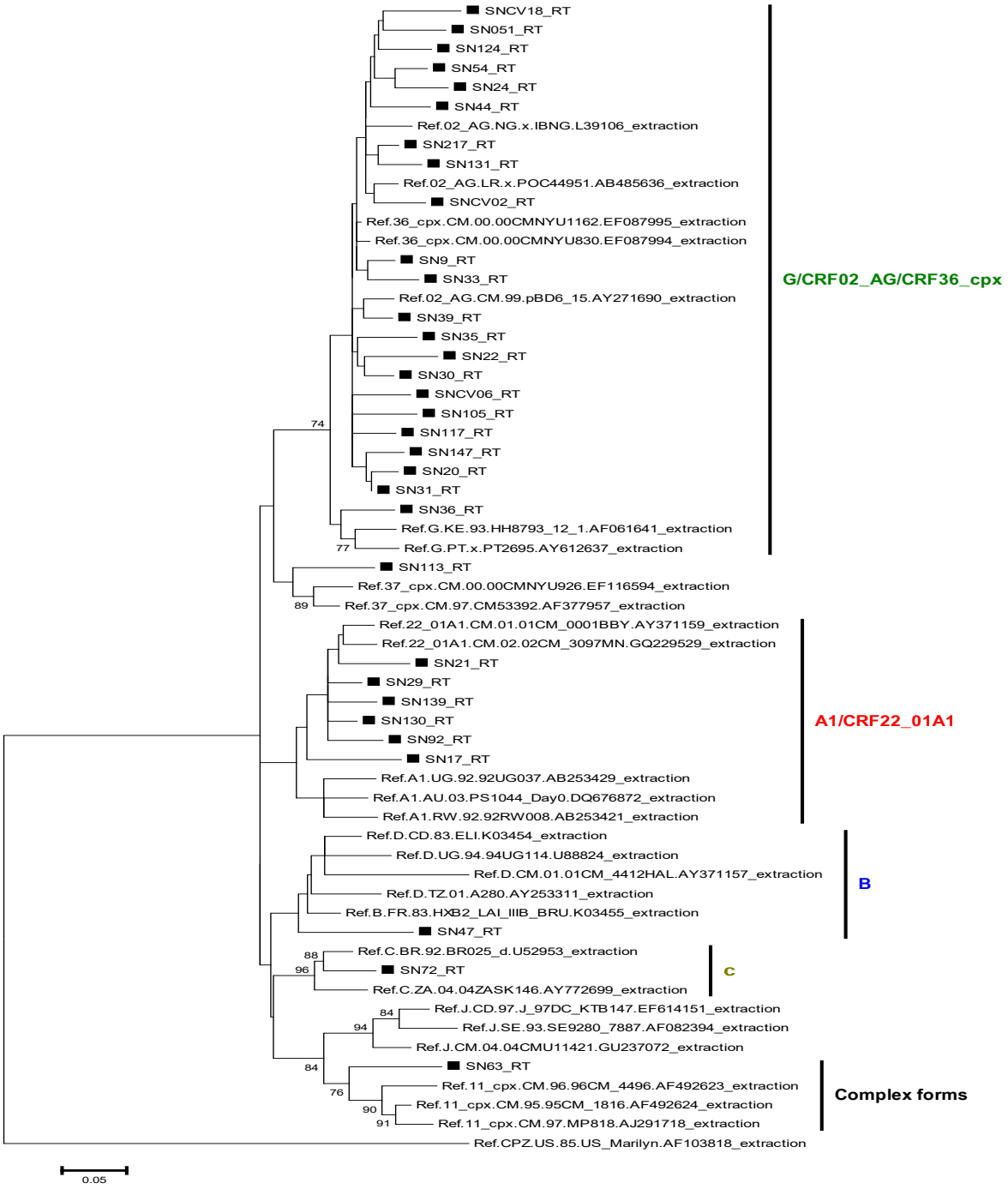


Figure 6.7. A neighbor-Joining phylogenetic reconstruction of 450 bp *pol* (RT) sequence and a bootstrap of 1000 replications. HIV subtypes are indicated with specific colours and include G/CRF02_AG which accounted for 40% (n = 22), CRF22_01A1 (10.9%; n = 6), C (1.8%; n = 1), B (1.8%; n = 1), other complex forms – 37_cpx/11_cpx (3.6%; n = 2). Twenty three samples (41.8%) could not be assigned to any subtype with confidence.

6.3.8 HIV-1 resistance associated mutations in the pol gene (PR and RT)

There were no major PIs RAMs detected in all analysed sequences for the PR fragment. However, in the RT fragment, three patients had NRTI RAMs and include; patient **SN35** who had; A62V, K65R and M184V/I mutations, patient **SN42** with M184V/I mutation while patient **SN47** had D67N, K70R, T215F and K219E mutations. In the NNRTI category, RAMS were noticed also in three patients whereby, patient **SN35** had K103N, Y188C and M230L mutations, while patient **SN47** and **SN52** had Y181C and G190R mutation respectively. A summary of level of drug resistance for the four patients identified is shown below in Figure8 6.8.

Adults: levels of drug resistance against NRTIs and NNRTIs

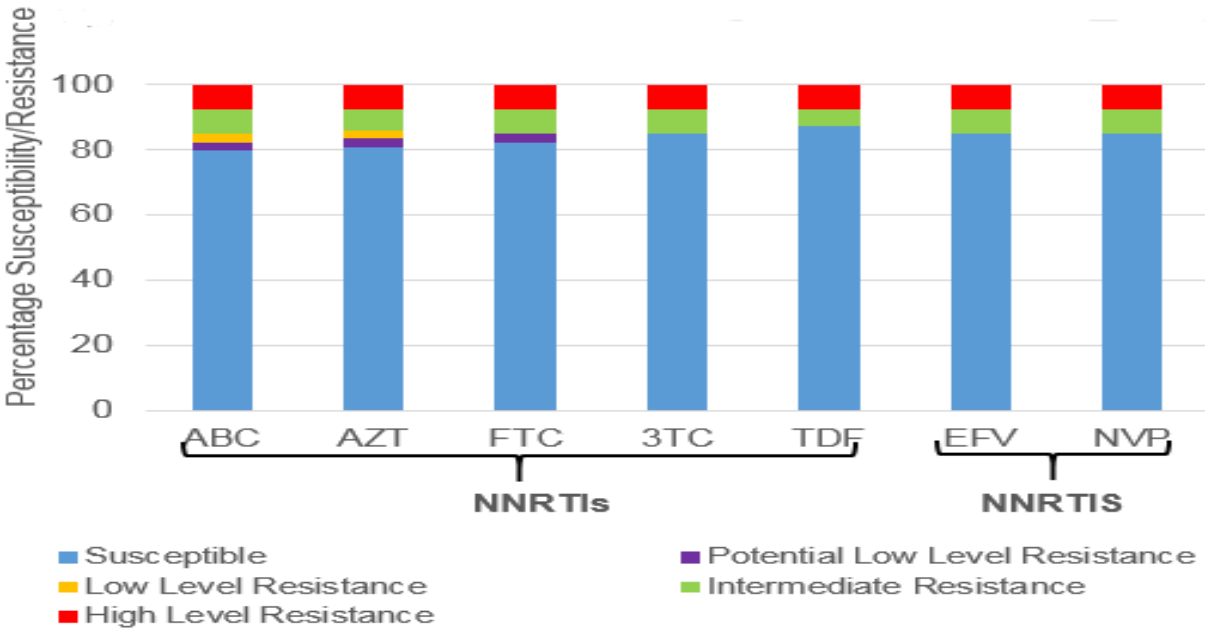


Figure 6.8 Observed NRTI and NNRTI drug resistance in adults with over 83.0% susceptible to commonly used cART. Approximately 4.0% of analysed samples for this cohort had intermediate to high level resistance to all classes of drugs available while less than 1.0% had Low to potential low level resistance to ABC, AZT and FTC.

CHAPTER 7 Discussion – Cohort II

7.1 Overview of Cohort II study

In this Cohort, HIV-1 diversity and drug resistance analysis were performed on 73/125 (58.4%) of therapy naïve adult male and female patients. The analyses were based on the partial *pol* gene fragments (PR and RT). In the preliminary HIV subtyping, six online subtyping programs that included jpHMM, REGA, COMET, SCUEAL, RIP and Stanford were used. The observed subtypes were further confirmed using a NJ phylogenetic inference as in cohort I. Genotypic drug RAMs were analyzing using Stanford drug resistance data base in the two partial *pol* gene fragments (PR and RT).

7.2 Observed HIV diversity

The online subtyping programs clustered the majority of subtypes with similar subtypes. The jpHMM, COMET, SCUEAL, RIP and Stanford subtyping programs inferred subtypes to more than 90.0% of submitted sample sequences in all the two *pol* gene fragments (PR and RT). However, with REGA subtyping program, less than 40.0% of submitted sample sequences subtypes could be inferred for subtypes. Noticeable difference in subtype inference was observed in two sample sequences of patient SN-47 and SN-54. In sample sequence of patient SN-47, a pure subtype D was inferred by jpHMM, COMET, SCUEAL and Stanford, while recombinant form BD was inferred by RIP and unknown by REGA subtyping program. In patient SN-54, the sample sequence was inferred as A1C, C/G and G by jpHMM, RIP and Stanford, respectively, while both COMET and Stanford inferred subtype G with RIP subtyping tool unable to infer any subtype. Although such small inconsistencies are observed

due to reasons discussed in chapter 4, majority of these subtyping programs provided similar subtypes across all sequences analyzed. A NJ phylogenetic analysis was performed to confirm the inferred subtypes by online subtyping programs. The results yielded by NJ phylogenetic tree were consistent with the majority of subtypes as observed in the online programs. The two samples that had a mixed subtype inference - sample sequence of SN-47 and SN-54 were phylogenetically inferred as subtype D and CRF02_AG, respectively. Results obtained in this cohort show the majority of sample sequences were G/CRF02_AG, accounting for 40% of all observed subtypes. The remaining subtypes were distributed among other subtypes and complex forms that included CRF22_01A1 (10.9%; n = 6), C (1.8%; n = 1), B (1.8%; n = 1), other complex forms – 37_cpx/11_cpx (3.6%; n = 2). Twenty three samples (41.8%) could not be assigned to any subtype with confidence. In previous studies done in both urban and rural areas of Cameroon, CRF02_AG was predominant accounting for more than 40.0% in their study cohorts of therapy naïve adult patients (Koizumi *et al*., 2006; Burda *et al*., 2010; Hemelaar *et al*., 2013). The findings of this study further reinforce the presence of high diversity and other circulating recombinant forms in Cameroon. Subtype C, F2 and 11_cpx and 13_cpx have mostly been observed as rare subtypes in rural areas of Cameroon (Powell *et al*., 2010), but their continued presence raises concern on transmission to other urban centres. The low prevalence of these subtypes could be due lack of standardised laboratory assays for non-B variants and this means that they would be higher in circulation than previously thought. Subtype C causes more than 50.0% of all HIV-1 infections globally with countries such as South Africa and India being the most affected by this subtype (Goudsmit & Jaap., 1997; Jacobs *et al*., 2009). The

presence of these rare subtypes would be evolving, and this might change and lead to an outbreak after adapting themselves well in this region.

7.3 Observed drug resistance associated mutations

As described earlier in cART management, Cameroon has adopted the 2013 WHO guidelines for HIV treatment in adults that include NRTIs – AZT, ABC, 3TC, FTC and TDF which and NNRTI that consist of NVP and EFV. The PIs for adults include Lopinavir boosted ritonavir or Atazanavir boosted ritonavir are preferred. In this study cohort RAMs observed against NRTIs were from three patients that include patient SN35 with A62V, K65R and M184V/I mutations, patient SN42 with M184V/I mutation while patient SN47 had D67N, K70R, T215F and K219E mutations. The K65R, M184V/I, K70R and T215F mutations show intermediate to high level resistance to AZT, TDF, ABC, 3TC and FTC drugs while D67N, K219Q/E mutations show low level to reduced resistance to AZT and A62V combines with K65R to increase resistance to NRTIs. In the NNRTI category, three patients also showed drug resistance mutations that were from patient SN35 with K103N, Y188C and M230L mutations, while patient SN47 had Y181C mutation and SN52 with G190R mutation. All these mutations to NNRTIs had intermediate to high level resistance to NVP and EFV, which are routinely used in Cameroon. A previous study conducted in Cameroon on drug resistance mutations in adult naïve patients (Ceccarelli *et al*., 2012; Agyingi *et al*., 2014) show that the K65R, M184V and Y188C were common mutations observed. As observed in this study cohort, the majority of patients (more than 95.0%) were susceptible to the adopted cART in Cameroon, while a small percentage of less than 2.5% patients had RAMs. The prevalence of drug resistance

in therapy naïve adults was 5.4% (3/56) for both NRTIs and NNRTIs. This prevalence is low compared to previous studies done in Cameroon on therapy naïve adult patients that showed a prevalence of between 4.0% and 8.0% for both NRTI and NNRTIs, respectively (Ndembi *et al*., 2008; Aghokeng *et al*., 2009; Agyingi *et al*.,2014). Similar studies in Africa on RT drug RAMs in adults show a prevalence of 1-11% with countries such as Kenya at 7.4% (Onywera *et al*., 2017), South Africa with 3.6% (Hamers *et al*., 2011) and Mali 11.5% (Derache *et al*., 2008). Since RT fragments could not be obtained for all 125 samples included in the study, the NRTI and NNRTI drug resistance would be underestimated or overstated but the results are comparable with other study cohorts in Cameroon. The PI mutations were minor with no RAMs.

CHAPTER 8 General discussion and conclusion

8.1 General discussion

The HIV-1 subtypes and recombinant forms display different phenotypic properties that are associated with viral evolution, entry through co-receptor utilization, efficient transmission, disease progression, drug resistance and vaccine development (Connor *et al*., 1997; Kiwanuka *et al*., 2010). This study focussed on HIV diversity and drug resistance based on the *pol* gene fragment in both Cohorts. The *pol* gene is highly polymorphic and characterized with majority of RAMs. The PIs, NRTIs/NNRTIs and INIs are designed to target these RAMs during the HIV life cycle. (Vondrasek *et al*., 1997; Gea- Banacloche *et al*., 2000). One of the major difficulties in managing the HIV-1 pandemic is high viral genetic variability. Globally, group M subtype C is the most prevalent strain accounting for 50.0% of all HIV infections. This is followed with subtype A at 12.3% and B, G, CRF01_AE, and CRF02_AG at 10.2%, 6.3% 4.7% and 5% respectively (Hemelaar *et al*., 2011)**.** The CRF02_AG is progressively causing more HIV infections in the West and Central Africa region where it accounts for 27.8 and 3.9% respectively (Taylor *et al*., 2008). In this study**,** two groups of HIV were identified- group M and group O**.** In group M**,** a high degree of viral diversity was observed with CRF02_AG accounting for 62.5% and 40.0% for cohort I and cohort II, respectively. This confirms the high prevalence of CRF02_AG strain in Cameroon which accounts for over 40.0% of HIV infections (Aghokeng *et al*., 2013). The preference towards CXCR4 receptor usage by CRF02_AG is thought to greatly contribute to its successful transmission. In a study done in Guinea-Bissau in 2010, it showed that the frequency of CXCR4 tropism in therapy naïve patients was at 86.0% (Esbjornsson *et al*., 2010). In another study that was done to evaluate competition

assays between CCR5 and CXCR4 viruses, it showed that the CXCR4 viruses have higher replication kinetics (Tebit *et al*., 2007). This could possibly explain the high frequency of CRF02_AG transmission in countries like Cameroon. More research is necessary to understand comprehensively the mechanism behind the CRF02_AG and other subtypes' high frequency transmission. Other HIV group M subtypes observed in this study include subtype A at 20.8%, F2 and C subtypes at 8.3% for cohort I and CRF22_01A1,B, C, B other complex forms – 37_cpx/11_cpx), at 10.9%, 1.8%, 1.8% and 3.6% respectively. HIV group O was identified in one child based on the IN fragment in Cohort I. HIV group O is thought to have resulted from cross-species transmission of SIVs that are hosted by lowland gorillas from the West and Central Africa region especially in southern Cameroon (D'Arc *et al*., 2015). It is the second prevalent HIV group after group M, accounting for about 1.0% of all HIV infections globally. Majority of the group O infections have been observed in Cameroon ([a]Vessiere *et al.,* 2010). In another study done in Cameroon, HIV group O was found to co-circulate with group M confirming the evolutionary dynamics of the HIV (Villabona-Arenas *et al*., 2015). The prevalence of HIV group O in Cameroon would be under estimated due to lack of appropriate diagnostic assays (Aghokeng *et al*., 2009). Likewise, in this study, group O and possibly other new recombinants would have been missed in other samples due to the small *pol* fragment that was used for analysis. A full length genome characterization will be necessary for complete characterization of the samples. Prevalence rate of RAMs was at 3.7% and 5.4% for ohort I and Cohort II respectively. Unfortunately, few studies have been done on drug resistance, especially on therapy naïve infants/children in resource-limited countries like Cameroon (Sánchez PR. & Holguín A., 2014). A study

conducted in Yaoundé, Cameroon between 1996 and 2007 shows an ever increasing drug resistance in therapy naïve individuals over time - 1.9%, 4.1% and 12.3% in 2001, 2002 and 2007 respectively (Aghokeng *et al*., 2011). In another study done in Cameroon in 2011 by Fokam and others, drug resistance in therapy naïve pediatric patients was ranging between 4.9% to 8.2% while high levels (9.0%) in patients failing first-line cART was observed (Fokam *et al*., 2011; Ceccarelli *et al*., 2012). In therapy naïve adult patients from Cameroon, previous studies have shown a prevalence of between 4.0% and 8.0% for both NRTI and NNRTI RAMs (Ndembi *et al.,* 2008; Aghokeng *et al*., 2009; Agyingi *et al*., 2014). The high level of RAMs in these previous studies would be due to high number of sample size as compared to what was analysed in this study. These findings suggest that there is a need to standardize resistance testing protocols, especially where there is a high endemic genetic diversity of HIV. Thus, infrastructure and access to facilities in rural Cameroon needs to be improved. Currently, the FCB is one of the few facilities that perform resistance testing in Cameroon. Since the current estimate shows that over a million children will require cART by 2020 (Penazzato*et al*., 2015), more resistance testing facilities are necessary to better manage these patients. The WHO programs should be robustly implemented, as they look to support regimens that are more simplified, less toxic, have higher genetic barriers and would ultimately require less clinical monitoring for each patient, while maintaining therapeutic efficacy.

8.2 Limitations

The study limitations included lack of follow-up and treatment and adherence data from the mothers. The relatively small sample size of the study cohort which was mainly hampered by difficulties in enrollment process, majority of the babies were under critical clinical observation with history of severe anaemia. Most caregivers also refused to participate in the study. Unsuccessful viral RNA amplification of nearly half of the samples further reduced the sample size. The major contributory factor for inadequate viral RNA amplification was difficulties in designing the right primers and the high cost attached to it. The PR and IN sequences in both cohorts were not long enough for a phylogenetic analysis and this was attributed to the difficulties in amplifying and sequencing samples with high diversity.

8.3 Conclusion

This study provides a summary of circulating HIV-1 group M subtypes plus a rare group O that was observed in infants/children. It also provides observed drug resistance patterns in the two different cohorts of therapy naïve patients in Cameroon. This study reveals that Cameroon might be harboring many HIV-1 types and subtypes as observed in both cohorts with the most prevalent HIV-1 clade being G/CRF02_AG with over 40.0%. Rare group O and other group M subtypes like subtype C were detected even in the small sample size, suggesting the improved detection methodologies currently used. This possibly suggests that there could be high presence of these assumed rare subtypes if more and reliable assays are

performed. Subtype C is especially common in South Africa but the detection of this subtype in west and central Africa, raises the possibility of transmission that may in future lead to an increase in the prevalence of subtype C. HIV diversity has implications on diagnostic assays as most assays developed are based only on HIV-1 subtype B which circulates in N. America & Europe. Further, HIV diversity may have implications in transmission, clinical management of the patient and vaccine development. Molecular studies in HIV have given an insight into HIV origin, how it has evolved over time and its global spread. New HIV infections are being reported and may be attributed to the ever evolving HIV. Continuous surveillance of the HIV diversity is therefore necessary to better manage the HIV-1 pandemic. Existence of drug resistance mutations especially in infants raises grave concerns on the availability and effectiveness of the current PMTCT program. The 90-90-90 program advocated by various governments and multi-international agencies should be strengthened. This will make it easier for newly infected individuals to know their HIV status and help to stop further transmission. There is also need to have improved, cost effective HIV diagnostic and resistant assays for the region.

8.4 Future research investigations

Since the sample size was a major limitation, a larger sample size will give better and significant statistical figure that can be used especially in policy formulation. Future studies will aim to include more sites in Cameroon for a better view of the prevalence in HIV diversity and drug resistance. Future investigations will also seek to sequence the whole genome to identify any recombination breaks/events that would have been missed due to the small *Pol* fragment that was analysed in this study. Some studies

in Cameroon have found recombinants of group M and O which informs the need to do a complete sequence analysis for further characterization. Drug resistance is a major challenge to currently used antiretroviral drugs as illustrated in this study by the detection of RAMS. Further research will be necessary especially in expectant mothers so as to further prevent the transmission of drug-resistant strains to unborn babies. This requires continuous analysis of evolving RAMs that could jeopardize the successful implementation of cART programs.

References

Aberg J., Gallant J., Ghanem K., Emmanuel P., Zingman B., Horberg M. (2014). Primary care guidelines for the management of persons infected with HIV: 2013 update by the HIV Medicine Association of the Infectious Diseases Society of America. *J. Clin Infect Dis*; 58 (1):

Agyingi L., Mayr L., Kinge T., Orock G., Ngai J., Asaah B., Mpoame M., Hewlett I., Nyambi P., (2014). The evolution of HIV-1 group M genetic variability in Southern Cameroon is characterized by several emerging recombinant forms of CRF02_AG and viruses with drug resistance mutations. *J. Med Virol*; 86 (3): 385 - 93.

Aghokeng, A., Vergne, L., Mpoudi-Ngole, E., Mbangue, M., Deoudje, N., Mokondji, E.,Nambei, W., Peyou-Ndi M., Moka, J., Delaporte, E., Peeters, M. (2009). Evaluation of transmitted HIV drug resistance among recently-infected antenatal clinic attendees in four Central African countries. *J. Antivir Ther*; 14, 401-411.

Aghokeng AF., Kouanfack C., Laurent C., Ebong E., Atem-Tambe A., Butel C., (2011). Scale-up of antiretroviral treatment in sub-Saharan Africa is accompanied by increasing HIV-1 drug resistance mutations in drug-naive patients. *AIDS*, 25, pp. 2183-2188.

Aghokeng A., Kouanfack C., Eymard-Duvernay S., Butel C., Edoul G., Laurent C., Koulla-Shiro S., Delaporte E., Mpoudi-Ngole E., Peeters M., (2013). Virological outcome and patterns of HIV-1 drug resistance in patients with 36 months' antiretroviral therapy experience in Cameroon. *J. Int AIDS Soc*; 16: 18004.

AIDSinfo / UNAIDS. (2015) Global factsheet; Epidermiolagical slides. (http://aidsinfo.unaids.org/) – accessed on 15th June,2016.

Alcantara L., Cassol S., Libin P., Deforche K., Pybus O., Van Ranst M., Galvão-Castro B., Vandamme A., de Oliveira T. (2009). A standardized framework for accurate, high-throughput genotyping of recombinant and non-recombinant viral sequences. Nucleic Acids Res.; 37 (Web Server issue): W634 - 42.

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

Arriaga, M, Carr J., Li P., Wang B., Saksena, N., (2006). Interaction between HIV-1 and APOBEC3 sub-family of proteins. *Curr HIV Res* 4, 401-409.

Arthos J., Cicala C., Martinelli E., Macleod K., Van Ryk D., Wei D., Xiao Z., Veenstra T., Conrad T., Lempicki R., McLaughlin S., Pascuccio M., Gopaul R., McNally J., Cruz C., Censoplano N., Chung E., Reitano K., Kottilil S., Goode D., Fauci A., (2008). "HIV-1 envelope protein binds to and signals through integrin alpha(4)beta(7), the gut mucosal homing receptor for peripheral T cells". *Nat. Imm*. In Press (3): 301–9.

Ayouba A., Souquieres B., Njinku P., Martin M., Muller-Trutwin P., Roques F., Barre-Sinoussi P., Mauclere F., et al. (2000). HIV-1 group N among HIV-1-seropositive individuals in Cameroon. *J. AIDS;* 14:2623-2625.

Baeten J., Chohan B., Lavreys L., Chohan V., McClelland R., Certain L., Mandaliya K., Jaoko W., Overbaugh J., (2007). HIV-1 subtype D infection is associated with faster disease progression than subtype A in spite of similar plasma HIV-1 loads. *J Infect Dis*; 195: 1177– 1180.

Barré-Sinoussi F., Chermann J., Rey F., Nugeyre M., Chamaret S., Gruest J., Dauguet C., Axler-Blin C., Vézinet-Brun F., Rouzioux C., Rozenbaum W., Montagnier L.(1983). Isolation of a T-Lymphotropic retrovirus from a patient at risk for acquired immune-deficiency syndrome *AIDS*. *J. Science*; 220: 868-871.

Bennett D., Camacho R., Otelea D., Kuritzkes D., Fleury H., Kiuchi M., et al. (2009). Drug Resistance Mutations for Surveillance of Transmitted HIV-1 Drug-Resistance. 2009 Update. PLoS ONE 4(3).

Billong S., Fokam J., Aghokeng A., Milenge P., Kembou E., Abessouguie I., Meva'a-Onglene F., Bissek A., Colizzi V., Mpoudi E., Elat J., Shiro K., (2013). Population-based monitoring of emerging HIV-1 drug resistance on antiretroviral therapy and associated factors in a sentinel site in Cameroon: low levels of resistance but poor programmatic performance. PLoS One; 8 (8).

Bloom A., (1984). Acquired immunodeficiency syndrome and other possible immunological disorders in European haemophiliacs. *J. Lancet 1*; 1452-1455.

Broder and Gallo (1984). A pathogenic retrovirus (HTLV-III) linked to AIDS. *N Engl J Med*.; 311:1292–1297.

Bulla I., Schultz A., Meinicke P., (2012). Improving Hidden Markov Models for classification of human immunodeficiency virus-1 subtypes through linear classifier learning. *Stat Appl Genet Mol Biol*; 11 (1).

Burda S., Viswanath R., Zhao J., Kinge T., Anyangwe C., Tinyami E., Haldar B., Powell R., Jarido V., Hewlett I., and Nyambi P., (2010). HIV-1 Reverse Transcriptase Drug-Resistance Mutations in Chronically Infected Individuals Receiving or Naïve to HAART in Cameroon. *J. med. virol*; 82 (2):187-196.

Bushman FD, Fujiwara T. and Craigie R. (1990). Retroviral DNA integration directed by HIV integration protein in vivo. *Science* 249:1555–1558.

Carr J., Nathan D., Judith N., Ubald T., Mpoudi-Ngole E., Lindsay E., (2010). HIV-1 recombinants with multiple parental strains in low-prevalence, remote regions of Cameroon: evolutionary relics? *J. Retrovirology*; 7:39.

Case K. (1986). Nomenclature: human immunodeficiency virus. Ann Intern Med. ;105(1):133.

CDC, (September 24, 1982). Current trends update on acquired immune deficiency syn-drome (AIDS)—United States. MMWR. 31(37): 507-508, 513-514. Available at: www.cdc.gov/mmwr/preview/mmwrhtml/00001163.htm - accessed on 23rd June 2017.

Ceccarelli L., Salpini R., Moudourou S., Cento V., Santoro M., Fokam J., et al., (2012). Characterization of drug resistance mutations in naïve and ART-treated patients infected with HIV-1 in Yaounde, Cameroon. *J. Med. Virol;* 84: 721–727.

Centers for Disease and Control (CDC). (1983). 'Current Trends Update: Acquired Immunodeficiency Syndrome (AIDS) - United States' MMWR Weekly 32(35):465-467.

Centers for Disease Control. (1982). "Opportunistic infections and Kaposi's sarcoma among Haitians in the United States". MMWR Morb Mortal Wkly Rep. 31 (26): 353–354; 360–361.

Chaix M.L., Desquilbet L., Descamps D., et al., (2007). Response to HAART in French patients with resistant HIV-1 treated at primary infection: ANRS Resistance Network. *Antivir Ther*.;12(8):1305–10.

Chan DC and Kim PS (1998). "HIV entry and its inhibition". Cell. 93 (5): 681–4.

Chiu, T. K., and D. R. Davies. (2004). Structure and function of HIV-1 integrase. *Curr Top Med Chem;* 4:965-977.

Clapham P. and McKnight A. (2002). Cell surface receptors, virus entry and tropism of primate lentiviruses. *J. Gen Virol*; 83: 1809–29.

Clavel F., Guetard F., Brun-Vezinet S., Chamaret M., MO S., Laurent C., et al., (1986). Isolation of a new human retrovirus from West African patients with AIDS. *J. Science*; Vol. 233, Issue 4761, pp. 343-346.

Clavel F.and Hance A. (2004). Medical progress: HIV drug resistance. *N. Engl J. Med*; 350:1023–1035.

Collier AC, Coombs RW, Schoenfeld DA, Bassett RL, Timpone J, Baruch A, et al., (1996). Treatment of human immunodeficiency virus infection with saquinavir, zidovudine, and zalcitabine. *N Engl J Med.;* 334:1011–1018.

Corado AdLG, Bello G, Leão RAC, Granja F, Naveca FG (2017). HIV-1 genetic diversity and antiretroviral drug resistance among individuals from Roraima state, northern Brazil. PLoS ONE12(3).

Connor R., Sheridan K., Ceradini D., Choe S., Landau N., (1997). Change in coreceptor use correlates with disease progression in HIV-1--infected individuals. *J Exp Med;* 185:621-628.

Cunningham A., Donaghy H., Harman A., Kim M., Turville S., (2010). Manipulation of dendritic cell function by viruses". *Current opinion in microbiology*. 13 (4): 524-529.

D'arc M., Ayouba A., Esteban A., Learn G., Boué V., Liegeois F., et al., (2015). Origin of the HIV-1 group O epidemic in western lowland gorillas. *Proc Natl Acad Sci U S A*. 2015; 112(11): E1343–E1352.

Dash K., Siddappa B., Mangaiarkarasi A., Mahendarkar V., Roshan P., Anand K., et a.,(2008). Exceptional molecular and coreceptor-requirement properties of molecular clones isolated from an Human Immunodeficiency Virus Type-1 subtype C infection. *Retrovirology*, *5*: 25.

Davaro R., and Thirumalai A. (2007). Life-threatening complications of HIV infection. *J. Intensive Care Med*; 22(2):73–81.

Derache A., Maiga A., Traore O., Akonde A., Cisse M., Jarrousse B., et al., (2008). Evolution of genetic diversity and drug resistance mutations in HIV-1 among untreated patients from Mali between 2005 and 2006. *J Antimicrob Chemother*. 2008;62(3):456–463.

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

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

de Oliveira T., Deforche K., Cassol S., Salminen M., Paraskevis D., Seebregts C.,et al.,.(2005). An automated genotyping system for analysis of HIV-1 and other microbial sequences. *Bioinformatics*.; 21 (19): 3797 - 800.

Delport W., Poon A., Frost S., Kosakovsky P., (2010): a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics*; 26 (19): 2455 - 7.

Douek D., Roederer M., Koup R., (2009). Emmerging Concepts in the Immunopathogenesis of AIDS. *Annu. Rev. Med*. 60: 471-84.

Dow D., & Bartlett J. (2014). Dolutegravir, the Second-Generation of Integrase Strand Transfer Inhibitors (INSTIs) for the Treatment of HIV. Infectious Diseases and Therapy. *J. Infect Dis Ther*; 3(2), 83–102.

Embretson J., Zupacic M., Ribas L., Burke A., Racz P., Tenner-Racz K. et al. (1993). Massive covert infection of helper T-lymphocytes and macrophages by HIV during the incubation period of AIDS. *Nature*;362, 359-62.

Esbjornsson J., Månsson F., Martínez-Arias W., (2010) 'Frequent CXCR4 Tropism of HIV-1 Subtype A and CRF02_AG During Late-Stage Disease—Indication of an Evolving Epidemic in West Africa', *J. Retrovirology;* 7: 23.

Evering T. and Markowitz M. (2007). Raltegravir (MK-0518): an integrase inhibitor for the treatment of HIV-1. *Drugs Today*; 43, 865–877.

Faria N., Rambaut A., Suchard M., Baele G., Bedford T., Ward M., et al., (2014). HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. *J.Science*; 346 (6205): 56 - 61.

Feng D.F. & Doolittle R.F., (1987). J. Mol. Evol.; vol. 25 (pg. 351-360).

Fischl M., Richman D., Grieco M., Gottlieb M., Volberding P., Laskin O., (1987). The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. A doubleblind, placebo-controlled trial. *N Engl J. Med*; 317, 185-191.

Fokam J, Salpini R, Santoro MM, Cento V, Perno CF, Colizzi V., et al., (2011). Drug resistance among drug-naive and first-line antiretroviral treatment-failing children in Cameroon. *Pediatr Infect Dis J*.; 30:1062-8.

Fokam J., Billong S., Bissek A., Kembou E., Milenge P., Abessouguie I., et al., (2013). Declining trends in early warning indicators for HIV drug resistance in Cameroon from 2008-2010: lessons and challenges for low-resource settings. *BMC Public Health*. 8; 13.

Fonjong L. (2001). Fostering women's participation in development through non-governmental efforts in Cameroon. *J. Geograpgy*; 167:223-234.

Fonsah J., Njamnshi A., Kouanfack C., Qiu F., Njamnshi D., Tagny C., et al., (2017). Adherence to Antiretroviral Therapy (ART) in Yaoundé-Cameroon: Association with Opportunistic Infections, Depression, ART Regimen and Side Effects. PLoS ONE 12(1).

Forsman A. and Weiss R.A.,(2008). Why is HIV a pathogen? *Trends Microbiol*.;16, 555-560.

Garcia-Calleja J., Zekeng L., Louis J., Trebucq A., Salla R., Owona R., et al., (1992). HIV infection in Cameroon: 30 months' surveillance in Yaounde. *J. AIDS;* 6:881–882.

Gakhar H., Kamali A. and Holodniy M., (2013). Health-related quality of life assessment after antiretroviral therapy: a review of the literature. *J. Drugs*; 73 (7): 651-72.

Gallo R., Sarin E., Gelmann M., Robert-Guroff E., Richardson V., Kalyanaraman D., (1983). Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *J. Science*; 220:865-867.

Galvani, A. P., and J. Novembre. (2005). The evolutionary history of the CCR5-Delta 32 HIV-resistance mutation. *Microbes and Infect.* 7:302–309.

Gea-banacloche J., Migueles S., Martino L., Shupert W., Mcneil A., Sabbaghian M., et al., (2000). Maintenance of large numbers of virus-specific CD8+ T cells in HIV-infected progressors and long-term nonprogressors. *J. Immunol*; 165(2)1082-92.

Geretti AM, Harrison L, Green H, Sabin C, Hill T, Fearnhill E, et al., (2009). Effect of HIV-1 subtype on virologic and immunologic response to starting highly active antiretroviral therapy. *Clin Infect Dis* 48:1296–1305.

Geretti A., Armenia D. and Ceccherini-Silberstein F. (2012). Emerging patterns and implications of HIV-1 integrase inhibitor resistance. Curr. Opin. *J. Infect Dis*; 25:677–686.

German Advisory Committee Blood (Arbeitskreis Blut), Subgroup "Assessment of Pathogens Transmissible by Blood." (2016). Human Immunodeficiency Virus (HIV). Transfusion Medicine and Hemotherapy, 43(3), 203–222.

Gheysen D., Jacobs E., de Foresta F., Thiriart C., Francotte M., Thines D., and De Wilde M. (1989). Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirusinfected insect cells. *Cell* 59, 103–112.

Grobler J., Stillmock K., Hu B., Witmer M., Felock P., Espeseth A., Hazuda D. (2002). Diketo acid inhibitor mechanism and HIV-1 integrase: Implications for metal binding in the active site of phosphotransferase enzymes. *Proceedings of the National Academy of Sciences of the United States of America*; 99(10), 6661–6666.

Gordon L., Gharibian D., Chong K., Chun H., (2015). Comparison of HIV virologic failure rates between Patients with Variable Adherence to Three Antiretroviral Regimen Types. AIDS Patient Care STDS. 29 (7):384–8.

Gottlieb M., Schanker H., Fan P., Saxon A. and Weisman J., (1981). Pneumocystis Pneumonia - Los Angeles. *MMWR* 30, 250-252.

Gottlieb M., Schroff R., Schanker H., Weisman J, Fan P., Wolf R., et al., (1981). Pneumocytis Carinii pneumonia and mucosal candidiasis in previously healthy homosexual men. N Eng. *J. Med;* 305, 1425-1431.

Goudsmit and Jaap. (1997). Viral Sex; The Nature of AIDS. Oxford University Press., New York, Pg. 51-58.

Götte M.; Li X.; Wainberg, M (1999). HIV-1 reverse transcription: A brief overview focused on structure-function relationships among molecules involved in initiation of the reaction. *Arch Biochem. Biophys; 365*, 199–210.

Graziosi C., Pantaleo G., Butini L., Demarest J., Saag M., Shaw G., et al., (1993). Kinetics of human immunodeficiency virus type 1 (HIV-1) DNA and RNA synthesis during primary HIV-1 infection. *J. Proc Natl Acad Sci;* 90:6405-6409.

Grant R., Hecht F., Warmerdam M., Liu L., Liegler T., Petropoulos C., et al., (2002). Time Trends in Primary HIV-1 Drug Resistance Among Recently Infected Persons. *J. AMA*; 288(2):181–188.

Gupta R., Jordan M., Sultan B., Hill A., Davis D., Gregson J., et al., (2012). Global trends in antiretroviral resistance in treatment-naive individuals with HIV after rollout of antiretroviral treatment in resource-limited settings: a global collaborative study and meta-regression analysis. *J. Lancet;* 380 (9849):1250-1258.

Hahn B., Shaw G., De Cock K., Sharp P., (2000). AIDS as a zoonosis: Scientific and public health implications. *J. Science*; 287:607-614.

Hamers R., Wallis C., Kityo C., Siwale M., Mandaliya K., Conradie F., et al., (2011). HIV-1 Drug Resistance in Antiretroviral-Naïve Individuals in sub-Saharan Africa after Rollout of Antiretroviral Therapy: A Multicentre Observational Study. *J. Lancet Infect Dis*; 11(10):750–9.

Hamers RL, Sigaloff KC, Wensing AM, Carole LW, Kityo C., Siwale M, et al., (2012). Patterns of HIV-1 drug resistance after first-line antiretroviral therapy (ART) failure in 6 sub-Saharan African countries: implications for second-line ART strategies, *Clin Infect Dis*, 2012, vol. 54 (pg. 1660-9).

Hazuda D., Young S., Guare J., Anthony N., Gomez R., Wai J., et al., (2004). Integrase inhibitors and cellular immunity suppress retroviral replication in rhesus macaques. *J. Science*; 305: 528-532.

Hellmann N., Johnson P. and Petropoulos C. (1999). Validation of the performance characteristics of a novel, rapid phenotypic drug susceptibility assay, PhenoSense HIV [abastract 51]. *J. Antivir Ther*; 4 (Suppl. 1):34-5.

Hemelaar J., (2013). Implications of HIV diversity for the HIV-1 pandemic. *J. Infect Dis.*; 66: 391–400.

Hemelaar J., Gouws E., Ghys P., Osmanov S., W-UNHI C., (2011). Global trends in molecular epidemiology of HIV-1 during 2000-2007. *J. AIDS;* 25: 679-689.

Holguin A., Lopez M., Soriano V., (2008). Reliability of rapid subtyping tools compared to that of phylogenetic analysis for characterization of human immunodeficiency virus type 1 non-B subtypes and recombinant forms. J. of Clinical Micro. 46 (2008), pp. 3896-3899.

Hughes A and Corrah T. (1990). Human immunodeficiency virus type 2 (HIV2). *Blood Rev*.;4(3):158–164.

Hunt P., (2009). Natural control of HIV-1 replication and long-term nonprogression: overlapping but distinct phenotypes. *J. Infect Dis*; 200:1636–8.

Hymes K., Cheung T., Greene J., Prose N., Marcus A., Ballard H., et al., (1981). Kaposi's sarcoma in homosexual men - a report of eight cases. *J. Lancet* 2, 598-600.

http://knoema.com/WBWDIGDF2015Aug/world-development-indicators-wdi-september-2015?tsId=1555540 – accessed on 12th June, 2017.

http://nicd.ac.za/assets/files/NICD%20Communicable%20Diseases%20Communique_Mar2016_final.pdf) - accessed on 24th June, 2017.

https://www.hiv.lanl.gov/content/sequence/HIV/HIVTools.html – accessed on 24th June, 2017.

https://www.hiv.lanl.gov/content/sequence/HIV/HIVTools.html) – accessed on 26th June, 2017.

http://www.nature.com/nri/journal/v2/n4/fig_tab/nri776_F1.htm)- accessed on 27 May, 2017.

https://aidsinfo.nih.gov/education-materials/glossary/1596/life-cycle)– accessed in 27th June, 2017).

https://aidsinfo.nih.gov/understanding-hiv-aids/fact-sheets/21/58/fda-approved-hiv-medicines) – accessed on 17th August, 2017.

http://apps.who.int/iris/bitstream/10665/255896/1/9789241512831-eng.pdf?ua=1- accessed on 17th August. 2017.

http://data.worldbank.org/indicator/SH.DYN.AIDS.ZS?locations=CM) – accessed on 6th May, 2017.

https://hivdb.stanford.edu/hivdb/by-mutations/ – accessed on 1st July, 2017.

http://mafft.cbrc.jp/alignment/server/ - accessed on 15th April, 2017.

http://www.hiv.lanl.gov – accessed on 10th November, 2016.

https://comet.lih.lu/ - accessed on 6th October, 2016.

http://dbpartners.stanford.edu:8080/RegaSubtyping/stanford-hiv/typingtool/-accessed on 8th November, 2016.

http://www.datamonkey.org/dataupload_scueal.php - accessed on 3rd February, 2017.

http://jphmm.gobics.de - accessed on 12th February, 2017.

http://www.hiv.lanl.gov/content/sequence/RIP/RIP.html - accessed on 18th February, 2017.

http://resources.psmile.org/resources/equipment/specific-equipment/facscount/Equ2.3-02%20FACSCount%20System%20Overview.pdf/view)- accessed on 14th October, 2017.

http://www.bioafrica.net/proteomics/POLprot.html - accessed on 12th October, 2017.

http://www.hiv.lanl.gov/content/sequence/QC//index.html) – accessed 12th November, 2016.

http://www.bioafrica.mrc.ac.za/rega-genotype/html/subtypinghiv.html- accessed on 8th November, 2016

http://www.comet.retrovirology.lu) - accessed on 6th October, 2016.

http://hivdb.stanford.edu – accessed on 22nd March, 2017.

http://www.datamonkey.org/dataupload_scueal.php) - accessed on 3rd February, 2017

http://www.who.int/countries/cmr/en/) - accessed on 5th September, 2016.

http://www.hiv.lanl.gov) - accessed10th November, 2016.

hiv-aids@who.int) - accessed on 24th June, 2017.opet J., Bicart-See A., Pasquier C., Sandres K., Bonnet E., Marchou B., (1999). Mutations conferring resistance to zidovudine diminish the anti-viral effect of stavudine plus didanosine. *J. Med Virol*; 59: 507-511.

Ikomey G., Julius A., Jacobs G., Mesembe M., Eyoh A., Lyonga E., (2016). FasMediated(CD95L) Periferal T-cell Apotosis Marker in Monitoring HIV-1 Disease Progression in Adults in Yaoundé, Cameroon. *J. of Immunol*; 4(1): 1-5.

Ikomey GM., Assoumou O., Gichana JO., Njenda D., Mikasi SG., Mesembe M., Lyonga E., Jacobs G.B., (2017). Observed HIV drug resistance associated mutations amongst naïve immunocompetent children in Yaoundé, Cameroon; Ref: GERMS-OA-2017-0019.R1. Accepted, in press.

Jacobs GB, Laten AD, van Rensburg EJ, Bodem J, Weissbrich B, Rethwilm A,.et al., (2008). Phylogenetic Diversity and Low Level Antiretroviral Resistance Mutations in HIV Type 1 Treatment-Naive Patients from Cape Town, South Africa. *AIDS Research and Human Retroviruses*.24: (1009-1012).

Jacobs G., Loxton A., Laten A., Robson B., van Rensburg E., and Engelbrecht S., (2009). Emergence and diversity of different HIV-1 subtypes in South Africa, 2000-2001. *J. Med Virol*; 81, 1852-1859.

Jacks T., Power M., Masiarz F., Luciw P., Barr P., and Varmus H., (1988). Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *J. Nature*, 331, 280-3.

James, M. N. G. & Sielecki, (1985). Structural homology. A. R. Biochemistry 24, 3701−3713.

Kandel, C. E., & Walmsley, S. L. (2015). Dolutegravir – a review of the pharmacology, efficacy, and safety in the treatment of HIV. *Drug Design, Development and Therapy*, *9*, 3547–3555.

Kantor R., DeLong A., Balamane M., Schreier L., Lloyd Jr., Injera W., et al., (2014). HIV diversity and drug resistance from plasma and non-plasma analytes in a large treatment programme in western Kenya. *JIAS;* 17:19262.

Kantor R., Machekano R., Gonzales M., Dupnik K., Schapiro J., and Shafer R., (2001). Human immunodeficiency virus reverse transcriptase and protease sequence database: An expanded model integrating natural language text and sequence analysis. *J. Nucleic Acids Res*; 29: 296-299.

Keele B., Van Heuverswyn F., Li Y., Bailes E., Takehisa J., Santiago M., et al., (2006). Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *J. Science*; 313:523-526.

Kityo C., Sigaloff E., Boender S., Kaudha E., Kayiwa J., Musiime V., et al., (2016). HIV Drug Resistance Among Children Initiating First-Line Antiretroviral Treatment in Uganda. *AIDS Research and Human Retroviruses*, 32(7), 628–635.

Kiwanuka N., Laeyendecker O., Robb M.,Kigozi G., Arroyo M., McCutchan F., et al., (2008). Effect of human immunodeficiency virus Type 1 (HIV-1) subtype on disease progression in persons from Rakai, Uganda, with incident HIV-1 infection. J. Infect. Dis. 197:707-713.

Kiwanuka N., Laeyendecker O., Quinn T., Wawer M., Shepherd J., Robb M., et al., (2009). HIV-1 subtypes and differences in heterosexual HIV transmission among HIV-discordant couples in Rakai, Uganda. *J. AIDS*; 23:2479–2484.

Kiwanuka N., Robb M., Laeyendecker O., Kigozi G., Wabwire-Mangen F., Makumbi F., et al. *(*2010). HIV-1 viral subtype differences in the rate of CD4+ T-cell decline among HIV seroincident antiretroviral naive persons in Rakai district, Uganda. *J Acquir Immu Defic Syndr* ; 54:180-184.

Klimas N., Koneru A. and Fletcher M., (2008). Overview of HIV. *J. Psychosom Med*; 70, 523-530.

Koizumi Y., Ndembi N. and Miyashita M., (2006). Emergence of antiretroviral therapy resistance-associated primary mutations among drug-naïve HIV-1-infected individuals in rural western Cameroon. *J. Acquir Immune Defic Syndr*; 43(1):15–22.

Korber B., Gaschen B., Yusim K., Thakallapally R., Kesmir C., and Detours V., (2001). Evolutionary and immunological implications of contemporary HIV-1 variation. *J.Br. Med. Bull*; 58:19-42.

Koup R., Safrit J., Cao Y., Andrews C., McLeod G., Borkowsky W., et al., (1994). Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. *J. Virol*; 68:4650-4655.

Kumar P. (2013). Long term non-progressor (LTNP) HIV infection. *J. of Medical Research, India*; 138(3), 291–293.

Kuritzkes D., Shugarts D., Bakhtiari M., Poticha D., Johnson J., Thomas R., (2000). Emergence of dual resistance to zidovudine and lamivudine in HIV-1 infected patients treated with zidovudine plus lamivudine as initial therapy. *J. AIDS*; 23, 26–34.

Laurent C., Kouanfack C., Vergne L., Tardy M., Zekeng L., Noumsi, et al., (2006). Antiretroviral Drug Resistance and Routine Therapy, Cameroon. Emerging Infectious Diseases, 12(6), 1001–1004.

Larder B. and Kemp S. (1989). Multiple mutations in HIV-1 reverse transcriptase confer high-level resistance to zidovudine. *J. Science*; 246, 1155–8.

Lemey, P., Rambaut, A. and Pybus, O.G., (2006). HIV Evolutionary Dynamics Within and Among Hosts. *AIDS Reviews*, *8*, pp.125-40.

Lessells, R., Katzenstein, D., & de Oliveira, T. (2012). Are subtype differences important in HIV drug resistance? *Current Opinion in Virology*, *2*(5), 636–643.

Levy J., (2007). HIV and the Pathogenesis of AIDS. *3rd ed. Washington DC*: *ASM Press.*

Leoz M., Feyertag F., Kfutwah A., Mauclère P., Lachenal G., Damond F., et al., (2015). The two-phase emergence of non-pandemic HIV-1 Group O in Cameroon. PLoS Pathogens; 11(8).

Lieberman-Blum S., Fung H. and Bandres J., (2008). A CCR5-receptor antagonist for the treatment of HIV-1 infection. *Clin Ther*.; 30:1228-50. 10.1016/S0149-2918(08)80048-3.

Lindström A. and Albert J. (2003). A simple and sensitive 'in-house' method for determining genotypic drug resistance in HIV-1. *J. Virol Methods*; 107 (1):45-51.

Loxton A., (2004). Characterization of new full-length subtype D virus from South Africa.

Lv Z., Chu Y., and Wang Y., (2015). HIV protease inhibitors: a review of molecular selectivity and toxicity. *J. HIV/AIDS (Auckland, N.Z.);* 7, 95–104.

MacArthur R., (2009). Understanding HIV Phenotypic Resistance Testing: Usefulness in Managing Treatment-Experienced Patients. *J. AIDS Rev*.; 11, 223–230.

MacDonald K., Embree J., Nagelkerke N., Castillo J., Ramhadin S., Njenga S., et al., (2001). The HLA A2/6802 supertype is associated with reduced risk of perinatal human immunodeficiency virus type 1 transmission. *J. Infect Dis*.; 183: 503–506.

Maiga A., Malet I., Soulie C., Derache A., Koita V., Amellal B., et al., (2009). Genetic barriers for integrase inhibitor drug resistance in HIV type-1 B and CRF02_AG subtypes. *J. Antivir. Ther*.; 14, pp. 123-129.

Malim M., (2006). Natural resistance to HIV infection: The Vif-APOBEC interaction. *C R Biol.;* 329, 871-875.

Martinez-Cajas J., Pai N., Klein M. and Wainberg M., (2009). Differences in resistance mutations among HIV-1 nonsubtype B infections: a systematic review of evidence (1996–2008). *J. Int AIDS Soc*; 12:11.

Marx P, Maul D, Osborn K, et al., (1984). Simian AIDS: isolation of a type D retrovirus and transmission of the disease. *Science*; 223:1083-6.

Mazauric, M.-H., Leroy, J.-L., Visscher, K., Yoshizawa, S., & Fourmy, D. (2009). Footprinting analysis of BWYV pseudoknot–ribosome complexes. *RNA*, *15*(9), 1775–1786.

Mbanya D., Sama M., and Tchounwou, P. (2008). Current Status of HIV/AIDS in Cameroon: How Effective are Control Strategies? *Inter J. of Envir Res and Public Health;* 5(5), 378–383.

McColl D. and Chen X., (2010). Strand transfer inhibitors of HIV-1 integrase: Bringing IN a new era of antiretroviral therapy. *J. Antivir. Res*.; 85:101–118.

Menéndez-Arias L.(2013). Molecular basis of human immunodeficiency virus type 1 drug resistance: Overview and recent developments. *Antiviral Res*. 98, 93–120.

Miura T., Sakuragi J., Kawamura M., Fukasawa M., Moriyama E., Gojobori T., et al., (1990). Establishment of a phylogenetic survey system for AIDS-related lentiviruses and demonstration of a new HIV-2 subgroup. *J. AIDS*; 4(12), 1257-1261.

Murphy E., Collier A., Kalish L., Assmann S., Para M., et al. (2001). Highly active antiretroviral therapy decreases mortality and morbidity in patients with advanced HIV disease. *J. Ann Intern Med*; 338:853–60.

Mushahwar, I. K. (2007). Human Immunodeficiency Viruses: Molecular Virology, pathogenesis, diagnosis and treatment. *Perspectives in Medical Virology* 13:75-87.

Nakayama E., Miyoshi H., Nagai Y., & Shioda T. (2005). A Specific Region of 37 Amino Acid Residues in the SPRY (B30.2) Domain of African Green Monkey TRIM5α Determines Species-Specific Restriction of Simian Immunodeficiency Virus SIVmac Infection. *J. Virology*, *79*(14), 8870–8877.

NAM (2005), AZT and d4T: the thymidine analogues; (http://www.aidsmap.com/AZT-and-d4T-the-thymidine-analogues/page/1729914/) - accessed on 15/06/2017.

[a]Ndembi N., Lyagoba F., Nanteza B., Kushemererwa G., Serwanga J., Katongole-Mbidde E., et al., (2008). Transmitted Antiretroviral Drug Resistance Surveillance among Newly HIV Type 1-Diagnosed Women Attending an Antenatal Clinic in Entebbe, Uganda. *J. AIDS Res Hum Retro*; 24(6):889–95.

[b]Ndembi N., Abraha A., Pilch H., Hiroshi I., Mbanya D., Kaptue L., et al., (2008). Molecular characterization of human immunodeficiency virus type 1 (HIV-1) and HIV-2 in Yaounde, Cameroon: Evidence of major drug resistance mutations in newly diagnosed patients infected with subtypes other than subtype B. *J. Clin Microbiol*; 46(1):177–184.

Nkenfou N., Mekue M., Nana T., & Kuiate R. (2013). Distribution of *CCR5-Delta32*, *CCR5 promoter 59029 A/G*, *CCR2-64I* and *SDF1-3'A* genetic polymorphisms in HIV-1 infected and uninfected patients in the West Region of Cameroon. *BMC Research Notes*, *6*, 288.

Nouhin J., Donchai T., Hoang K., Ken S., Kamkorn J., Tran T., et al. (2011). Natural polymorphisms of HIV-1 CRF01_AE integrase coding region in ARV-naive individuals in Cambodia, Thailand and Vietnam: an ANRS AC12 working group study. *J. Infect. Genet. Evol;* 11, pp. 38-43.

Oelrichs RB, Shrestha IL, Anderson DA, Deacon NJ (2000). The explosive human immunodeficiency virus type 1 epidemic among injecting drug users of Kathmandu, Nepal, is caused by a subtype C virus of restricted genetic diversity. *J Virol* 74: 1149-1157.

Okome-Nkoumou M., Guiyedi V., Ondounda M., Efire, N., Clevenbergh, P., Dibo, M., (2014). Opportunistic Diseases in HIV-Infected Patients in Gabon following the Administration of Highly Active Antiretroviral Therapy: A Retrospective Study. The *American J. of Trop Med and Hyg*, 90(2), 211–215.

Oleske J., Muimefor A., Cooper R., Thomas K., dela Cruz A., Ahdieh H., (1983). Immune deficiency syndrome in children. *J. AMA*. 249, 2345- 2351.

Ombretta T., Claudia M., Francesca F., Mauro B., Gianluca R., Miriam L., et al., (2012). Short Communication: Analysis of the Integrase Gene from HIV Type 1-Positive Patients Living in a Rural Area of West Cameroon. AIDS Research and Human Retroviruses, 28(12): 1729-1733.

Oroszlan, S. & Luftig, R. B. (1990). Retroviral proteinases. Curr Top Microbiol Immunol 157, 153–185.

Onywera H., Maman D., Inzaule S., Auma E., Were K., Fredrick H., (2017). Surveillance of HIV-1 pol transmitted drug resistance in acutely and recently infected antiretroviral drug-naïve persons in rural western Kenya. PLoS ONE 12(2).

Palmisano L., and Vella S., (2011). A brief history of antiretroviral therapy of HIV infection. Success and challenges. Ann 1st Super Sanita; 47: 44-8.

Pandey, K. K. (2014). Critical appraisal of elvitegravir in the treatment of HIV-1/AIDS. *HIV/AIDS (Auckland, N.Z.)*, *6*, 81–90.

Pantaleo G., and Fauci A., (1995). New concepts in the immunopathogenesis of HIV infection. *Annu*. *Rev Immunol*.13:487-512.

Pape J., Liautaud B., Thomas F., Mathurin J., St Amand M., Boncy M., (1983). Characteristics of the acquired immunodeficiency syndrome (AIDS) in Haiti. *J. Med. N Engl;* 309, 945-950.

Pau A. and George J., (2014). Antiretroviral Therapy: Current Drugs. *J. Infect Dis Clinics of North America*; 28(3), 371–402.

Penazzato M., Dominguez K., Cotton M., Barlow-Mosha L., Ford N., (2015). Choice of antiretroviral drugs for postexposure prophylaxis for children: a systematic review. *J. Clin Infect Dis*; 60 Suppl 3: S177-81.

Pettit S., Everitt L., Choudhury S., Dunn B., Kaplan A., (2004). Initial Cleavage of the Human Immunodeficiency Virus Type 1 GagPol Precursor by Its Activated Protease Occurs by an Intramolecular Mechanism. *J. of Virol*; 78(16):8477-8485.

Pineda-Pena AC., Faria NR, Imbrechts S, Libin P, Abecasis AB, et al. (2013). Automated subtyping of HIV-1 genetic sequences for clinical and surveillance purposes: performance evaluation of the new REGA version 3 and seven other tools. Infect Genet Evol 19: 337–348.

Piot P, Quinn TC, Taelman H, Feinsod FM, Minlangu KB, Wobin O, et al., (1984). Acquired immunodeficiency syndrome in a heterosexual population in Zaire. *Lancet*.;2(8394):65–69.

Plantier J., Dachraoui R., Lemée V., Gueudin M., Borsa-Lebas F., Caron F., (2005). HIV-1 resistance genotyping on dried serum spots. *J. AIDS*; 4; 19 (4): 391-7.

Plantier J., Leoz M., Dickerson J., de Oliveira F., Cordonnier F., Lemee V., (2009). A new human immunodeficiency virus derived from gorillas. *J. Nat Med*; 15, 871-2.

Pommier Y., Johnson A., and Marchand C., (2005). Integrase inhibitors to treat HIV/AIDS. *J. Nat Rev Drug Discov*; 2005; 4:236–248.

Powell R; Barengolts D; Mayr L; Nyambi P. (2010). The Evolution of HIV-1 Diversity in Rural Cameroon and its Implications in Vaccine Design and Trials. *Viruses*; 2, 639–654.

Rhee S-Y, Jordan MR, Raizes E, Chua A, Parkin N, Kantor R, et al. (2015). HIV-1 Drug Resistance Mutations: Potential Applications for Point-of-Care Genotypic Resistance Testing. *PLoS ONE*10(12).

Roberts J., Bebenek K., Kunkel T., (1998). The accuracy of reverse transcriptase from HIV-1. *J. Science*; 242: 1171–1173.

Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, et al.,(2000). HIV-1 nomenclature proposal. *J. Science*; 288: 55–56.

Roederer M., Staal F., Raju P., Ela S., Herzenberg L., Herzenberg L., (1990). Cytokine-stimulated human immunodeficiency virus replication is inhibited by N-acetyl-L-cysteine. Proceedings of the National Academy of Sciences of the United States of America, 87(12), 4884–4888.

Ross L., Lim M., Liao Q, Wine B., Rodriguez E., Weinberg W., Shaefer M., (2007). Prevalence of antiretroviral drug resistance and resistance-associated mutations in antiretroviral therapy-naive HIV infected individuals from 40 United States cities. *HIV Clin Trials:* 8:1–8.

Salemi M, Strimmer K, Hall WW, Duffy M, Delaporte E, Mboup S, et al., (2001). Dating the common ancestor of SIVcpz and HIV-1 group M and the origin of HIV-1 subtypes using a new method to uncover clock-like molecular evolution. *FASEB J.* 2001; 15 (2):276–8.

Sama C., Feteh V., Tindong M., Tanyi J., Bihle N., Angwafo F., (2017). Prevalence of maternal HIV infection and knowledge on mother–to–child transmission of HIV and its prevention among antenatal care attendees in a rural area in northwest Cameroon. PLoS ONE 12(2):

Sambrook J. and Rusell D., (2001). Molecular Cloning; *a laboratory manual, Cold Spring Harbor Laboratory Press, New York, USA, 3rd edition,* Vol 1 pp 5.2-5.13 & Vol 2, p 8.4.

Sambrook J. and Michael R. Green (2012). Molecular Cloning: A Laboratory Manual (Fourth Edition). Cold Spring Harbor Laboratory Press. pp. 2,028. l.

Sánchez PR. & Holguín A., (2014). Drug resistance in the HIV-1-infected paediatric population worldwide: a systematic review, J. of Antimicrobial Chemother., Vol. 69, Issue 8, 1 August 2014, Pages 2032–2042.

Sarafianos S., Marchand B., Das K., Himmel D., Parniak A., Hughes, et al., (2009). Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. *Journal of Molecular Biology*, *385*(3), 693–713.

Sanger F., Nicklen S. and Coulson A., (1977). DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA;* 74, 5463-5467.

Schultz A., Zhang M., Bulla I., Leitner T., Korber B., Morgenstern B., Stanke M., (2009). jpHMM: improving the reliability of recombination prediction in HIV-1. *Nucleic Acids Res*.; 37 (Web Server issue): W647-51.

Schacker T., Collier A., Hughes J., Shea T., Corey L., (1996). Clinical and epidemiologic features of primary HIV infection. *J. Ann Intern Med*; 125:257-264.

Seth C., Hamers R., Kityo C., Rinke de Wit T.,Roura M., (2016). Long-Term Antiretroviral Treatment Adherence in HIV-Infected Adolescents and Adults in Uganda: A Qualitative Study. *PLoS ONE*, 11 (11).

Sharp P. and Hahn B., (2011). Origins of HIV and the AIDS pandemic. *Cold Spring Harb Perspect Med*; 1 (1), a006841.

Shafer R., Rhee S., Pillay D., Miller V., Sandstrom P., Schapiro J., (2007). HIV-1 protease and reverse transcriptase mutations for drug resistance surveillance. *J. AIDS;* 21, 215-223.

Shafer R., (2002). Genotypic Testing for Human Immunodeficiency Virus Type 1 Drug Resistance. *Clinical Microbiology Reviews*; 15 (2):247-277.

Shafer R., (2006). Rationale and uses of a public HIV drug-resistance database. *J Infect Dis;* 194 Suppl 1: S51-58.

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

Sheth A., Ofotokun I., Buchacz K., Armon C., Chmiel J., Hart R., et al., (2016). Antiretroviral regimen durability and success in treatment-naïve and treatment-experienced patients by year of treatment initiation, United States, 1996–2011. *J. AIDS*; (1999), 71(1), 47–56.

Siliciano J., Kajdas J., Finzi D., Quinn T., Chadwick K., Margolick J., (2003). Long-term follow-up studies confirm the stability of the latentreservoir for HIV-1 in resting CD4+ T cells. J*. Nat Med*; 9: 727–8.

Spira S., Wainberg M., Loemba H., Turner D., Brenner B., (2003). Impact of clade diversity on HIV-1 virulence, antiretroviral drug sensitivity and drug resistance. *J. Antimicrob Chemother;* 51: 229–240.

Steigbigel RT., Cooper DA., Kumar PN., Eron JE., Schechter M., Markowitz M, et al., (2008). Raltegravir with optimized background therapy for resistant HIV-1 infection. *N Engl J Med*. 359:339–354.10.

Stevenson M. (2003). HIV-1 pathogenesis. *Nat. Med*. 9, 853–860.

Stremlau M., Owens M., Perron J., Kiessling M., Autissier P., Sodroski J, (2004). The cytoplasmic body component TRIM5α restricts HIV-1 infection in old world monkeys. *Nature*. 427:848–853.

Struck D., Lawyer G., Ternes A., Schmit J., and Bercoff D., (2014). COMET: adaptive context-based modeling for ultrafast HIV-1 subtype identification. *Nucleic Acids Research;* 42(18).

Takow S., Atashili J., Enow-Tanjong R., Mesembe M., Ikomey G., Ndip L., (2015). Time for Option B+? Prevalence and characteristics of HIV infection among

attendees of 2 antenatal clinics in Buea, Cameroon. *J. IntAssocProvid AIDS Care*; 2015. 14 (1): 77 - 81.

Tamura K., Peterson D., Peterson N., Stecher G., Nei M., & Kumar S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution, 28(10), 2731–2739.

Taylor, B. S., Sobieszczyk, M. E., McCutchan, F. E., & Hammer, S. M. (2008). The Challenge of HIV-1 Subtype Diversity. *The New England Journal of Medicine*, *358*(15), 1590–1602.

Tebit DM, Nankya I, Arts EJ, Gao Y. (2007). HIV diversity, recombination and disease progression: how does fitness "fit" into the puzzle? *AIDS Reviews* 9: 75-87.

Tongo M. and Burgers W., (2014). Challenges in the design of a T cell vaccine in the context of HIV-1 diversity. *J. Viruses*; 6(10):3968-90.

UNAIDS, 2017- http://www.unaids.org/sites/default/files/media_asset/20170720_Data_book_2017_en.pdf) – accessed on 9th October 2017.

UNAIDS, 2017: HIV and AIDS estimates, 2016 - http://www.unaids.org/en/regionscountries/countries/cameroon/) - accessed on 9th October 2017.

UNAIDS, epidemic update December 2015 - www.unaids.org – accessed on 15th April 2016.

UNAIDS/WHO, December 2015 – (www.unaids.org) – accessed on 4th May, 2016.

UNAIDS, AIDS epidemic update December 2015. (www.unaids.org) – accessed on 15th April 2016.

UNAIDS/WHO, (December 2015). Working group on global HIV/AIDS & STD surveillance. (2013 / 2015). AIDS epidemic update. Geneva, Switzerland, (www.unaids.org) – accessed on 4th May, 2016.

UNAIDS - *Global AIDS Update 2016* – available online at*:* (http://www.unaids.org/sites/default/files/media_asset/global-AIDS-update-2016_en.pdf) - accessed on 18th June 2016.

UNAIDS, 2017: Fact Sheet, 2017- Available online at: http://www.unaids.org/sites/default/files/media_asset/UNAIDS_FactSheet_en.pdf – accessed on 20th October 2017.

Vallari A., Holzmayer V., Harris B., Yamaguchi J., Ngansop C., Makamche, F., et al., (2011). Confirmation of Putative HIV-1 Group P in Cameroon. *J. of Virology*; 85(3), 1403–1407.

Van L., Van K., Schmit J., Sprecher S., Hermans P., De Vroey V., et al., (1999). Phenotypic assays and sequencing are less sensitive than point mutation assays for detection of resistance in mixed HIV-1 genotypic populations. *J. Acquir Immune Defic Syndr*; 22(2):107-18.

Van Heuverswyn F, Li Y, Bailes E, Neel C, Lafay B, Keele BF, et al., (2007). Genetic diversity and phylogeographic clustering of SIVcpzPtt in wild chimpanzees in Cameroon. *J. Virology*. 2007;368 (1):155–171.

[a]Vessiere A., Leoz M., Brodard V., Strady C., Lemée V., Depatureaux A., et al., (2010). First evidence of a HIV-1 M/O recombinant form circulating outside Cameroon. *J. AIDS*; 24 (7), 1079–1082.

[b]Vessière A., Rousset D., Kfutwah A., Leoz M., Depatureaux A., Simon F, et al., (2010). Diagnosis and Monitoring of HIV-1 Group O-Infected Patients in Cameroun. *J. AIDS*; Vol.53 (1):pp 107-110.

Villabona-Arenas C., Domyeum J., Mouacha F., Butel C., Delaporte E., Peeters., et al., (2015). HIV-1 group O infection in Cameroon from 2006 to 2013: Prevalence, genetic diversity, evolution and public health challenges. Infection, Genetics and Evolution : *J. of Mol Epide and Evol Gene in Infect Dis;* 36, 210–216.

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

Wang C., Mitsuya Y, Gharizadeh B., Ronaghi M., and Shafer R., (2007). Characterization of mutation spectra with ultra-deep pyrosequencing: application to HIV-1 drug resistance. *J. Genome Res;* 17: 1195-1201.

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

Wei P, Garber ME, Fang SM, Fischer WH, Jones KA, (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(4):451–62.

Weiss R., (1993). "How does HIV cause AIDS?". *J. Science*; 260 (5112): 1273–9. PMID 8493571.

Wensing A., van Maarseveen N. and Nijhuis M., (2010). Fifteen years of HIV Protease Inhibitors: Raising the barrier to resistance. *J. Antivir. Res*., 85, 59–74.

Wensing A., Calvez V., Gunthard H., Johnson V., Paredes R., Pillay D., et al. 2017 update of the drug resistance mutations in HIV-1. Top Antivir Med. 2017;24(4):132–141.

Westby M., Smith-Burchnell C., Mori J., Lewis M, Mosley M, Stockdale M, et al., (2007). Reduced maximal inhibition in phenotypic susceptibility assays indicates that viral strains resistant to the CCR5 antagonist maraviroc utilize inhibitor-bound receptor for entry. *J. Virol*; 81:2359–71.

Worobey M., Gemmel M., Teuwen D. E., Haselkorn T., Kunstman K., et al., (2008). Direct Evidence of Extensive Diversity of HIV-1 in Kinshasa by 1960. *J. Nature*; *455*(7213), 661–664.

World Bank- (*http://data.worldbank.org/indicator/SH.HIV.ARTC.ZS?locations=CM*) - accessed on 6th May 2017.

World Development Indicators (WDI). (2015). (http://knoema.com/WBWDIGDF2015Aug/world-development-indicators-wdi-september-2015?tsId=1555540 – accessed on 12th June, 2017.

WHO, 2017: http://www.who.int/mediacentre/factsheets/fs360/en/) – accessed on 9th October, 2017

World Bank - https://data.worldbank.org/indicator/SH.HIV.ARTC.ZS?locations=CM - accessed on 12th October 2017].

www.unaids.org) – accessed on 4th May, 2016

World Health Organization (WHO, 2015). Country report. (http://www.who.int/countries/cmr/en/) - accessed on 5th September 2016.

WHO, 2017. HIV drug resistance report, 2017. Available at: http://apps.who.int/iris/bitstream/10665/255896/1/9789241512831-eng.pdf?ua=1- accessed on 17th August 2017.

WHO, 2017. Infant, Newborn: http://www.who.int/topics/infant_newborn/en/- accessed on 15th October 2017.

Wyatt, R. and Sodroski, J. (1998). The HIV-1 envelope glycoproteins: Fusogens, antigens, and immunogens. *Science*, 280(5371), 1884-1888.

Yamaguchi J., Bodelle P., Vallari A., Coffey R., Mcrthur C., Schochetman G., et al., (2004). HIV Infections in North-western Cameroon: Identification of HIV Type 1 Group O and Dual HIV Type 1 Group M and Group O Infections. *J. AIDS Res and Hum Retroviruses;* 20(9):944-57.

Yasutsugu S. and Youichi S., (2011). Gene Regulatable Lentiviral Vector System, in Dr. Ke Xu (Ed.), Viral Gene Therapy, *InTech*; pp.1-25.

Yılmaz M., Ozic C., and Gok I., (2012). Principles of Nucleic Acid Separation by Agarose Gel Electrophoresis, Gel Electrophoresis - Principles and Basics, Dr. Sameh Magdeldin (Ed.), InTech, DOI: 10.5772/38654. Available from: https://www.intechopen.com/books/gel-electrophoresis-principles-and-basics/principles-of-nucleic-acid-separation-by-agarose-gel-electrophoresis.

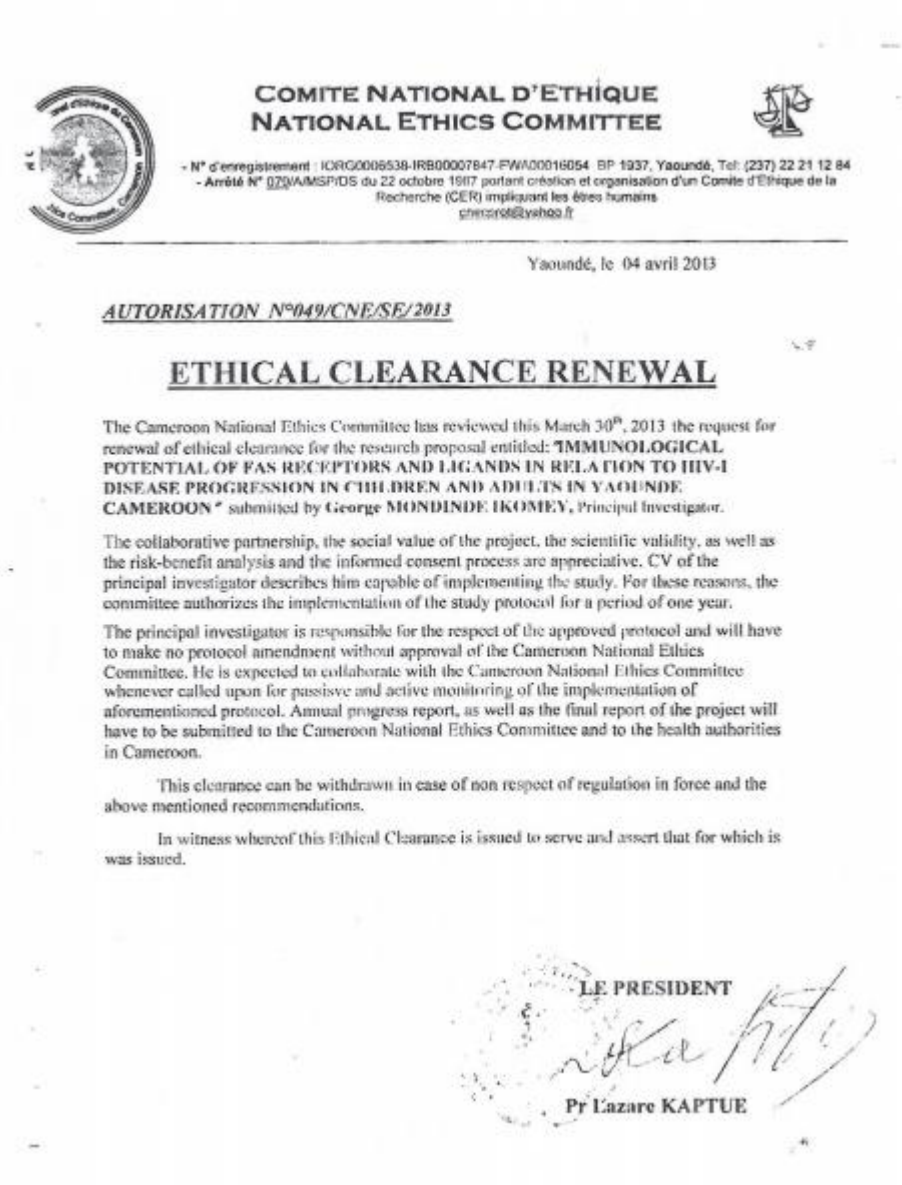
Yasukawa K, Nemoto D. and Inouye K. (2008). Comparison of the thermal stabilities of reverse transcriptases from avian myeloblastosis virus and Moloney murine leukaemia virus. J. Biochem.; 143:261–268.

You J, Wang H, Huang X, Qin Z, Deng Z, Luo J, et al. (2016). Therapy-Emergent Drug Resistance to Integrase Strand Transfer Inhibitors in HIV-1 Patients: A Subgroup Meta-Analysis of Clinical Trials. PLoS ONE; 11(8).

Zheng Y., Lovsin N., Peterlin B., (2005). "Newly identified host factors modulate HIV replication*". Immunol. Lett*. 97 (2): 225–34.

Zhu T., Korber B., Nahmias A., Hooper E., Sharp P., Ho D., (1998). An African HIV-1 sequence from 1959 and implications for the origin of the epidemic. *J. Nature*; 391:594-597.

Appendix I



COMITE NATIONAL D'ETHIQUE
NATIONAL ETHICS COMMITTEE

- N° d'enregistrement : IORG0006538-IRB00007847-FWA00016054 BP 1937, Yaoundé, Tel: (237) 22 21 12 84
- Arrêté N° 079/A/MSP/DS du 22 octobre 1997 portant création et organisation d'un Comité d'Ethique de la Recherche (CER) impliquant les êtres humains
cnerzoa@yahoo.fr

Yaoundé, le 04 avril 2013

AUTORISATION N°049/CNE/SE/2013

ETHICAL CLEARANCE RENEWAL

The Cameroon National Ethics Committee has reviewed this March 30th, 2013 the request for renewal of ethical clearance for the research proposal entitled: "**IMMUNOLOGICAL POTENTIAL OF FAS RECEPTORS AND LIGANDS IN RELATION TO HIV-1 DISEASE PROGRESSION IN CHILDREN AND ADULTS IN YAOUNDE CAMEROON**" submitted by **George MONDINDE IKOMEY**, Principal Investigator.

The collaborative partnership, the social value of the project, the scientific validity, as well as the risk-benefit analysis and the informed consent process are appreciative. CV of the principal investigator describes him capable of implementing the study. For these reasons, the committee authorizes the implementation of the study protocol for a period of one year.

The principal investigator is responsible for the respect of the approved protocol and will have to make no protocol amendment without approval of the Cameroon National Ethics Committee. He is expected to collaborate with the Cameroon National Ethics Committee whenever called upon for passisve and active monitoring of the implementation of aforementioned protocol. Annual progress report, as well as the final report of the project will have to be submitted to the Cameroon National Ethics Committee and to the health authorities in Cameroon.

This clearance can be withdrawn in case of non respect of regulation in force and the above mentioned recommendations.

In witness whereof this Ethical Clearance is issued to serve and assert that for which is was issued.

LE PRESIDENT

Pr Lazare KAPTUE

NATIONAL ETHICS COMMITTE ON HUMAN HEALTH RESEARCH (NECHHR)

SECRETARIAT

Tel : 22 76 21 14 E-mail : cnethique_minsante@yahoo.fr

- Registration N° : IORG0006538-IRB00007847-FWA00016054
- Degree N° 0977/A/MSP/SESP/SG/DROS of 18 April 2012 on the creation, organization and functioning of the National Ethics Committee on Human Health Research within the structure of the Ministry in Charge of Public Health

Administratives procedure for issuing an ethical clearance

Title of protocol ----------

Submission Date_____ Name of Principal Investigator (PI) _____

Tel: _____ E-mail _____ Institution _____

Application file Number: _____

Number of

REQUIRED DOCUMENTS

N°	DOCUMENTS	YES	NA*
1	An Application Addressed to the President of the NECHHR		
2	04 copies of protocols signed by the PI and the Co-Investigators (or 02 copies of protocols for students) with a soft copy sent to the email address of the secretariat of the NECHHR. The Protocol should be written in French or English		
3	A summary of the protocol		
4	A paragraph on ethical consideration (methodology)		
5	Information sheet and consent form in French and English		
6	Tools for data collection (questionnaires, Discussion and/or Interview guides, CRFs ...)		
7	Detailled Budget of the project		
8	Updated CVs of the PI and all the Co-Investigators		
9	An autorisation letter from the health institution where the research work would be carried out		
10	The Source of funding and the address		
11	Receipt of payment of processing fee (The payment should be made to the account **N° 10002 00001 97662023130 57 located within the « Centre des Affaires PME/PMI de la direction SCB Cameroun » situated opposite « BICEC centrale » or Saint-Paul Library, Yaoundé**		
	For Clinical Trials:		
12	Investigator's Brochure		
13	Insurance cover for any harm incured by the participant		
14	Insurance cover for errors which could occur during the implimentation of the protocol		
15	Agreement to receive treatment at the end of the trial		
	Studies for which the Promoter is abroad:		
16	Ethical Clearence from a competent ethical committee		
17	An Agreement to transfer biological data, Material Transfer Agreement (MTA)		
18	An agreement to share the data generated from the study, Data Sharing Agreement (DSA)		

* NA=Not Applicable

NB: Evaluation meetings for application files take place once every two months, the fourth Thursday of the month. The research protocol should be deposited at the secretatriat of the National Ethic Committee at least 10 days before the meeting for it to be considered. !

Signature of the Secretary:

A copy is given back to the Principal Investigator or his representative and another archived

Fees for National Research protocol: One Hundred thousand Francs CFA (100 000)

Fees for International Research Protocol: Two hundred thousand Francs CFA (200 000)

Appendix II

UNIVERSITEIT•STELLENBOSCH•UNIVERSITY
jou kennisvennoot • your knowledge partner

Ethics Letter

05-May-2017
Jacobs, Graeme GB

Ethics Reference #: N14/10/130
Title: The investigation of the molecular diversity and resistance patterns of HIV in different cohorts in Cameroon.

Dear Dr. Graeme Jacobs

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

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

Approval date: 17 March 2017

Expiry date: 16 March 2018

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

Please remember to use your **protocol number (N14/10/130)** on any documents or correspondence with the HREC concerning your research protocol.

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

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

Sincerely,
Franklin Weber
REC Coordinator
Health Research Ethics Committee 1

Appendix III

GERMS

Observed HIV drug resistance associated mutations amongst naïve immunocompetent children in Yaoundé, Cameroon

Journal:	*GERMS*
Manuscript ID	Draft
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	IKOMEY, GEORGE; Microbiology,Parasitology,Virology and Infectious Diseases Okomo Assoumou, Marie Claire; Center for the Study and Control of Communicable Diseases (CSCCD) Faculty of Medicine and Biomedical Sciences, University of Yaoundé 1, Microbiology, Parasitology, Infectiuos Diseases Gichana, Josiah; Division of Medical Virology University of stellenbosh, Pathology; University of Nairobi, College of Health Sciences, Oral surgery Njenda, Duncan; Division of Medical Virology University of stellenbosh, Pathology Mikasi, Sello; Division of Medical Virology University of stellenbosh, Pathology Mesembe, Martha; Microbiology,Parasitology,Virology and Infectious Diseases mbamwi, Emilia; Center for the Study and Control of Communicable Diseases (CSCCD) Faculty of Medicine and Biomedical Sciences, University of Yaoundé 1, Microbiology,Parasitology and Infectious Diseases Jacobs, Graeme; Division of Medical Virology University of stellenbosh, Pathology
Keywords:	HIV, Infants, Cameroon, Mothet-to-child-transmission, treatment-naïve, immune competent
Abstract:	Abstract Background Emerging drug resistance mutations (DRMs) has been a major threat for successful lifelong combination antiretroviral therapy (cART), especially for HIV-vertically infected children within the context of the prevention of Mother-to-Child transmission (PMTCT). This study aimed at evaluating DRMs amongst immune competent treatment-naïve children in Cameroon. Methods A cross-sectional study was conducted between 2015 and 2016 amongst 55 proxy consented HIV-1 positive prenatal children, aged 9 - 72 months old. They were all immune competent, cART naïve and with unknown PMTCT exposure. CD4 cell counts and genotypic drug resistant testing were

doneusing standard methods.

Results
Levels of DRMs to Protease (PR) inhibitors (PIs), nucleoside reverse transcriptase inhibitors (NRTIs) and non-NRTIs were 27.6%, 3.7% and 40.7%, respectively. Only minor DRMS were observed for PR.The observed mutations for NRTI were K65R, T215I and K219E (33.0% each) and for NNRTI; V106M, Y181C and Y188H (6.0% each). Only minor accessory mutations were found in the Integrase (IN) region.

Conclusion
Despite widely available cART we still observe naïve HIV children, especially from the rural communities. We observe that a proportion of study participants had HIV-1 drug resistance associated mutations (RAMs). Data generated could help strengthen the current PMTCT programmes within the country. There is a need to upscale approaches for drug resistance testing for children in Cameroon and many other resource-limited settings.