ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

COLLEGE OF HEALTH SCIENCES

FACULTY OF MEDICINE

GENOMIC CHARACTERIZATION OF HIV-1 ISOLATES FROM ETHIOPIAN PATIENTS: BASELINE STUDIES ON ANTIRETROVIRAL DRUG RESISTANCE AND SUB-TYPE VARIATIONS

BY

WOLDAREGAY ERKU ABEGAZ

MAY 2011

ADDIS ABABA

ADDIS ABABA UNIVERSITY

School of Graduate Studies

College of Health Sciences

Faculty of Medicine

Genomic Characterization of HIV-1 Isolates from Ethiopian Patients: Baseline Studies on Antiretroviral Drug Resistance and Sub-type Variations

By

Woldaregay Erku Abegaz

A Thesis Submitted to the school of Graduate Studies of the Addis Ababa University in Partial Fulfillment of the Requirement for the Degree of Doctor of Philosophy in Medical Microbiology

May 2011

Addis Ababa

ADDIS ABABAUNVERSITY

SCHOOL OF GRADUATE STUDIES

Genomic Characterization of HIV-1 Isolates from Ethiopian Patients: Baseline Studies on Antiretroviral Drug Resistance and Sub-type Variations

By

Woldaregay Erku Abegaz

A Thesis Submitted to the School of Graduate Studies of the Addis Ababa University in Partial Fulfillment of the Requirement for the Degree of Doctor of Philosophy in Medical Microbiology

Approved by the Examining Board:

Dr. Yohannes Mengistu (Supervisor) _____

_____(External Examiner) _____

_____(Internal Examiner) _____

_____(Chairperson) _____

Acknowledgments

This work was the result of many people's participation, however small the contribution might be. First, I would like to thank those study participants who were very helpful with no hesitation during specimen collection. Secondly, this study would have never come to an end without unfailing efforts from my dearest advisor, Dr Yohannes Mengistu, who tirelessly sought for and finally succeeded in securing laboratory supplies, kits and reagents which were the only ways to finish the work. He was responsible to successfully establish external links in relation to this study so that a portion of specimens were analyzed at foreign laboratories during the early period of the study. In the final phase of the thesis write-up, his amazing speed of reviewing the draft thesis amidst the tight schedule of his job, and the valuable comments he provided helped me to finalize and reshape the write-up in the current form. Finally, I would like to acknowledge that while the strengths were contributed by Dr. Yohannes as a supervisor, the weaknesses that might be seen in this thesis should be my own responsibilities.

The contributions of Dr. Jonathan Kaplan from the CDC Atlanta in the early phase and Dr. Almaz Abebe from EHNRI in late phase of the work were very crucial for materializing this study, without whose involvement it would have been impossible to start and finish the project at all. I am also grateful to Dr. Peter Fonjungo from the CDC-Ethiopia, Professor Damen H/Mariam and Dr. Solomon G/Selassie from the Faculty of Medicine, and Professor Getachew Tilahun (former director of ALIPB) from ALIPB, Dr. Mirutse Giday from ALIPB (current Director), and Ato Tamirat Bekele from International Clinical Laboratories (General Manager), for these people were the persons who brought additional positive differences to the momentum of the research work when it was in a state of stale. The help from Mrs. Dorothy Fitzmaurice

from Emory University in linking me with Dr. Kaplan of the CDC was decisive to make the first progress in this study. Also deserving appreciation is Dr. Carlos del Rio from Emory AIDS International Training and Research Program (AITRP) for giving me the first opportunity to visit laboratories working on HIV research at Atlanta. Technical assistances from Ato Dawit Asseffa from EHNRI (then MSc. student), W/ro Hiwot Berhanu of former EHNRI staff(then MSc. student), Ato Tesfaye Tilahun from EHNRI, Drs. Zehava Grossman, Daniela Ram, and Professor Ella Mendelson of Central Virology Laboratory at Tel Hasomer (Israel) and all the technicians working there, and technical staff of CDC-Atlanta deserve special thanks. I would also thank Dr. Dawit Wolday for his attempts to link the work with other foreign laboratory in Canada. Special thanks should go to the physicians and laboratory workers at the collaborating health providing sites where sample were collected and processed: Dr. Aster Shewamere, Ato Dnaiel Dejene, and Ato Girmame Wodimneh from Zewditu Hospital, Dr. Daniel Fekadu from Tikur Anbessa Hospital, the physicians at St. Paul Hospital, and all the laboratory technicians at the seven ANC sites selected in this study. I immensely thank Ao Terefe Tsegaye for the technical assistances I received in viral load determination at the International Clinical Laboratories (ICL).

I am indebted to my friends and colleagues at ALIPB who were always beside me to support and boost up my morale whenever I faced frustrations, which happened several times throughout the duration of this lengthy study. Very special thanks to W/ro Zufan Sisay, Dr. Mshesha Balkew, Dr. Mulugeta Bellay, Ato Girmay Medhin, and all other staff members of ALIPB for keeping me motivated and concentrated. Other friends and relatives who had been providing me with encouragements and who had been solving my private problems, so that I should focus only on the study work, deserve due credit. Especially my heartfelt appreciation goes to my friends Ato Musa Mohammod, W/ro Helen Fetene, Dr. Berhanu Bogale and W/ro Meseret Teshome, who in

one way or another saved me in several occasions from problems that otherwise would have distracted my attention from the work that matters.

I am fully aware that the sacrifies that my family paid for this study is priceless. I thank my wife, W/ro Atsede Bezabih, for taking care of our kids especially during my intermittent absence from home in the final year of the thesis write-up. Of course, nobody was victimized by the protracted study period as were my two beloved kids, Biruk and Ephrata Woldaregay, for they were deprived of the attentions, care and nurturing they are naturally entitled to receive from a father. I ask of their apology, and promise to do my outmost ability to make up my previous long absence. I thank them deeply for they have been the foundation for my endurance during the long painful period of keeping the study alive.

Finally, I am grateful to the following institutions for financial, technical, logistic and other assistances they provided me with: Faculty of Medicine and Department of Microbiology, Immunology & Parasitology (AAU); Ethiopian Health and Nutrition Research Institute (EHNRI); Aklilu Lemma Institute of Pathobiology (ALIPB); School of Graduate studies, Addis Ababa University; International Clinical Laboratories (ICL); Zewditu Memorial Hospital; Tikur Anbessa General Teaching Hospital; St Paul Teaching Hospital; Center for Disease Control and Prevention (CDC), Atlanta; CDC-Ethiopia; Central Virology Laboratory at Tel Hasomer (Israel); and Emory AIDS International Training and Research Program (AITRP).

Table of Contents

List of Tables

List of Figures

List of Acronyms

3TC	lamivudine
AB	antibody
ABC	abacavir
AIDS	acquired immunodeficiency syndrome
ANC	antenatal clinic
ART	antiretroviral therapy
ARV	antiretroviral
ATV	atazanavir
AZT	zidovudine (also known as ZDV)
CD4+ cell	T-lymphocyte bearing CD4 receptor
CDC	Centers for Disease Control and Prevention
CNS	central nervous system
D4T	stavudine
DBS	dried blood spot
ddI	didanosine
DNA	deoxyribonucleic acid
DRV	darunavir
EFV	efavirenz
EHNRI	Ethiopian Health and Nutrition Research Institute
EIA	enzyme immunoassay
ETV	etravirine
FDC	fixed-dose combination
FSW	Female Sex Workers
FTC	emtricitabine
HAPCO	HIV/AIDS Prevention & Control Office
HAART	Highly Active Antiretroviral Therapy
HIV	human immunodeficiency virus
HIVDR	HIV drug resistance
HIVRNA	human immunodeficiency virus ribonucleic acid
HIVDR-TS	HIV drug resistance threshold survey
IDU	injecting drug user
IDV	indinavir
INI	integrase inhibitor
LPV	lopinavir
LPV/r	lopinavir/ritonavir
MOH	Ministry of Health
FMOH	Federal Ministry of Health
MTCT	mother-to-child transmission (of HIV)
NFV	nelfinavir
NGO	Non-Governmental Organization
NNRTI	non-nucleoside reverse transcriptase inhibitor
NRTI	nucleos(t)ide reverse transcriptase inhibitor
NSS	National Sentinel Surveillance
NVP	nevirapine

OI	opportunistic infection
PCP	*Pneumocystis jiroveci* pneumonia
PEPFAR	President's Emergency Plan for AIDS Relief
PI	protease inhibitor
PLHIV	people living with HIV
PMTCT	prevention of mother-to-child transmission (of HIV)
PR	protease
RAL	raltegravir
RBV	ribavirin
RNA	ribonucleic acid
RPR	Rapid Plasma Reagin
RT	reverse transcriptase
RTI	reverse transcriptase inhibitor
RTV	ritonavir
SQV	saquinavir
TDF	tenofovir disoproxil fumarate
VL	viral load
VCT	Voluntary Counselling and Testing
UNAIDS	Joint United Nations Program on HIV/AIDS
WHO	World Health Organization

Abstract

Background

Several studies were conducted in the past two decades in Ethiopia to understand the genotypic characteristics of prevalent HIV subtypes in the country. However, the majority of the information gathered relied on sequencing of the envelope and a certain portion of the gag gene fragments of the HIV genome. Considering the facts that relying only on sequencing of the env and gag regions has inherent inadequacy in characterizing HIV subtypes, that long time has elapsed since genomic characterization of HIV isolates in Ethiopia, that non-C subtypes have expanded in the neighboring countries, and that a new selective pressure coming from the recently initiated ARV treatment has been introduce into current HIV isolates circulating in the country as a consequence of which newHIV subtypes might be currently circulating in the country, this PhD study was, therefore, undertaken to understand the genomic characteristics of HIV isolates from ARV drug-naïve and drug-experienced persons by sequencing the protease (PR) and reverse transcriptase (RT) genes using standard RT-PCR/PCR amplification and genome sequencing protocols. In addition, two test evaluation studies were carried out: one on use of In-house brewed genotyping system, and the other on use of DBS as source of specimen for genotypic HIV drug resistance monitoring.

Materials and Methods

Different set of criteria were used to select eligible HIV-infected people to take part in three study groups: recently infected drug-naïve antenatal care (ANC) attendee pregnant women (recent drug-naïve), chronically infected drug-naïve but who were eligible to start ART (chronic drug-naïve), and chronically infected heavily treated patients (chronic drug-experienced). For drug resistance assays, list of mutations from three algorithms were utilized. For evaluations of the performance of In-house genotyping system and DBS as source sample, the commercial genotyping system ViroSeq™ and plasma samples, respectively, were used as standards. Mean nucleotide similarities of paired sequences generated from In-house/ViroSeq™ genotyping system and DBS/plasma were determined on pairwise alignment sequences using ClustalW multiple alignment program in BioEdit Version 7.0.9.0 software. For other purposes, FASTA-formatted nucleotide and amino acid sequences were aligned and analyzed using various on-line and off-line software. Descriptive, χ^2 test, and Student t test statistical data analyses were used as appropriate, with significant level of 0.05.

Results

Among chronically infected participant groups, 74% of drug-naïve and 84% of drug-experienced participants were in the advanced WHO clinical stages III and IV at baseline. Nearly 92% of chronic drug-naïve patients had CD4+ count of $\leq$200 cells/µL, and 70% had over 100,000 copies/ml of plasma viral RNA. Mean baseline CD4+ T cell count for chronic drug-experienced patients was ~114 cells/µL. Moreover, mean plasma viral RNA and mean CD4+ T cell counts for these patients at the time of genotyping were 289,128 copies/ml and 238 cells/µL, respectively, after median treatment duration of 42 months. The performance of the In–house genotyping system relative to the performance of the commercial ViroSeq™ genotyping system both in amplifying and genotyping the specimens in this study was 100%. It had also mean concordance value of 98.72% at the nucleotide level. Similarly, the performance of DBS as specimen for drug resistance genotyping relative to the performance of plasma specimens was 100% for both amplification and genotyping. The mean nucleotide concordance rate between paired DBS/plasma sequences was 98.82%.

The prevalence of transmitted drug resistance mutations was 4.9% among recent drug-naïve persons. Among chronically infected drug-naïve patients, 22 major resistance mutations were detected: eight against PIs, eight against NRTIs, and 6 against NNRTIs. Virological treatment failure among treatment-experienced patients was 31%; 70% of successfully genotyped viremic specimens harbored at least one confirmed major drug resistance mutation from any class. Drug class-wide mutations were 3% against PIs, 62% against NRTIs, 68% against NNRTIs, 61% against any NRTI and NNRTI double class

mutations, and 3% against all the three classes of drugs in use. Nearly 37% of D4T- and 20% of AZT-containing initial regimens were responsible for selection of at least one NRTI resistance mutations. Of the initial drug regimens, ~27% of treated patients with NVP-containing regimens or ~55% of NVP-exposed patients with genotyped viral isolates harbored at least one major NNRTI resistance mutation. Likewise, ~29% of all EFV-treated patients or 96% of EFV-exposed patients with genotyped viral isolates had at least one NNRTI resistance mutation.

With regards to subtype distribution among the study participants, ~96% of the sequences tested were of subtype C both at PR and RT genes. Sequences from two individuals (1%) were identified as subtype B, while the remaining 3% were subtyped, at least by one genotyping algorithm, as mosaic forms at the PR and RT genes; including four (2%) BCs, one (0.5%) CRF_02AG (or BG? or ABG?), and one (0.5%) A1D. Analysis of genome diversity among the three groups of study participants showed that while certain non resistance mutations had high prevalence in all the three groups, others showed differential occurrences. Synonymous/non-synonymous substitution analysis and test of Shannon Entropy have also shown genome diversity differences among the three groups with distinct positional patterns, diversity at certain patches in the RT region being most prominent.

Discussions and Conclusions

Among chronically infected patients, the high baseline viral RNA load, low CD4+ T cell count, and more persons in clinical stages III and IV during treatment initiation show that treatment was started late after the patients' virological, immunological and clinical conditions have already been deteriorated. These poor baseline conditions were probably reflected by the high viral load values and low CD4+ T cell counts observed after treatment period of over three years among heavily treated patients. The transmitted drug resistance mutation rate documented in this study was in the WHO's 'low prevalence category'. However, more surveillance works of greater scope are required in terms of both geographic coverage and study population in order to understand the current prevalence of transmitted drug resistance. The detection of 22 drug resistance mutations among drug-naïve patients starting ART indicates the need for undertaking resistance testing before initiating therapy whenever settings allow. Among drug experienced patients, the most important drugs which were associated with NRTI and NNRTI resistance mutations were D4T and EFV, respectively. Suboptimal therapy could be the source of resistance mutations against these two drugs. It could also be due to impaired absorption of the two drugs, as the pharmacokinetics of ART drugs have not been investigated under the settings of Ethiopian patients. This therefore calls for the need to conduct therapeutic drug monitoring (TDM) at least in research settings on Ethiopian patients taking the various drugs. A lot of so-called non-resistance mutations occurred at positions known to confer resistance. Since most resistance/non-resistance mutation identification was done based on subtype B genetic background, and since virus variants behave differently under different settings, it would be of great benefit if phenotypic drug resistance profiles of isolates from Ethiopia with these non-resistance mutations are investigated.

With regards to subtype distribution, the overwhelming majority of the isolates belonged to subtype C at the pol region investigated in this study, showing that still this subtype has not been replaced by any other in Ethiopia. However, variants with mosaic genomic regions have been detected, showing the possibility of finding subtypes other than C, particularly recombinant forms, if further sequencing is made from genomic region larger than the PR and RT regions. Because of discrepancies between various genotyping algorithms in assigning subtypes, and because none of these mosaic forms had exact similarity with those previously known subtype identity, the mosaic variants identified in this study might be considered as Unique Recombinant Forms (URFs). Worth noting, however, is the population sequencing method used in this study might have masked presence of minority populations both in terms of variants with drug resistance mutations and those with different subtype profiles. In the latter case, use of only the short pol region might have contributed to the ambiguity of subtype identification in the non-C isolates detected in this study.

CHAPTER 1: INTRODUCTION

1. 1 INTRODUCTION

HIV is an RNA retrovirus belonging to the genus Lentivirus in the family of Retroviridea, the group constituting viruses characterized by chronic course of diseases, longer clinical latency, and persistent viral replication (Rubbert *et al*., 2007). There are also other immunodeficiency viruses, which are relatives of HIV and may or may not cause immunodeficiency diseases to other animals. These include Simian Immunodeficiency virus (SIV) in monkeys, feline immunodeficiency virus (FIV) in cats, Bovine Immunodeficiency Virus (BIV), Small Ruminant Lentivirus (SRLV), and other lentiviruses (Blacklaws and Harkiss, 2010; Craigo and Ross, 2010; Elder *et al*., 2010; Miller *et al*., 2000; Yamamoto *et al*., 2010).

The earliest cases of AIDS were documented on previously health homosexual men in 1981 in the United States, who presented with unusual infection of *Pneumocystis carinii* pneumonia and unusual cancer of Kaposi's sarcoma (CDC, 1981). The disease was characterized by a marked reduction of CD4 cell number and increased B-cell proliferation along with elevated plasma antibody concentration (Levy, 2009). It was soon established that the virus is transmitted through intimate sexual contact, blood and blood products, and mother-to-child transmission (Jaffe *et al*., 1983). In 1983, two years after the recognition of AIDS, the causative agent of this syndrome was identified as a retrovirus by various groups of workers (Barre-Sinoussi *et al*., 1983; Gallo *et al*., 1983). Following its characterization, the virus was given different names by different authors: Human T Cell Leukemia Virus III (HTLV-III) (Gallo *et al*., 1983), Lymphadenopathy-Associated Virus (LAV) (Basavapathruni *et al.*, 2007), AIDS-Associated Retrovirus (ARV-2) (Levy *et al*., 1984), and finally the naming settled on the present Human Immune Deficiency Virus (HIV) (Basavapathruni *et al.*, 2007). Studies conducted at the later years on autopsies of

early dead patients have indicated that HIV/AIDS was already circulating among different populations of the world before its recognition in 1981. For instance, a 15-year-old male who died of AIDS-like illness in 1969 (Kolata, 1987), and a Norwegian sailor and truck driver, who was probably infected time between 1961 and 1965 and died in 1976 (Froland *et al*., 1988) are few among a number of suspected/proven HIV infections prior to 1981. Since then, over 20 million people have died (Marison, 2001), 6 million have been newly infected annually (Hahn *et al*., 2000), and over 33 million were living with HIV/AIDS worldwide at the end of 2009 (UNAIDS, 2010).

It appears that no other infectious disease has been so extensively studied as has been HIV in the history of humankind (Farmer, 1996). Important discoveries such as characterization of the viral genome, the natural history, and pathogenesis of the virus were illuminated shortly after its discovery (Basavapathruni *et al.,* 2007). Because of these extensive studies, the routes of HIV transmission have been well characterized (Kamps and Hoffmann, 2007). An important fact to most African countries in this regard is the fact that the most common route for new HIV infection is unprotected heterosexual intercourses, with estimated likelihood of transmission from male to female being as high as 8-fold more likely than that from female to male (Hansasuta, and Rowland-Jones, 2001). It has also been documented that among the vertically transmitted new infections, as many as 42% of the children born to infected mothers in non-industrialized countries are infected with HIV (Hansasuta, and Rowland-Jones, 2001), and about 10-15 % of them acquiring it from breastfeeding (Kamps and Hoffmann, 2007).

As no curing medication has so far been made available, the best alternative to reduce the surge of HIV infection is to develop prophylactic and therapeutic vaccines based on scientific findings

(Wahern and Landy, 2002). Such vaccine development relies, among others, on complete understanding of the virus involved with regards to the nature and extent of its interaction, the various epitopes against which the desired vaccines are targeted at; and most of all, the extent of genotypic and phenotypic diversities and the immunologic implications of these diversities must be considered before any attempt is made to design and present for clinical trials (Gaschen *et al*., 2002; Novitsky *et al*., 2002). Unfortunately, in spite of documentation of huge wealth of information about HIV in general and HIV-1 in particular, still our understanding of this rampant virus is far from complete. This is largely because of the virus's inherent capability to evolve fast (Liu *et al*., 2002; Mani *et al*., 2002), owing to its higher rate of mutations coupled with its high rate of multiplication, error-prone reverse transcription, inter- and intra-subtype recombination, and effect of selective pressure from both the host immune system and widespread use of antiretroviral drugs (Blackard *et al*., 2002; Su *et al*., 2000).

Besides the virus's rapid evolution, however, there has been another source of problem in our understanding of HIV-1 diversity: only limited region(s) of the genome have been utilized most often for studying the various subtypes and recombinant forms (Blackard *et al*., 2002). In fact, a large number of phylogenetic studies conducted in the previous DNA analyses on HIV-1 variants were based on envelope sequences, most notably the C2V3 region, and/or partial *gag/pol* sequences, the results of which cannot be reliably utilized for vaccine development (Lole *et al*., 1999; Su *et al*., 2000). Not only is diversity linked to vaccine design and trial, but its direct significance to other biological properties make HIV genome study very important. In this respect, numerous studies have found genomic diversity of HIV to be closely associated to the biological differences that exist between the various types, groups and subtypes: for example, differences in the level of virus load (Baeten *et al*., 2007; Kiwanuka *et al*., 2008; Senkaali *et al.*

2004); disease progression (Matheron *et al*., 2003); clinical management (Keller *et al*., 2009); chemokine receptor use (Abraha *et al*., 2009; Esbjörnsson *et al*., 2010); vertical and horizontal transmission rate (Hawes *et al*., 2008); transcriptional activation rate and consequently replication rate (Gretti, 2006); immunological response to HIV infection and vaccine design (Li *et al*., 2009); response to ARV treatment (Hu *et al*. 2001); and drug resistance pattern (Chaplin *et al*., 2010); Martinez-Cajas *et al.,* 2009).

Given these biologically significant differences among the different HIV-1 subtypes, it is important for any country to understand and intensify studies on the genomic characteristics of the serotypes circulating among its population. Thus, this PhD study was largely designed to fill the present gap of knowledge on subtype diversity of Ethiopian HIV-1 isolates by sequencing the protease and reverse transcriptase genes, unlike most of the previous sequence studies which focused largely on the envelop region (Abebe *et al* 2000; Ayehunie *et al*., 1993; de Wit *et al*., 2002; Hussien *et al*., 2000). In doing so, the project was also intended to characterize drug resistance mutation patterns of the isolates from ARV drug-naïve and heavily treated individuals.

1.2. LITERATURE REVIEW

1.2.1 HIV Epidemiology

Since its identification in the early 1980's (Brown, 1986; CDC, 1981), the Human Immunodeficiency Virus pandemic has been mounting its attack globally at an unprecedented rate. The pandemic has been reported to have infected nearly 60 million and killed about 22 million people in the first 20 years of its existence (Marison, 2001). The global HIV incident rate reached its peak in 1997 (Figure 1.2.1.1), when 3.2 million (estimated range 3.0 million-3.5 million) new infections occurred (UNAIDS, 2010).

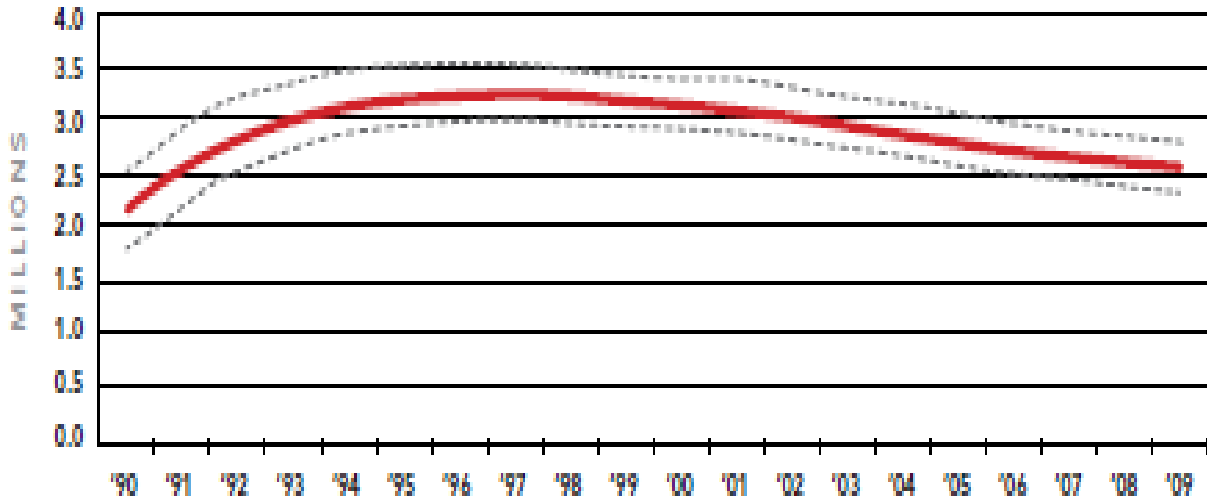


Fig. 1.2.1.1 Number of people newly infected with HIV 1990-2009.
(Source UNAIDS, 2010)

According to the most recent report from United Nations Program on HIV/AIDS (UNAIDS, 2010), the global burden of HIV at the end of 2009 was estimated to be 33.3 million (estimated range 31.4 million-35.3 million) (Figure 2.1.2); of this 2.5 million (range 1.7 million - 3.4 million) were children under the age of 15 while the rest 30.8 million constituted adults. The overall number of people living with the virus in 2009 was 27% higher than the number in 1999 (UNAIDS, 2010).

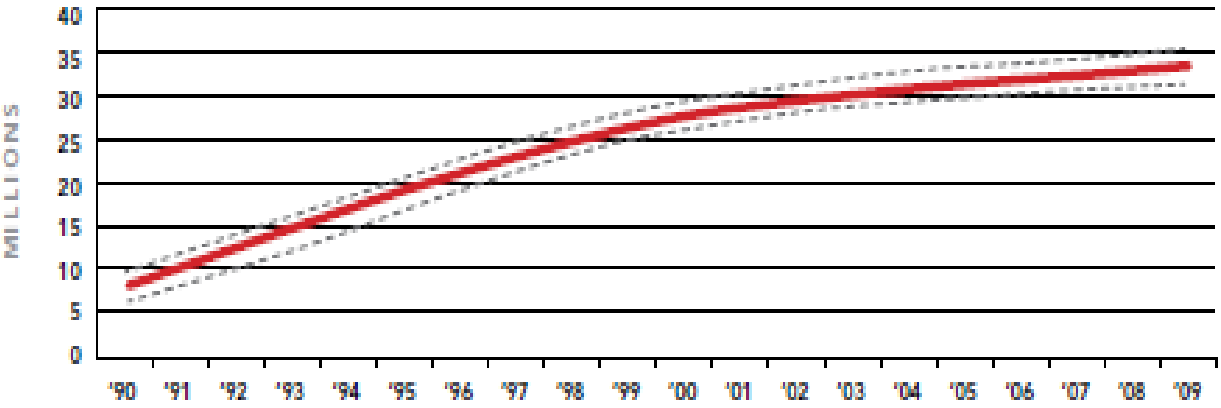


Figure 1.2.1.2 Global prevalence of HIV infection in the years 1990-2009.
(Source: UNAIDS, 2010)

Although the current prevalence study shows a decrease by over 25% in 33 previously high burden countries (22 of which being from sub-Saharan Africa) or stabilization in some others (UNAIDS, 2010), a shift in prevalence between different regions of the world has been recorded with respect to rate and mode of transmission. For example, the prevalence rate of 0.5% recorded in 2001 in Eastern Europe and central Asia was elevated to 0.7% in 2008 due to a high rate of new HV infections (UNAIDS/WHO, 2009). These new infections are believed to be caused by increased transmission significance of sexual route in addition to the once predominant injection drug use (Kilmarx, 2009). In spite of a decrease in overall prevalence in Africa, sub-Saharan Africa, where only 10% of the world populations live, is still the region carrying disproportionately 68% of all global HIV burden and 72% of global AIDS-related deaths (UNAIDS, 2010). Currently, an estimated 22.5 million (20.99 million-24.2 million) people are living with HIV and 1.8 million reported to be newly infected in 2009 (UNAIDS, 2010).

1.2.2 Emergence and Evolution of HIV/AIDS

The source of AIDS and origin of HIV have been subjects of debates and arguments ever since they were first recognized in early years of 1980s, involving everybody - scientists and nonscientists alike. Thus, scientific search for the link of HIV with its origin began soon after its discovery. Although epidemiological evidences are absent showing direct transmission of HIV from animal reservoirs to humans, researchers have gathered sufficient evidences for tracing the possible sources of HIV (both type 1 and 2). The springboard for these studies was the fact that related lentiviruses have been found infecting numerous non-human primates of sub-Saharan Africa origin (Sharp *et al*., 2005; Sharp *et al*., 2001). The closest lentiviruses known to infect over 30 species of primates, although not known to be pathogenic to the natural primate hosts, are Simian Immunodeficiency Viruses (SIV) (Sharp *et al*., 2001; Wain *et al*., 2007). In addition, the

observation that a wild-born chimpanzee from Gabon harboring a lentivirus (SIVcpzGAB1), which was isogenic to HIV-1, turned researchers to suspect chimpanzees as the possible source of HIV-1 (Sharp *et al*., 2005). Close examination of the genome from SIVcpzGAB1 has identified that this virus carried the same nine reading frames as HIV-1, including an accessory gene termed *vpu*, which had thus far been identified only in HIV-1 (Sharp *et al*., 2005). Moreover, phylogenetic analyses indicated that HIV-1 was more closely related to SIVcpzGAB1 than to any other SIV strains (Sharp *et al*., 2005).

Data from phylogenetic and genomic studies have further shown that West-central African chimpanzees (*Pan troglodytes troglodytes*) harbor strains of simian immunodeficiency viruses (SIVcpz) that are closely related to all the three groups of HIV-1 (Group M, N, and O) (Santiago *et al*., 2002). Many more studies conducted by several researchers have collected ample amount of evidences that substantiate the zoonotic transmission of primate lentiviruses to humans: similarities in viral genome organization; phylogenetic relatedness; prevalence in the natural host; geographic coincidence; and plausible routes of transmission (Gao *et al*., 1999). With regard to HIV-2, it is generally accepted that this virus is a close relative of SIVsm, a virus that is genomically indistinguishable and phylogenetically closely related to the HIV-2 (Sharp *et al*., 2001). This virus was found in wild-living sooty mangabeys (*Cercocebus atys*) whose natural habitat coincides with the epicenter of the HIV-2 epidemic (Sharp *et al*., 2001).

Similarly, it has been shown that all HIV-1 strains belonging to groups M and N are phylogenetically closely related to SIVcpz (SIVcpz*Ptt*) strains infecting *Pan troglodytes troglodytes*, a primate whose natural range coincides precisely with areas of HIV-1 group M, N, and O endemicity (Brookfield *et al*., 2006). However, recently new data emerged identifying two

more non-human primates as reservoirs of SIVs: *Pan troglodytes schweinfurthii* as a reservoir for SIVcpz (SIVcpz*Pts*) (Keele *et al*., 2006); and wild-living gorillas (*Gorilla gorilla gorilla*) as reservoir for SIVgor (Takehisa *et al*., 2009; Van Heuverswyn *et al*., 2006). While the SIVcpz*Pts* isolate from *P. t. schweinfurthii* resembles very closely with HIV-1 group M (Gao *et al*., 1999; Keele *et al*., 2006), SIVgor has been linked to HIV-1 group O (Takehisa *et al*., 2009; Van Heuverswyn *et al*., 2006) and a new type of HIV-1 recently isolated from a west African woman which was closer to SIVgor in all regions of the genome, and was designated a new group, Group P but distinct from groups M, N, and O (Plantier *et al*., 2009).

Two very interesting findings have emerged recently regarding the transmission of HIV-1 strains from non-human primates to humans. The first one is related to the genetic composition and original source before the involvement of chimpanzees in transmission of HIV-1 to humans. Bailes *et al*. (2003) have presented evidence that SIVcpz arose through successive cross-species transmission and recombination events of SIVs infecting monkeys on which chimpanzees prey; in *pol* region they resemble SIVrcm from red-capped mangabeys *(Cercocebus torquatus)*, whereas in *env* region they are closer to SIVgsn from greater spot-nosed monkeys (*Cercopithecus nictitans*) (Bailes *et al*., 2003). The authors speculated that because chimpanzees are known to hunt smaller monkey species, the simplest explanation appears to be that both SIVrcm and SIVgsn have been acquired by chimpanzees and recombined in that host which then followed transmission of the recombined SIVcpz from chimpanzee to humans (Bailes *et al*., 2003). The authors further reasoned that this finding has important implications: it provides evidence that, in addition to humans, another ape species acquired SIV by cross-species transmission under natural conditions; that the endemic infection of two chimpanzee subspecies indicates substantial

secondary spread of the initial hybrid; and that the recombinant chimpanzee virus was capable of spreading to humans (Bailes *et al*., 2003).

The other important finding is evidence generated by Wain *et al*. (2007) regarding the evolution and adaptation of HIV-1 in the human host after transmission from chimpanzees to humans. With the aim of looking for viral genetic changes associated with cross-species transmission that may represent adaptations of SIVcpz to its new human host, Wain *et al*. (2007) compared inferred sequences of the ancestors of the three independent groups of HIV-1 (group M, N, and O) with those of SIVcpz strains, and found a single site in the HIV-1 proteome (Gag-30 which was Meth in the chimpanzees) which appears to have undergone identical changes on each of the three occasions when virus was transmitted from apes to humans (replaced into either Arg or Lys), except in HIV-1 subtype C, in which Meth was conserved (Wain *et al*., 2007). This genetic evidence clearly showed that genetic modification was needed for the virus to adapt from its natural hosts (nonhuman primates) to the new host (human). Relations between and genetic diversity in HIVs and SIVs, and patterns of cross-species transmission is depicted in Figure 1.2.2.1.

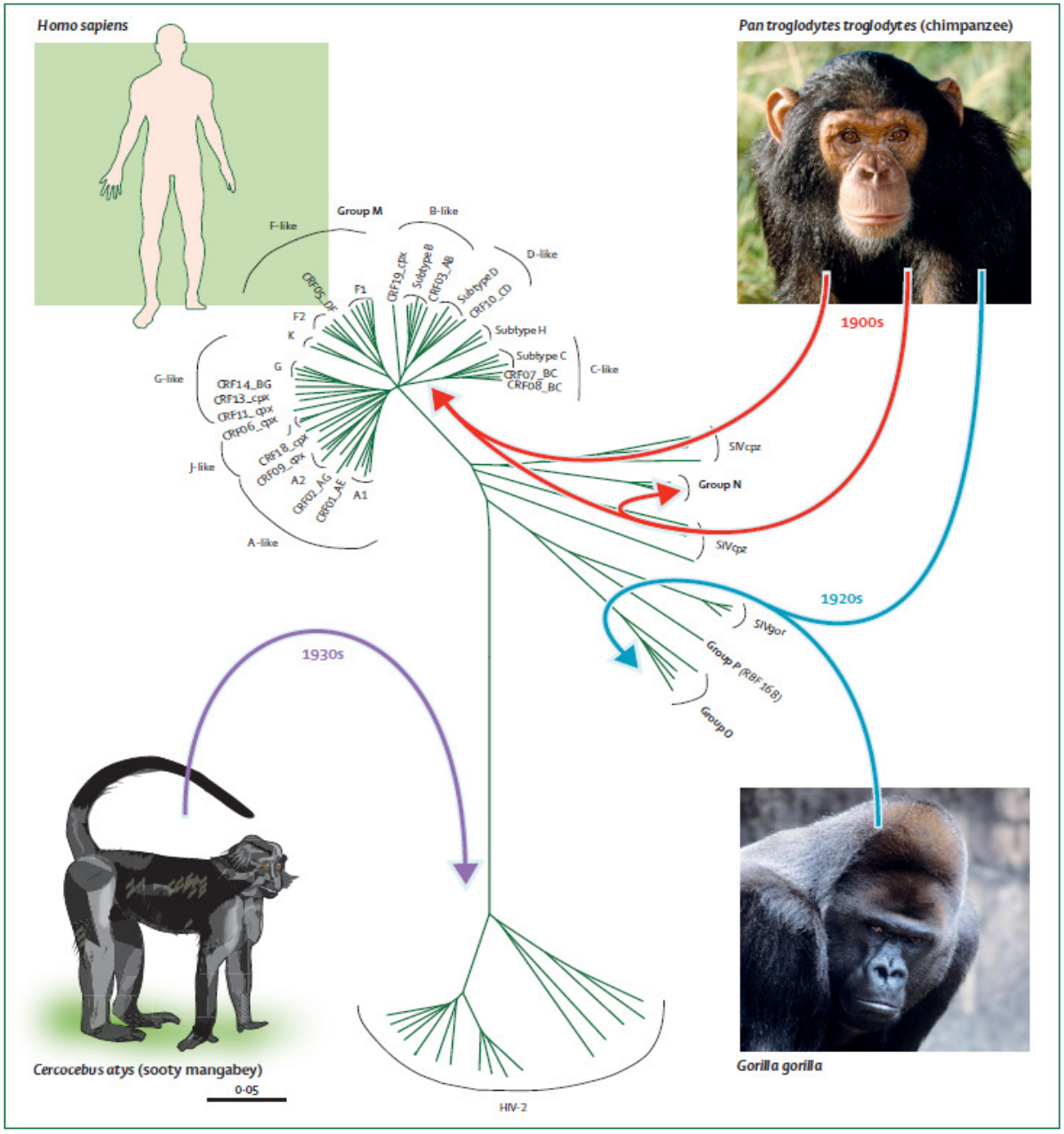


Figure 1.2.2.1 Relations between and genetic diversity in HIV-1 groups M, N, O, and P, HIV-2, and SIVs, and patterns of cross-species transmission
(Adapted from Tebit and Arts, 2011)

1.2.3 Structural Components of HIV

HIV is a spherical particle (Fig 1.2.3.1) surrounded by lipoprotein membrane containing 72 glycoprotein complexes integrated into the lipid layer, and each composed of an external gp 120 glycoprotein trimer non-covalently linked to a transmembrane spanning gp 41 proteins (Chan *et al*., 1997). The lipoprotein layer can also acquire such host proteins as HLA class I and II and adhesion proteins during budding (Rubbert *et al*., 2007). Beneath the lipoprotien membrane are attached the matrix protein (p17) and a conical core protein (p24), within which is located a protein-nucleic acid complex composed of two copies of HIV RNA bound to nucleoprotein p7 and the reverse transcriptase (p66) (Gottlinger, 2001). The core also contains the enzymes integrase (p32) and protease (p11), both of which required, along with the reverse transcriptase, for replication of the viral particles (Kuiken *et al*., 2010).

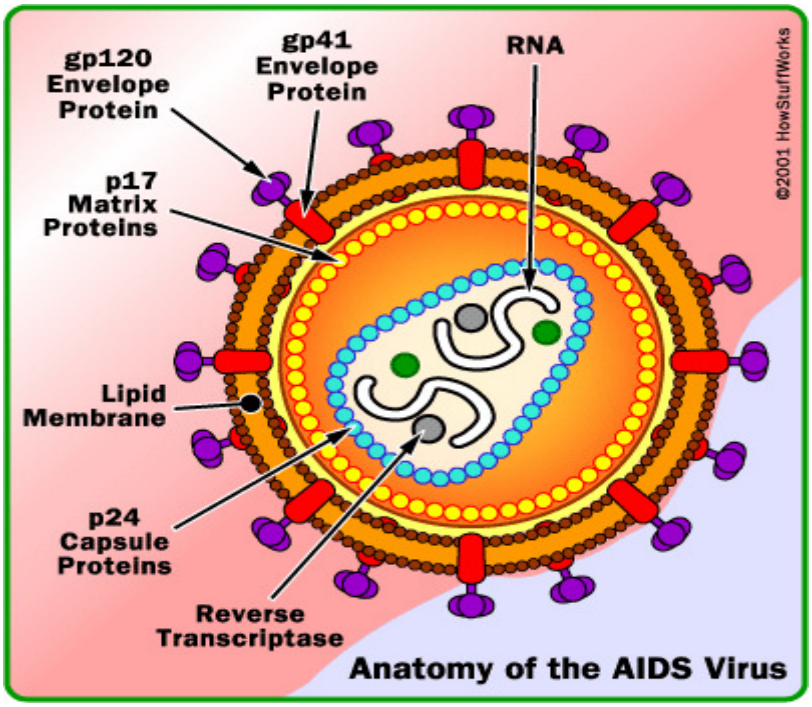


Figure:1.2.3.1 Structural components of HIV Virion
(Obtained from http://health.howstuffworks.com/ diseases-conditions/infectious/aids2.htm)

HIV RNA is a nearly 10-kb nucleic acid containing three coding regions for at least 15 protiens, three of which for structural and catalytic proteins: *gag* (for group antigen), *pol* (for polymerase), and *env* (for envelope) proteins, and the remainder for regulatory proteins (Basavapathruni and Karen, 2007). While *gag* and *env* genes code for nucleocapsides and membrane glycoprotiens, respectively, *pol* codes for the viral enzymes (reverse transcriptase, protease and integrase) (Kuiken *et al*., 2010). The initial *gag* product is a precursor protein (p55), which soon associates with the plasma membrane and processed to Matrix (p17), Capsid (p24), Nucleocapsid (p7) and p6 proteins by the viral protease (Gottlinger, 2001). Similarly, the viral enzymes are produced as Gag-Pol precursor polyprotein as the result of ribosomal frame shift near the 3' end of *gag* and processed by the viral protease to their respective primary structures (Kuiken *et al*., 2010). The primary *env* gene product is also a larger precursor gp160, which is processed to gp120 and gp41 by the viral protease (Rubbert *et al*., 2007). In addition, six other genes are carried within the RNA genome (*vif*, *vpu*, *vpr*, *tat*, *rev* and *nef)* (Kuiken *et al*., 2010). While *tat* and *rev* are regulatory in function, *vif, vpu (*in HIV-1, but replaced by *vpx* in HIV-2*), vpr,* and *nef* are accessory (auxiliary) genes, which determine the replicative and pathogenic characteristics of the viruses as well as the virus's ability to counteract host restriction factors (Levy, 2009; Strebel, 2003). Although *in vitro* studies in tissue culture showed that the latter four genes are dispensable to the virus reproduction, the fact that they are conserved in different strains indicates their importance *in vivo* (Rubbert *et al*., 2007).

Like any other lentivirus, the sequential arrangement of HIV coding genes in the RNA genome from 5' to 3'end is *gag-pol-env* (Kuiken *et al*., 2010) (Figure 1.2.3.2). These genes are flanked in both ends by the LTR regions (Long Terminal Repeat), non-coding regions that are connected to

host DNA after integration. Although LTR regions do not code for any protein, they play very crucial roles in the regulation of transcription (transcription initiation and polyadenylation) of proviral DNA into viral RNA and reverse transcription of cDNA from genomic RNA during replication since the promoter sites are located in these regions (Desfosses *et al*., 2005; Jeeninga *et al*., 2000). The regulatory and accessory genes are distributed within *pol, env* and 3'LTR; however, *tat* and *rev* exist as two physically separate exons within the *env* region (Rubbert *et al*., 2007). Other sequences that encode for more regulatory elements are also distributed here and there. A section of RNA sequence within LTR of the proviral DNA called TAR (Transactivation-responsive region), which forms a hairpin stem-loop structure with a side bulge, constitutes a binding site or target sequence for viral transactivation initiated by Tat (coded by the *tat*) and cellular proteins (Desfosses *et al*., 2005). Rev responsive element (RRE), which contains seven high affinity site for Rev protein, is another 200 nucleotide long RNA element encoded within *env* region spanning the boarder of gp120 and gp41 (Kuiken *et al*., 2010). Other relevant elements include Psi elements (PE), a TTTTTT slippery site (SLIP), Cis-acting repressive sequence (CRS), and Inhibitory/Instability RNA sequence (INS) (Kuiken *et al*., 2010).

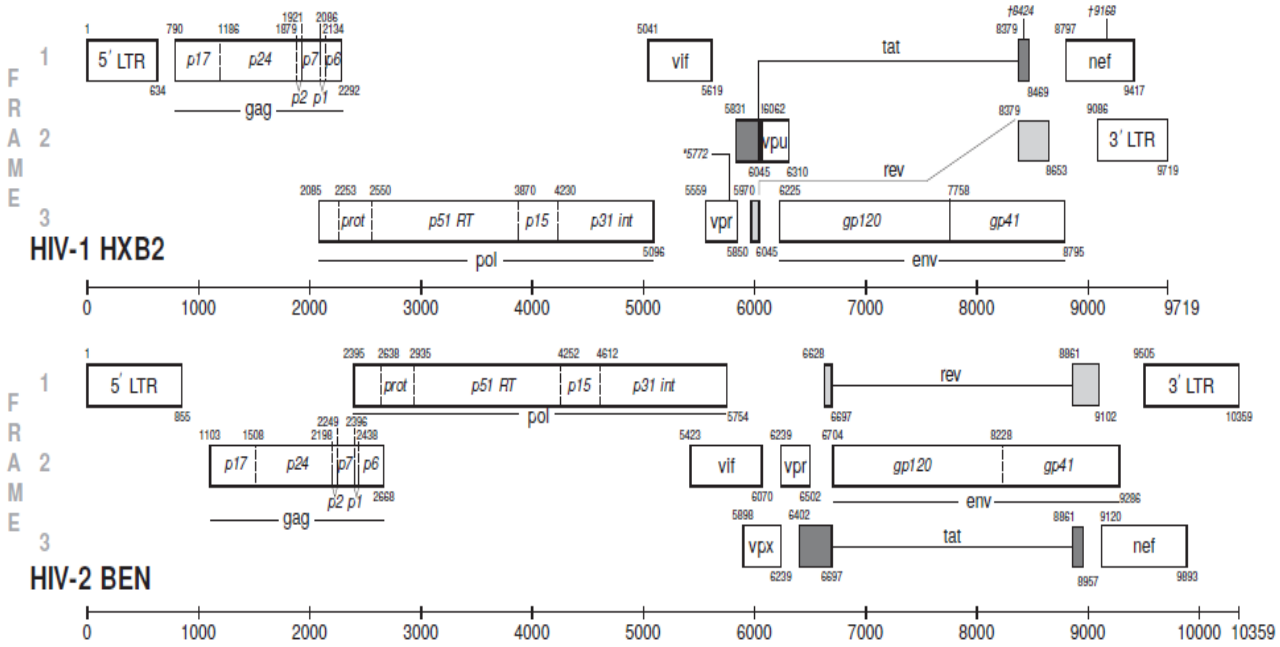


Figure. 1.2.3.2 Gene Maps of HIV-1 and HIV-2
(Adapted from from Kuiken *et al*., 2010)

Very recently, a study, using a technique called SHAPE (selective 2'-hydroxyl acylation analysed by primer extension), has elucidated the complete HIV-1 genome structure, and has found HIV RNA to have highly organized motifs with secondary structure of stems and loops in regulatory and protein coding-regions (Al-Hashimi, 2009). This study revealed that the structured regions of the HIV-1 genome are concentrated in about 21 large domains (Watts *et al*., 2009). In between these structured domains are located unstructured regions that are thought to slow down the movement of ribosomes during the synthesis of polyprotien, which gives time for individual proteins to assume their secondary structure (Watts *et al*., 2009). Figure 1.2.3.3 shows structure of the HIV-1 NL4-3 genome as depicted by Watts *et al*. (2009).

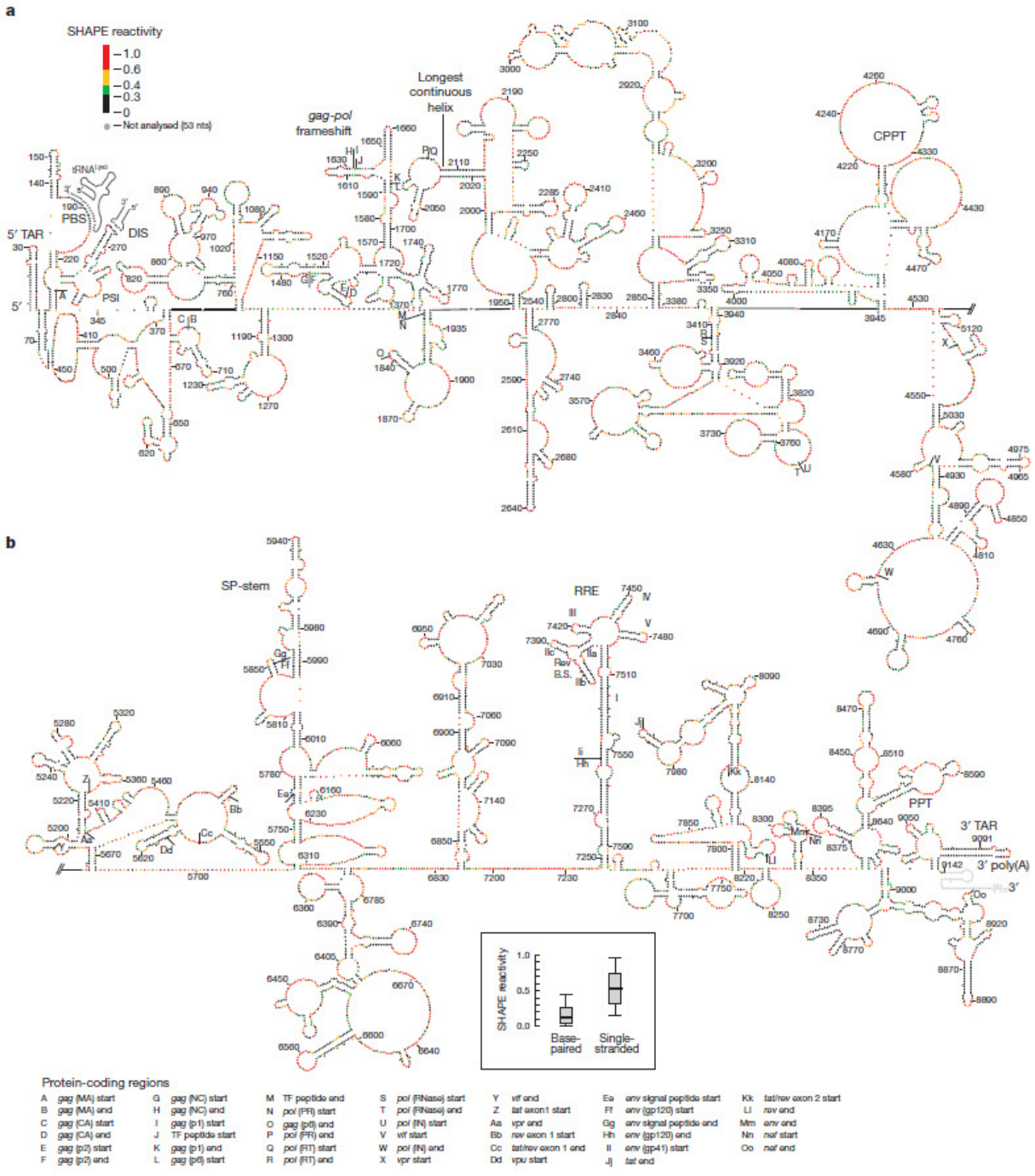


Figure1.2.3.3 Structure of the HIV-1 NL4-3 genome.

The 59 (a) and 39 (b) genome halves are shown. Nucleotides are coloured by their absolute SHAPE reactivities (see scale in a). Every nucleotide is shown explicitly as a sphere; base pairing is indicated by adjacent parallel orientation of the spheres. Protein domains are identified by letters; TF, transframe peptide; nts, nucleotides. Important structural landmarks are labelled explicitly. (Adapted from Watts *et al.,* 2009)

1.2.4 Replication

Like any member of retrovirus family including the lentiviruses, HIV replication in general can be viewed most simply as a process proceeding in an ordered, step-wise manner, which is characterized by key events, including viral binding and entry, reverse transcription of the RNA genome, integration of the resulting proviral DNA into the host genome, viral gene expression, virus assembly, and progeny virus budding and maturation (Gomez and Hope, 2005). These series of events can be divided into two overall phases: "early" and "late" (Freed, 2001). Figure 1.2.4.2 summarizes a simplified life cycle of HIV-1.

1.2.4.1 Entry

Entry of HIV-1 into host cells is mediated by an HIV-1 envelope protein complex hat is initially produced as the precursor gp160, which is extensively glycosylated and proteolytically cleaved into two subunits by a cellular convertase: surface subunit (gp120) and the transmembrane subunit envelope proteins (gp41) that remain noncovalently associated and oligomerize as trimers, on the surface of the virion (Chan and Kim, 1998; Gomez and Hope, 2005). The transmembrane gp41 contains two helical regins (HR1 and HR2) and a fusion peptide (Gomez and Hope, 2005). The early events of HIV-1 replication begin with receptor binding in which HIV-1 surface glycoprotien gp120 is adsorped to the CD4 receptor. This induces a change in shape within gp120 that brings the chemokine receptor binding domains of the gp120 into proximity with the host cell chemokine receptor, which could be, depending on the viral tropism determined by the V3 loop of gp120, CXCR4 (for T lymphocytes tropic) or CCR5 (for macrophage tropic). Consequently, the co-receptors will bind to the virion particle (Chan and Kim, 1998; Wyatt and Sodroski, 1998).

This engagment trigers conformational changes in the structure of gp41 such that the previously buried portion of the transmembrane glycoprotein gp41, which is the fusion peptide, enables the viral envelope to fuse with the host cell membrane (Bergamaschi and Pancino, 2010), followed by further conformationl change in gp41 that allows HR1 and HR2 to interact with each other to form stable six-helix bundle structure (Figure 1.2.4.1) (Chan and Kim, 1998). This fusion leads to the release of genome-containing viral core into the cytoplasm of the infected cell, where along with various enzymes (including reverse transcriptase, integrase, ribonuclease, and protease) and the single-stranded RNA genomes within the core or capsid of the virus, are released into the cytoplasm following uncoating (Gomez and Hope, 2005; Taylor *et al*., 2008). A different version of viral entry into the host cell also occurs occasionally through endocytosis pathway (Pope and Haase, 2003; Simon *et al*., 2006); however, this usually results in inactivation or degradation in the lysosomes of the infected host cell (Gomez and Hope, 2005).

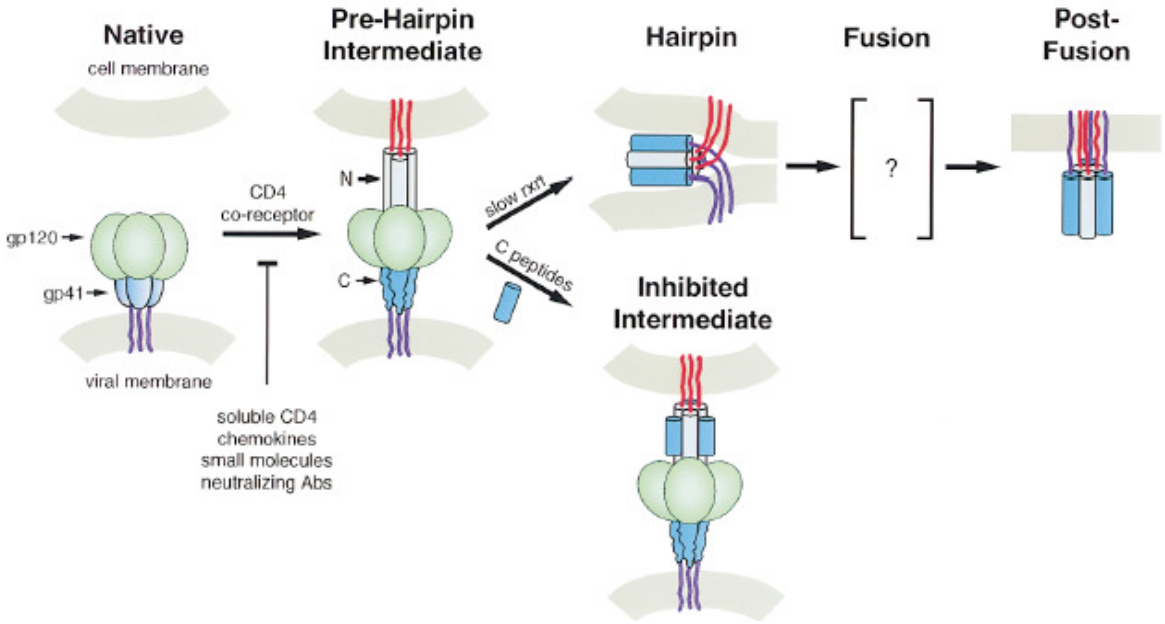


Figure 1.2.4.1 Model of HIV membrane fusion
(Adapted from Chan and Kim, 1998)

1.2.4.2 Replication and transcription

As soon as the viral RNA genome is released into the cytoplasm, it moves to the nucleus through microtubule-based transport, during which time the viral RNA genome-associated error-prone enzyme reverse transcriptase makes a complementary minus-strand cDNA copy of the RNA genome by reverse transcription using its RNA-dependent DNA polymerase activity (Rubbert *et al*., 2007; Sarafianos *et al*., 2009). Reverse transcription is initiated by a tRNA of the host cell serving as the primer; the primer tRNA is incorporated into virions during viral assembly and is annealed to an 18-base region located at the 5' terminal region (Oh *et al*., 2008; Wei *et al*., 2005). Simultaneous to synthesizing the (-) cDNA, the same reverse transcriptase degrades the (+) RNA template by its ribonuclease activity before synthesizing the complementary (+) DNA from (-) cDNA by its DNA-dependent DNA polymerase activities. Once the (+) DNA is formed, it remains paired with the (-) cDNA making a double stranded intermediate DNA, which then gives rise to pre-integration complex (PIC) composed of the double stranded viral DNA, integrase, matrix, Vpr, RT, and the high-mobility group DNA-binding protien (HMGI/Y) (Zheng *et al*., 2005). The double stranded viral DNA is then transported into the cellular nuleus via nuclear pore and becomes integrated with the host genome and forms a provirus with the help of the enzyme integrase (Rubbert *et al*., 2007; Sarafianos *et al*., 2009). This ends the early phase of HIV replicaction.

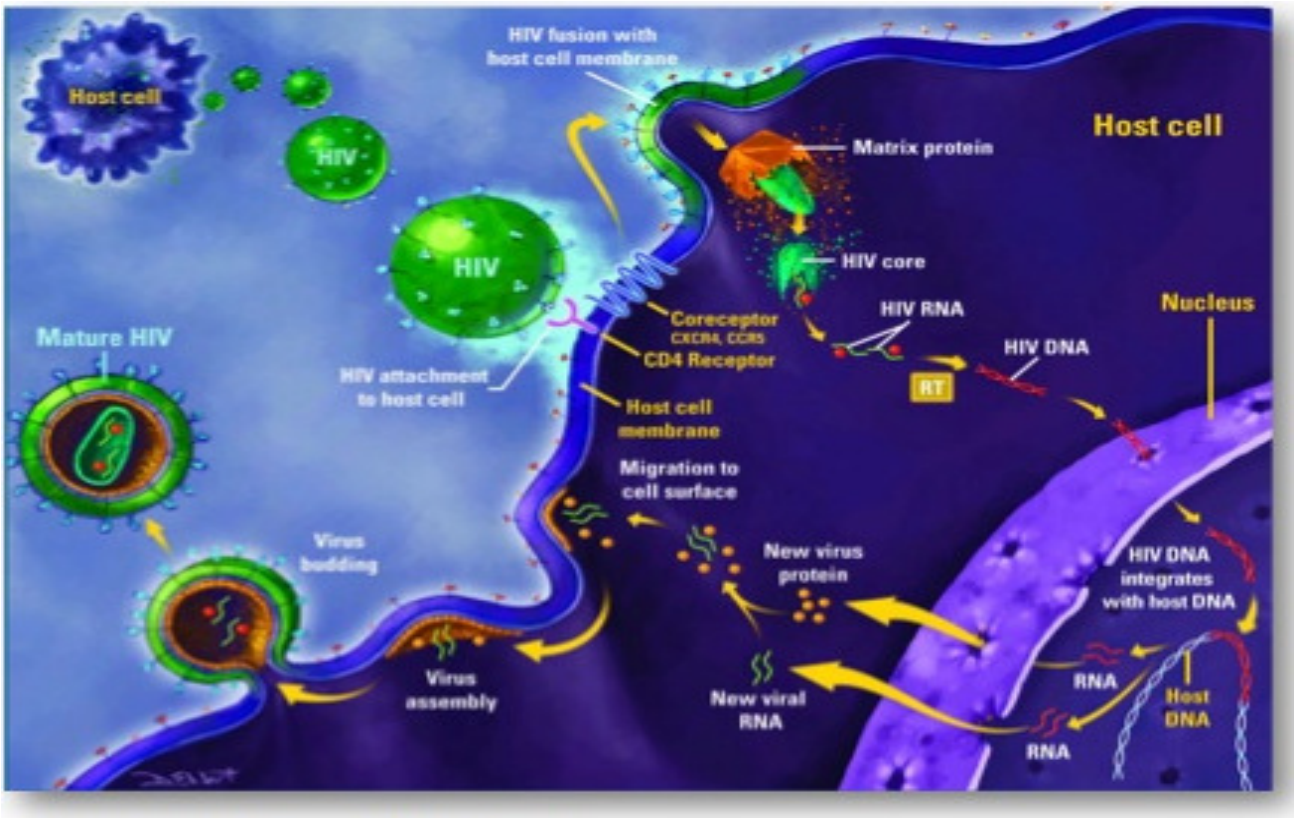


Figure 1.2.4.2 A simplified Schematic diagram of HIV-1 Life cycle
(Adapted from HIV Web Study: www.HIVwebstudy.org)

After integration, the provirus may exist in one of the two forms – either a latent or a productive state (Stevenson, 2003). Both host and viral factors determine the immediate fate of the integrated provirus as to which of the two forms it might assume: thease are; genetic factor of the viral strain, the type of cell infected, and the production of host cell proteins (Spina *et al*., 1997; Stevenson, 2003). The provirus that has infected memory T4 cells, monocytes, macrophages, and denditic cells remain latent provirus and serve as stable reservoirs of HIV, while that integrated in activated T4 cells (which comprises of up to 95% of infected cells) is in a state of productive infection (Simon and Ho, 2003; Stevenson, 2003). Thus the productively infected cells are immediately engaged in producing new viruses through series of coordinated steps (Stevenson,

2003). On the other hand, the provirus in the latently infected cells needs to be activated in order to be involved in viral production by replication (Simon and Ho, 2003; ; Stevenson, 2003). This activavtion is achieved by antigenic stimulation of the infected cells, or their activation by cytokines, endotoxins or Lipopolyaccharides (LPS), and superantigens (Rubbert *et al*., 2007).

The late phase of HIV-1 replicaction involves all post-proviral integration events following activavtion of the provirus, in which the host cell's genetic machineries are commandeered by the provirus (Weiss, 2001). The host cell RNA polymerase II transcribes viral mRNA from the (-) sense proviral DNA (Rubbert *et al*., 2007). The activation of the provirus and transcription of mRNA involves intricate cascades of biochemical processes that require interplays between host factors and viral proteins and proviral sequences at LTR regions (Hiscott *et al*., 2001; Suhasini and Reddi, 2009; Zheng *et al*., 2005). Up on transcription, a 9-kb unspliced mRNA is formed that is used for three viral functions: i) synthesis of the Gag polyproteins p55, which will be cvleaved later by HIV protease to matrix proteins (MA, p17), capsid proteins (CA, p24), and nucleocapsid proteins (NC, p7); ii) synthesis of the Gag-Pol polyproteins p160, which will be cleaved later by HIV protease to matrix proteins (MA, p17), capsid proteins (CA, p24), proteinase molecules (protease or PR, p10), reverse transcriptase molecules (RT, p66/p51), and integrase molecules (IN, p32); and iii) becoming the genome of RNA later when the new viruses mature (Gomez and Hope, 2005).

Another avenue of the 9-kb mRNA is getting spliced to form a 4.4-kb mRNA and a 2-kb mRNA, each of which being responsible for the synthesis of their respective structural, regulatory and accessory proteins. The 4.4-kb mRNA is used to synthesize the Env polyprotien gp160, which is cleaved later by protease to gp120 and gp41; and also to synthesize the regulatory/accessary proteins Vif, Vpr, and Vpu (Rubbert *et al*., 2007). Likewise the 2-kb mRNA is used to synthesize

the other regulatory/accessory proteins Tat, Rev, and Nef (Rubbert *et al*., 2007). While multiply spliced mRNAs (2-kb) transcripts can exit the nucleus freely, unspliced (9-kb) and singly spliced (4.4-kb) transcripts require Rev and the Rev Response Element (RRE) RNA structure for their transportation to the cytoplasm (Suhasini and Reddy, 2009; Zheng *et al*., 2005). With the help of accumulated Rev protein in the nucleus, unspliced mRNAs are bound and transported to the cytoplasm, where they are retained unspiced until viral particle assembly is accomplished (Suhasini and Reddy, 2009).

1.2.4.3 Viral assembly, release, and maturartion

After synthesis of viral proteins and genomic RNA, assembly of the viral particles begins at cellular membrane. The HIV Gag polyprotien is identified to orchestrate this process by serving as a scaffold to promote assembly of progeny virions at cellular membranes and recruits components of the vesicular protein sorting pathway to facilitate virus budding (Gottwein and Krausslich, 2005). It is generally accepted that HIV-1 viral particles takeover a system normally involved in the intracellular sorting of membrane proteins to exit infected cells (Gomez and Hope, 2005). The detection of Gag proteins in plasma membrane and the membranes of endosomes suggetst that viral assembly and budding may not be localized to plasma membrane only (Porter *et al*., 2010).

In this regard, two models have been proposed: the first model proposes that following synthesis, Gag traffics to endosomal membranes, and upon exocytosis, is deposited on the PM, where it serves as a site for productive virus assembly; the other model (the lipid raft budding model) proposes that Gag is first trafficked to the PM, where virus assembly occurs, and then excess Gag is internalized to intracellular endosomal compartments called multivesicular bodies (MVB) that

serve as sites of productive virus assembly (the Trojan exosome hypothesis) (Gomez and Hope, 2005; Hurley *et al*., 2010; Porter *et al*., 2010). Because of the presnce of supporting evidences for both of these models, it is believed that both of them could be plausible models that may not be mutually exclusive (Gomez and Hope, 2005; Gousset *et al*., 2008).

In general, HIV-1 Gag recruits site-specific cellular lipids and proteins to the budding virion, in addition to viral proteins (Hurley *et al*., 2010). Moreover, viral assembly and release has been recognized to occur at cell-cell junction called Virological Synapsis, where Gag and Env proteins are accumulated during assembly (Gomez and Hope, 2005; Gousset *et al*., 2008). Wherever the assembly sites may be, the polyprotiens are cleaved into their respective functional proteins by viral protease: Env polyprotein (gp160) cleaved and processed into the two HIV envelope glycoproteins gp41 and gp120 and transported to the plasma membrane of the host cell, where gp41 anchors the gp120 to the membrane of the infected cell; the Gag (p55) and Gag-Pol (p160) polyproteins also associate with the inner surface of the plasma membrane along with the HIV genomic RNA as the virion under formation begins to bud from the host cell (Gelderblom, 1997). In the end, mature viral particles, with full capability of infection, are produced after budding following the vleavage of the various polyprotiens (Gelderblom, 1997).

1.3. HIV/AIDS Natural History and Pathogenesis

1.3.1 Natural History of HIV Infection

The natural history of HIV infection has been arbitrarily separated into primary or acute infection, latent or asymptomatic, and Advanced HIV infection or AIDS phase (Pilcher *et al*., 2004). These phases and events along with their associated markers are depicted in Figure 1.3.1.1 and described briefly below.

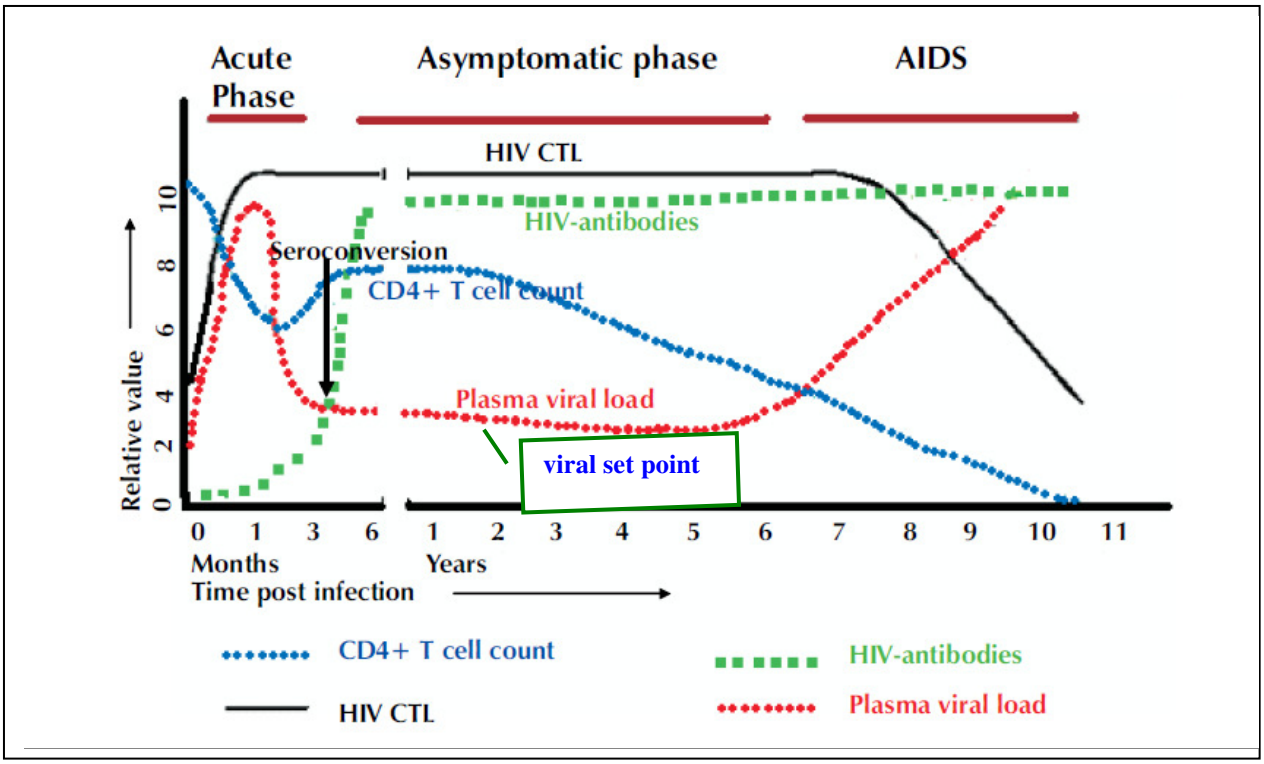


Figure 1.3.1.1 Natural History of HIV-1 infection
(Adapted from WHO, 2009)

For the purpose of guiding the diagnosis, evaluation, and management of HIV/AIDS, WHO and the CDC have their own definition of the clinical stages of HIV infection. WHO's staging definition was developed in 1990 and revised in 2007, in which the stages are categorized as 1-4, progressing from primary HIV infection to advanced HIV/AIDS (WHO, 2007). The CDC on the other hand has classified case definitions into stages 1-3 and a fouth one stage unknown, based on clinical and laboratory findings (CDC, 2008).

Primary/acute HIV infection

Soon after HIV infects its target cells at the mucosal surfaces via various routes (except for direct parenteral or in utero transmission) an incubation period of primary infection ensues lasting for two to four weeks (Altfeld and Walker, 2007), which heralds beginning of a persist

infection throughout the remainder of the person's life. This is followed by clinically acute HIV infection period, which is subclinical for most of the infected person (80-90%) with insginifican consequences or 'flu-like' symptomes and hence passes unnoticed (Daar *et al*., 2001), while small proportion of individals may develop rash, fever and lymphadenopathy within 2-3 weeks; and still some get pharyngitis, erythematous maculopapular rash, arthralgia, myalgia, retro-orbital headache, malaise, diarrhea and vomiting (Kahn *et al*., 1998). In the first few weeks of the acute infection phase, HIV can be detected in blood serum and plasma before the formation of antibodies routinely used to diagnose infection, and high levels of viremia and shedding can be demonstrated at mucosal sites (Pilcher *et al*., 2004).

During the primary HIV infection, the virus replicates profusely reaching to the level of 100 million copies of RNA/ml (Altfeld and Walker, 2007). This high viral titer is accompanied by a marked drop in the numbers of circulating $CD4^+$ T cells, and is associated with the activation of $CD8^+$ T cells, which kill HIV-infected cells (Pantaleo *et al*., 1997). Concomitant to this response of CD8+ T cells and appearance of neutralizing antibodies developed later over the following weeks, the high level RNA drops down until it reaches a viral setpoint, which is defined as the relatively stable level of plasma viremia that is observed during the chronic stage of infection and predictive of long-term prognosis, with higher levels of viremia being associated with a more rapid loss of CD4+ T Lymphocytes and faster progression to AIDS (Simon and Ho, 2003). As the result, the acute infection is resolved, giving way to the next clinically laten phase (Klatte, 2010). Before the clearing happens, however, the virus is already seeded into various target tissues and organs that will serve as reservoirs from which the viruses are continuously trickled into the blood stream throughout the course of the disease (Weber, 2001).

Viral replication during acute infection and viral setpoint are influenced by several factors including viral replicative fitness, host immune responses and host genetic factors (Stevenson, 2003). With regards to host immune response, antibody response against HIV-1 with neutralizing capacities lag behind for weeks, and hence are rarely detectable during primary HIV-1 infection, and if ever, they are largely of non-neutralizing types targeting debris of HIV (Letvin and Walker, 2003). Nonetheless, a number of studies have demonstrated a crucial role of HIV-1-specific cellular immune responses for the initial control of viral replication during this stage of infection (Altfeld and Waker, 2007). The CD8+ T cells eliminate HIV-infected cells directly by MHC classI-restricted cytolysis or indirectly by secreting cytokines, chemokines or soluble factors (Altfeld and Waker, 2007). In addition, host genetic factors like presences of CCR5delta32 (Biti *et al*., 1997), HLA class I alleles such as HLA-B27 and HLA-B57 are crucial in determining viral set point and rate of disease progression (Altfeld *et al*., 2006; Altfeld *et al.*, 2003; Migueles and Connors, 2010), which are all associated with lower viral setpoints and slower disease progression.

Latent/asymptomatic or chronic infection

Following the primary/acute infection, the CD4+ T cells increase in number albeit not to the level of pre-infection (Weber, 2001). This marks the beginning of a clinically latent period of asymptomatic/chronic infection phase, whose length varies from few weeks upto 20 years (Altfeld and Waker, 2007; Klatte, 2010). Although peripheral viral level is very much reduced compared to the acute infection phase, HIV viral particles are still active within lymphoid organs, where a large amount of viruses are trapped by Follicular Dendritic Cells, from where CD4+ T cells in the surrounding tissue are infected (Burton *et al*., 2002; Stevenson, 2003). As the result of this persistent viral infection, the CD4+ T cell level continues declining slowly, though not at a

constant rate, until it reaches to a level (≤ 200 cells/uL) where the immune ystem can no longer controls microbial infections, leading to emergence of several opportunistic infections (Klatte, 2010). This marks the beginning of late stage infection or progression to AIDS.

Late stage HIV infection

Once the infection has progressed to this stage and the level of CD4+ T cells declines to ≤200 cells/uL, the rate of further CD4 decline and the level of plasma viral load increament is heightened tremendously (Moroni and Antinori, 2003). The principal feature of this stage is the emergence of myriad of opporunuistic infections and neoplasms of AIDS. In addition, direct damage of organs (kidney, brain, gastrointestine, bonemarrow, and heart has been recognized (Moroni and Antinori, 2003).

It is generally recognized that HIV infected persons are categorized into four strata depending on the rate of their progression, beginning from seroconversion to AIDS phase and subsequently death: these are; typical progressors, rapid prgressors, nonprogressors, long servivors, and elite controllers (Klatte, 2010). While rapid progressors (accounting for about 10% of all HIV-infected persons) reach AIDS stage within short period of 2-3 years following initial infection, the typical progressors and nonprogressors (or also known as long survivors) can remain in the chronic stage without a significant and progressive decline in immune function for average of 8-10 and more than 10 respectively. Long survivors are characterized by a stable CD4 lymphocyte count, negative plasma cultures for HIV-1, fewer HIV-infected cells, and a strong virus-inhibitory CD8+ T-lymphocyte response (Klatte, 2010). On the other hand, the elite controllers of HIV suppress viremia, although viremic initially following primary infection with HIV, below the limit of detection (<50 copies/mL) even in the absence of antiretroviral therapy. They are also

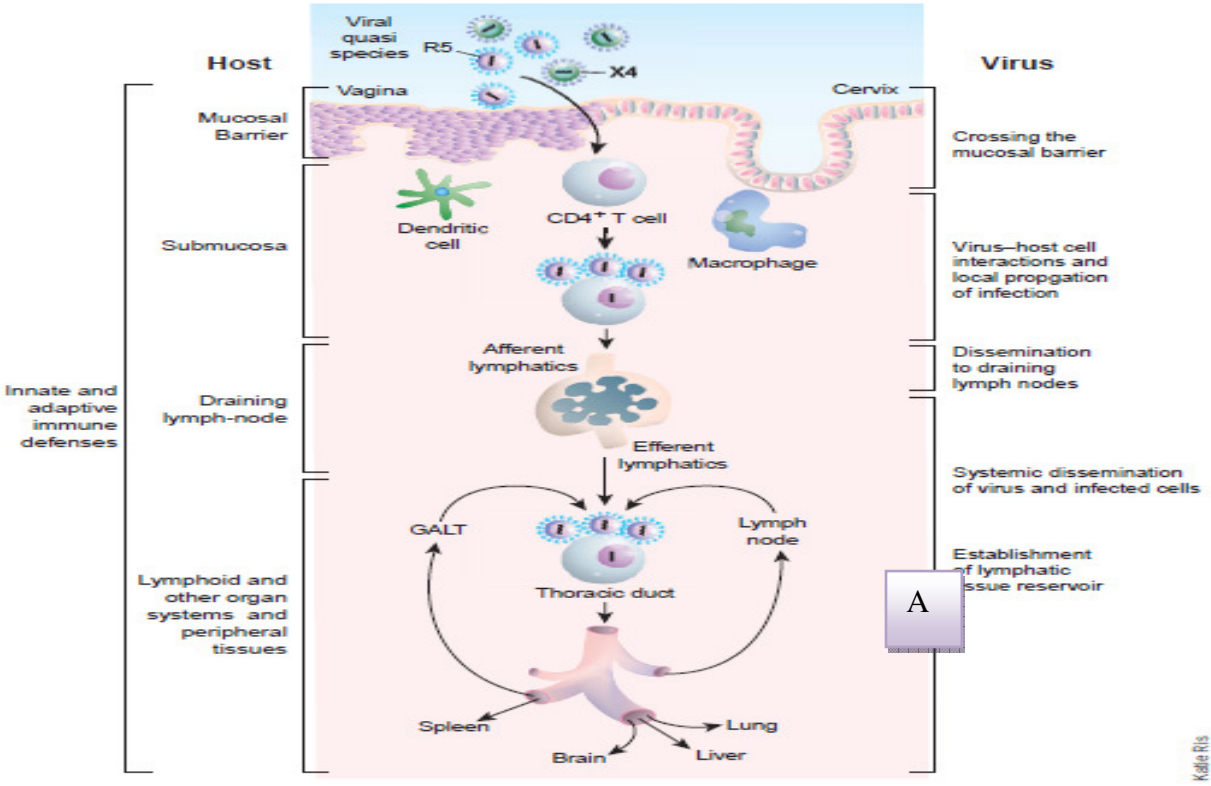
characterized by minimal decline in CD4+ lymphocytes, mounting polyfunctional CD4 response without loss of specific CD4+ clones; immune systems with preserveed cytolytic NK cells; and mounting a more polyfunctional CD8+ cytotoxic lymphocytic response to HIV (Klatte, 2010).

1.3.2 HIV Pathogenesis

It is now a well-established fact that HIV primarily infects cells with CD4 surface receptor molecules, which are used to gain entry into the cells. This receptor is shared by many cells in the human body and so a number of target cells are available for the entry of the virus, although CD4 bearing lymphocytes are the primary cellular homes; cells of the mononuclear phagocyte system, principally blood monocytes and tissue macrophages, T lymphocytes, B lymphocytes, natural killer (NK) lymphocytes, dendritic cells (Langerhans cells of epithelia and follicular dendritic cells in lymph nodes), hematopoietic stem cells, endothelial cells, microglial cells in brain, and gastrointestinal epithelial cells are the primary targets of HIV infection (Stebbing *et al*., 2004). Thus, anatomical sites afflicted by HIV are lymphoid organs such as lymph nodes, gut, spleen, the central nervous system, and the genitourinary tract (Simon and Ho, 2003; Stebbing *et al*., 2004).

During initial HIV infection, the virus crosses the mucosal barrier through any of the following three methods, resulting in infection of a small number of CD4+ T lymphocytes, macrophages, and dendritic cells: i) infection of intraepithelial lymphocytes; ii) capture or infection of DCs by virus attachment to mannose C-type lectin receptors (eg, dendritic-cell–specific intercellular adhesion molecule 3–grabbing nonintegrin or DC-SIGN) located on dendritic cells and macrophages; or iii) transcytosis, or access of subepithelial target cells through disruptions in mucosal integrity (Pope and Haase, 2003; Simon *et al*., 2006). Once the virus crosses the mucosal barrier, it begins propagating locally in the infected cells and disseminates into draining

lymphnodes and then, via blood, into various target organs including other lymphoid organs, where tisse reservoir is established (Figures 1.3.2.1A and B) (Pope and Haase, 2003). HIV-1 that is bound and internalized by a DC may allow infection of nearby CD4+CCR5+ cells and/or efficient presentation of HIV-1 to susceptible target cells after DCs migrate to draining lymph nodes (Pilcher *et al*., 2004). Entry of HIV-1 into target cells by fusion or endocytosis results with sharply different functional consequences: while fusion promotes productive infection of the host cell, endocytosis generally leads to virion inactivation in acidified endosomes or degradation in lysosomes. Both processes are restricted to the CD4-expressing subset of cells and that both pathways require the initial binding of HIV virions to surface CD4 receptors (Schaeffer and Greene, 2003).



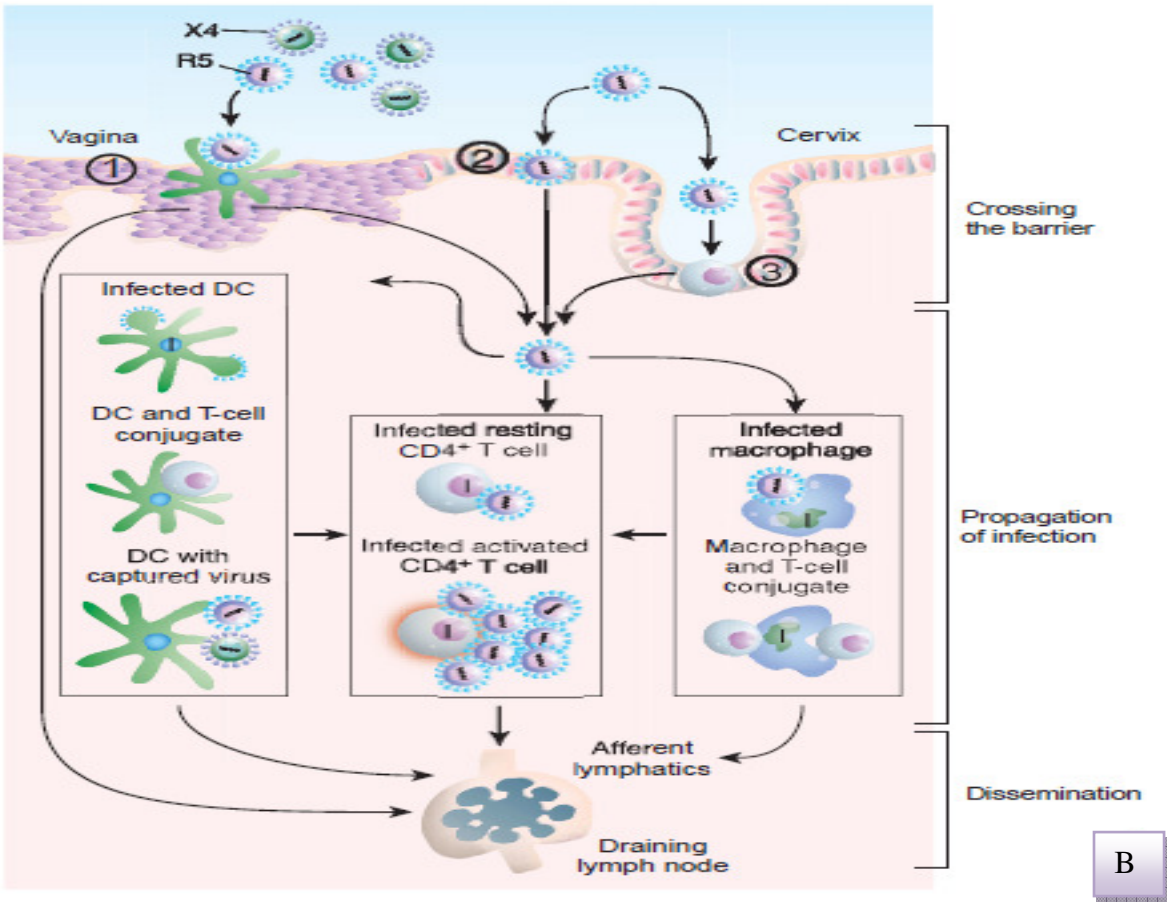


Figure 1.3.2.1 HIV transmission in the female genital during acute infection.
Panel A: shows the path through which the initial infecting virus passes and where it is disseminated. Only R5 viruses are preferentially taken up at the mucosal surface. Panel B: shows how the initial HIV-1 infects target cell, propagates and disseminantes into the submucosal and regional lymphnodes. (1) Infection of DCs. (2) Transcytosis (3) Infection of intraepithelial lymphocytes (Adapted from Pope and Haase, 2003)

The principal strategy of HIV-1 is its taking advantage of cellular pathways while neutralising and hiding from the different components of the immune system. It does this through various tricky ways. In the early steps of the infection, it enters to target cell without causing immediate lethal damages but the entry process can stimulate intracellular signal cascades, which in turn might facilitate viral replication (Simon *et al*., 2006). In addition, the initial replication is composed of homogenous population with few viral variants, which leads to modest primary amplification probably because of the viral isolates being slow growing at early in the course of

infection (Simon *et al*., 2006). However, four weeks after infection, viral loads in blood, genital secretions and other compartments peak at very high levels (Pilcher *et al*., 2004), because of migration of infected T lymphocytes or virions into the bloodstream, and consequently secondary amplification in the gastrointestinal tract, spleen, and bone marrow results in massive infection of susceptible cells (Simon *et al*., 2006). Increasing HIV-1 replication during this stage leads to a general activation of CD4+ and CD8+ cells, including HIV-1–specific CD4+ cells, a condition that predisposes activated CD4+ cells to be successfully infected with HIV-1 (Pilcher *et al*., 2004). This high rate of amplification is paralleled with viral quasispecies diversification in distinct biological compartments that results in the generation and archiving in long-lived cells (ie, viral reservoirs) of mutant viruses that are resistant to antibody neutralisation, cytotoxic T cells, or antiretroviral drugs (Simon *et al*., 2006). At this peak viremic stage, the half-life of a virion is so short that half of the entire plasma virus population is replaced every 6 hours or less, as the result of which the total number of virions that are produced and released in an untreated HIV-1-infected person can reach to the order of $\geq 10^{10}$ particles per day (Simon and Ho, 2003).

The various infected cell types have been observed to contribute to the plasma viral level differently and at different time point in the course of HIV-1 infection (Figure 1.3.2.2). This is because propagation of HIV-1, which is determined by the activities of integrated provirus, depends on metabolic and activavtion state of the infected cell; and the longevity of the provirus determines the persistence of high level viral replication, which in turn is dependent on the lifespan of the infected cell (Stevenson, 2003). These two characterstics are expressed by the various HIV-1 target cells differently. It has been demonstrated that viral replication is rapid and efficient in activavted infected cells (Pope and Haase, 2003). Thus, activated T cells and macrophages are the principal sources of viremia during the acute infection (Stevenson, 2003).

Thus, with continuous activavtion and rapid turnover rate of CD4+ T lymphocytes during the prolonged chronic phase, HIV-1 always gets pool of target cells to infect and destroy. While activated T cells are highly cytopathic paralleling their high viral productivity, resting T (G1) cells and quiescent (G_0) cells have reduced virus productivity or are latently infected, but have extended longevity, which make them the principal reservoirs and persistent viral production sites during the chronic phase of the infection (Pope and Haase, 2003). The reason for this persistence viral production and long-lived reservoir for the resting and quiescent states of CD4+ T cells is that these cells simultaneously avoide viral cytopathic effects and immune clearance mechanism (Stevenson, 2003). In the bsence of treatment, a gradual destruction of the naive and memory CD4+T-lymphocyte populations, and eventually diminishing of the immune system, advance HIV-1 infection further into AIDS phase (Fauci, 2003).

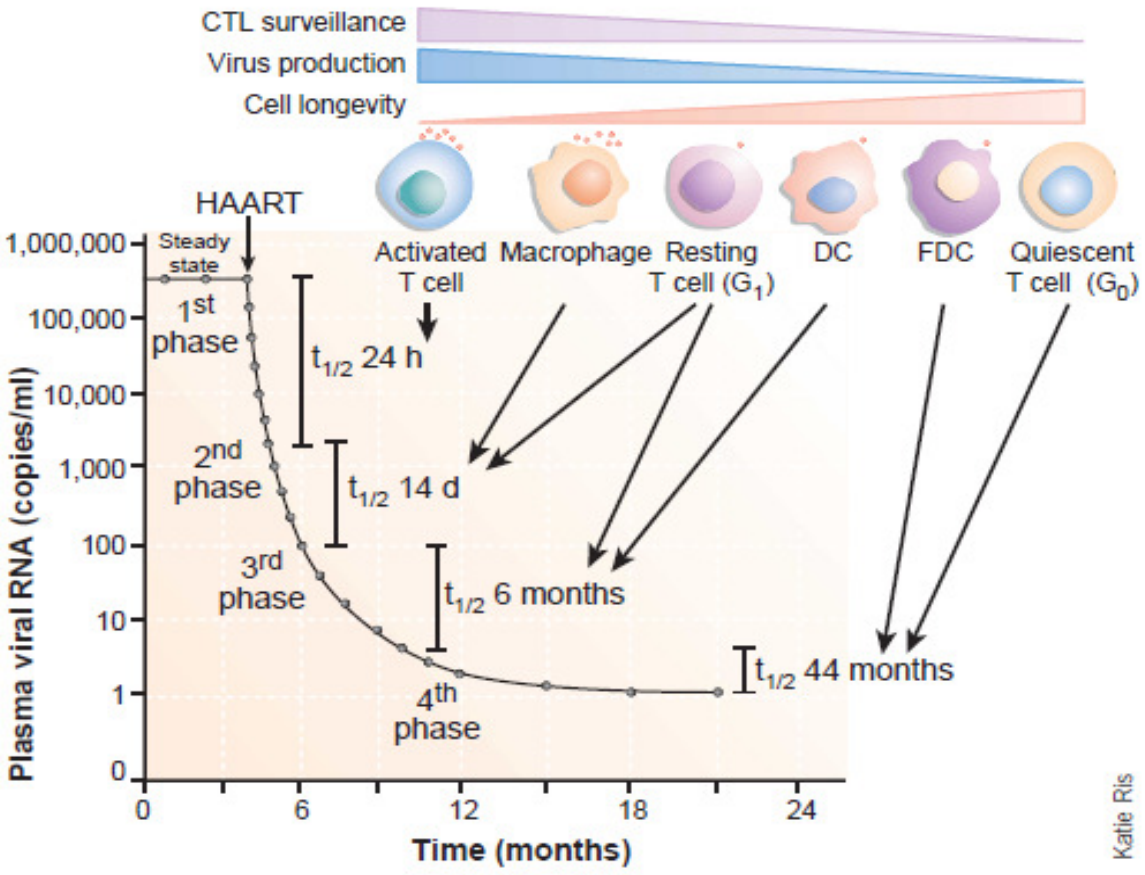


Figure 1.3.2.2 Viral reservoirs and their relative contribution to plasma viremia.
Steady-state levels of plasma viral RNA reflect the cumulative production of virus from the various cellular reservoirs and the turn over of virus producing cells in those reservoirs. 1st phase largely by activated CD4+ T cells; 2nd phase largely by Macrohages and resting T cels; 3rd phase, extremely low and chronic production from stable reservoirs (resting T cells and DC); 4th phase, FDCs harboring trapped virions and quicent T cells. The numbers pointed by the arrows indicate the lifespan of the respective cells, and so show the duration of their viral production [Adapted from Stevenson, 2003]

1.3.3 HIV Diagnosis, ARV Treatment and Drug Resistance

1.3.3.1 HIV diagnosis

HIV diagnosis is performed in a two-level procedures: the initial procedures are for screening purposes (screening assays) and are designed to detect all infected individuals, whereas the second ones are designed to be confirmatory (supplemental assays) that identify individuals who are not infected but who have reactive screening test results (Preiser and Korsman, 2007). While

screening tests possess a high degree of sensitivity with accompanied few false-negative results, confirmatory assays have a high specificity with accompanied few false-positive results (MacDonald *et al*., 1989). The sequential use of these two tests offers results that are highly accurate and reliable (CDC, 1989; Chou *et al*., 2005). HIV diagnosis can also be categorized into indirect and direct assays. The indirect assays demonstrate presence of humoral immune response markers (virus-specific antibodies) against HIV, which are found in virtually 100% of HIV-infected individuals, although negligible number of infected persons fail to produce detectable antibody (Novitsky *et al*., 2007). The direct assays on the other hand demonstrate presence of infectious virus using either of the following methods; cell culture, the viral antigen p24 by ELISA, or viral nucleic acid/viral genome (nucleic acid testing) by RT-PCR/PCR (Preiser and Korsman, 2007). The latter diagnostic tests are appropriate for suspected primary or vertically transmitted infections (Owen *et al*., 2008).

Several indirect and direct alternative assays are available to conduct HIV tests that can detect all potentially occurring virus types (HIV-1, HIV-2), groups (HIV-1-N, HIV-1-O, HIV-1-M) and group M subtypes (Owen *et al*., 2008). The most commonly used HIV screening tests involve Enzyme-Linked Immunosorbent Assays/Enzyme Immunoassays (ELISA/EIA), because these assays utilize relatively simple methodology with inherent high sensitivity and suitability for testing large numbers of samples (MacDonald *et al*., 1989). HIV ELISA/EIA assays have been developed progressively from first to the current fourth generation, based on the principle used in the assays as well as the type of antigens used, featuring improved performance with each subsequent generations (WHO, 2009). While first generation assays employed HIV "whole virus" antigen obtained from cell cultures to react with antibodies of patient serum, those from the second generation onwards used recombinant virus proteins or synthetic peptides representing

immunodominant epitopes (Preiser and Korsman, 2007). As all ELISA/EIA antibody assays are based on the principle of a specific antigen-antibody reaction, the assays may not detect the target antibodies during the 'window period', 6 to 12 weeks after infection in the case of early generation tests, or 3 to 4 weeks in the case of newer generation assays including third generation ELISA (Owen *et al*., 2008). Since recently, the 'window period' has been shortened further by several days using antigen detection tests and by several more days using nucleic acid detection methods (Owen *et al*., 2008). It may even be only 2 to 3 weeks if an all-inclusive testing strategy, involving both direct and indirect tests, is used such as combining nucleic acid detection assays and fourth generation ELISA tests that detect IgM antibody simultaneously with IgG, and demonstration of circulating p24 antigen (Owen *et al*., 2008).

Since screening tests are liable to give false-positive results due to their non-specific reactivity as the result of their high sensitivity, samples that produce repeatedly reactive results by screening tests must be further tested using a confirmatory test, which in most cases is a Western blot or an immunofluorescence assay (IFT or IFA) (Preiser and Korsman, 2007). In this case, purified and denatured viral proteins produced from cell culture-propagated HIV, are separated according to their molecular weight by electrophoresis, immobilized and blotted onto a nitrocellulose membrane which is then cut into strips and incubated with patient's serum (CDC, 1989). Upon binding of patient antibodies reactive to the viral proteins in the strip, bands are formed when an antihuman enzyme-labeled secondary antibody and matching substrate is used (CDC, 1989). The viral proteins used in this assay include the env or envelope glycoproteins (gp41, gp120, gp160), the gag or nuclear proteins (p18, p24/25, p55) and the pol or endonuclease-polymerase proteins (p34, p40, p52, p68) (Preiser and Korsman, 2007). Interpretation of the test is variable according to the country's' accepted guideline: the American Red Cross demands at least three bands, one

from each group (i.e. one gag, one pol and one env band); FDA demands the p24, the p34 as well as the gp41 or gp120/160 bands; WHO recommends only two env bands (CDC, 1989; Preiser and Korsman, 2007).

A different approach to these conventional testing methods is a recently introduced technique called point-of-care diagnoses or rapid tests. The rapid test strategy is superior to conventional testing in certain respects; namely, it is less time-consuming (taking only about 20 minutes), inexpensive with no requirement for special equipment, and being technically feasible in developing countries, offering access to screening for the vast majority of populations in these resource-poor settings (Reynolds *et al*., 2002). Yet, they have similar accuracy to the EIA tests, with the same sources of specimen (i.e., saliva, serum, plasma, and whole blood from finger sticks) (Chou *et al*., 2005). Like the EIAs, rapid test reactive specimens have to be tested with other confirmatory tests in order to definitely decide the serostatus of tested individuals (Greenwald *et al.,* 2006). The US Food and Drug Administration has approved four rapid tests OraQuick® (and its newer version OraQuick® Advance) Rapid HIV-1/2 Antibody Test (OraSure Technologies, Inc., Bethlehem, PA); Reveal™ (and its newer version Reveal™ G2) Rapid HIV-1 Antibody Test (MedMira, Halifax, Nova Scotia); Uni-Gold Recombigen® HIV Test (Trinity BioTech, Bray, Ireland); and Multispot HIV-1/HIV-2 Rapid Test (Bio-Rad Laboratories, Redmond, WA) (Greenwald *et al.,* 2006).

1.3.3.2 Highly Active Antiretroviral Therapy (HAART)

No evidence is thus far available for the presence of curing therapy to HIV infection, except for the one instance in which the patient received stem cell bone marrow transplant in Germany (Hutter *et al*., 2009), and remained sero-negative until now, 45 months since the transplant

(Allers *et al*., 2010). Nonetheless, the breakthrough made in 1996 in the introduction of combined antiretroviral treatment has helped millions to have a prolonged survival and improved quality of life as the therapy has substantially reduced the morbidity and mortality of HIV-infected individuals (Lihana *et al*., 2009a; Simon and Ho, 2003). The first approved anti-HIV agent in 1987 was zidovudine (De Clercq, 2009). Since then, about 25 antiretroviral drugs that fall into six classes have been approved for use against HIV (De Clercq, 2009). These drugs interact with different target biological events in the replication cycle of HIV and disrupt the completion of their processing. Two of the classes, Nucleoside/nucleotide Reverse Transcriptase Inhibitors (NRTIs/NtRTIs) and Non-Nucleotide Reverse Transcriptase Inhibitors (NNRTIs), target the process of reverse transcription of genomic RNA into cDNA by the enzyme reverse transcriptase (Sarafianos *et al*., 2009). The other four classes target the proteolytic cleavage of viral poly-proteins by the enzyme protease into their respective functional proteins (Protease Inhibitors [PIs]), viral entry through virus–cell fusion (Fusion Inhibitor [FIs]) and interaction of the virus with its (co-)receptors (Co-receptor Inhibitors [CRIs]), and integration of the proviral DNA into the host cell genome with the help of the enzyme integrase (Integrase Inhibitor [INIs]) (Sarafianos *et al*., 2009).

The NRTIs/ NtRTIs interact with the catalytic site or the substrate-binding site of the enzyme (De Clercq, 2009). The NRTIs/NtRTIs are 2’3’dideoxynucleoside analogues that should be phosphorylated into their respective dideoxyribonucleotide triphosphates (by undergoing 3 and 2 phosphorylaltion reactions, respectively, for nucleoside and nucleotide analogues) to be converted to their active form with the help of cellular kinases (Ravichandran *et al*., 2008). These are then incorporated into the newly formed cDNA as dideoxriboynucleoside monophosphate. Since NRTIs and NtRTIs lack a 3'-hydroxyl group on the deoxyribose moiety, unlike their

physiological counterpart deoxynucleotide substrates, the next incoming deoxynucleotide cannot form the next 5'-3' phosphodiester bond needed to extend the DNA chain, thereby terminating the chain elongation step of the reverse transcription process as competitive inhibitors to the natural deoxyribonucleotide triphosphates (De Clercq, 2009). Seven nucleoside RT inhibitors are currently in use globally: zidovudine (AZT); didanosine (ddI); zalcitabine (ddC); stavudine (d4T); lamivudine (3TC); abacavir (ABC); and emtricitabine (FTC) (Clavel and Hance, 2004; Thompson *et al*., 2010). Tenofovir disoproxil fumarate (TDF) is the prototype and the only Nucleotide Reverse Transcriptase Inhibitor in current use (De Clercq, 2009).

The NNRTIs, on the other hand, interact with an allosteric site located at a short distance away from the polymerase catalytic site of the enzyme reverse transcriptase (Clavel and Hance, 2004). Therefore, the mechanism of action of NNRTIs is different from that of NRTIs/NtRTIs; it disturbs the normal functioning of the reverse transcriptase by distorting the active (catalytic) site of the enzyme during the process of cDNA synthesis (Ravichandran *et al*., 2008). Efavirenz, Nevirapine, Etravirine, and Delavirdine are the four NNRTI drugs currently in use (Thompson *et al*., 2010).

Ten protease inhibitors (PIs) have been licensed for clinical use in the treatment of HIV infections: saquinavir, ritonavir, indinavir, nelfinavir, amprenavir, lopinavir, atazanavir, fosamprenavir, tipranavir and darunavir (De Clercq, 2009). PIs contain a hydroxyethylene scaffold ((-CH_2-CH(OH)) which mimics the normal peptide linkage, that would be cleaved by the HIV protease under normal circumstance, but which itself cannot be cleaved (De Clercq, 2009). They occupy the active site of the enzyme protease, which has a symmetrical homodimeric structure, and interfere with the proteolytic processing of precursor viral proteins into mature

viral proteins (Clavel and Hance, 2004). Other antiretroviral agents in current use include two entry inhibitors: Maraviroc, (also called chemokine receptor antagonist or CCR5 inhibitor that binds to CCR5 and prevent interaction of gp120 with the co-receptor), and Enfuvirtide (T-20) or also called fusion inhibitor that binds to gp41 thereby inhibiting virus-cell fusion] (De Clercq, 2009). In addition, another recently approved agent called Raltegravir was introduced into clinical use; it blocks provirus formation through preferentially inhibiting strand transfer by binding to the target DNA site of the enzyme integrase (Thompson *et al*., 2010).

It is now proven that each of the approved antiviral drugs can substantially reduce plasma HIV RNA a few months after initiation of antiretroviral therapy to below detection limit with a corresponding improvement in CD4+ T cell counts (Clavel and Hance, 2004). However, this successful treatment outcome sustains only for short period, as drug resistant variants appear shortly. Maximum benefit of the drugs is therefore obtained only by combining three or more drugs, so called Highly Active Antiretroviral Therapy (HAART) that has the potency sufficient to generate sustained suppression without outgrowth of resistant virus (Clavel and Hance, 2004). The ability of combination therapy to retard the rate of emergence of drug resistant variants stems from two properties of HIV drug resistance: first, occurrence of resistance for most of the drugs requires multiple mechanisms each with different mutations (Molla *et al*., 1996); and second, the naturally occurring error-prone rapid rate of HIV replication, that fuels the emergence and expansion of mutant viruses, is successfully suppressed by multiple drugs (Clavel and Hance, 2004). Thus, at present HAART is based on the administration of at least three different compounds, which jointly block viral activities required for completion of its replication (Sarafianos *et al*., 2009). Unfortunately, since HIV-1 infections cannot be cured even under intensive treatment with HAART, ARV drug therapy is a life-long therapy (Sarafianos *et al*.,

2009). The prominent obstacle to the eradication of HIV are the existence of persistent low-level viral replication, persistence of latently infected resting CD4+ T lymphocytes, and anatomical sanctuaries (for example, brain and testes) (Clavel and Hance, 2004).

Recognizing the pressing need to expand access to HIV treatment, WHO launched a very important program of reaching 3 million people in low and middle-income countries with ARVs by 2005, although this number was achieved only after 2007 (WHO, 2006). The '3 by 5' campaign was latter replaced by 'universal access to HIV treatment, prevention and care by 2010', where 80% of patients needing ARV treatment would get access to the treatment, but only few countries achieved the goal by the specified time (WHO/UNAIDS/UNICEF, 2010). Currently, WHO has raised a new motto called 'halting and beginning to reverse the spread of HIV/AIDS by 2015', and it is now included as part of the 'Millennium Development Goal 6' (UNAIDS, 2010). So far, 700,000 people from resource-rich and 5.2 million from resource-constrained countries have been on HAART, making the total people on treatment about 6 million (UNAIDS, 2010).

In order to coordinate and make the treatment efforts evidence-based, countries formulate ART guidelines and update them as new evidences are accumulated. To this end, the International AIDS Society-USA Panel has issued its 2010 updated guideline, which would be followed by the developed world. According to this guideline, ART should be initiated regardless of CD4 count for all symptomatic patients and those specific conditions and comorbidities such as increased risk of disease progression associated with a rapid decline in CD4 cell count (ie, >100 cells/μL per year) or a plasma HIV-1 RNA level > 100,000 copies/mL; older than 60 years; pregnancy (at least by the second trimester); or chronic HBVor HCV coinfection (Thompson *et al*., 2010). In

addition, all asymptomatic patients whose CD4+ counts is ≤500 cells/μL should also begin the therapy regardless of the presence or absence of symptoms (Thompson *et al*., 2010). The panel also recommends that the initial regimen be individualized according to resistance testing results and predicted virologic efficacy, toxicity and tolerability, pill burden, dosing frequency, drug-drug interactions, comorbidities, and patient and practitioner preference (Thompson *et al*., 2010).

Similarly, panel of experts at WHO have revised the 2006 ART guideline intended for use in resource-constrained countries to be implemented according to the countries' peculiar health system settings. This update was developed based on the following principles; do no harm, accessibility, quality of care, equity of access, efficiency in resource use, and sustainability (WHO, 2010). It is thus recommended for resource-constraint countries to start ART in all adolescents and adults at CD4+ T cell count of ≤350 cells/μL (unlike the previous recommendation of ≤200 cells/μL) with first-line regimen of one NNRTI and two NRTIs (one being either zidovudine (AZT) or tenofovir (TDF) (WHO, 2010). As the second-line regimen, the same guideline recommends a ritonavir-boosted protease inhibitor (PI) (preferably atazanavir (ATV/r) or lopinavir/ritonavir (LPV/r) and two NRTIs, one of which being AZT or TDF, whichever was not used in the first line (WHO, 2010).

1.3.3.3 Emergence and transmission of antiretroviral drug resistance

Despite administration of potent antiretroviral treatment, HIV is not completely eradicated from patients (Hatano *et al*., 2010). This is because replication-competent HIV-1 can be recovered from resting CD4+ T lymphocytes from most HAART-treated patients, and in most patients the latent reservoir remains stable with an estimated decay rate that ranging between 6 and 44 months even with prolonged treatment (Simon and Ho, 2003). Above all, HIV's high mutation

rate coupled with error-prone reverse transcription provide an opportunity for emergence of quasispecies that have acquired mutations conferring resistance to the drugs in use, and so abet the expansion of new variants in the presence of the drugs (Hupfeld and Efferth, 2009).

In fact reports on the last 25 years of HAART utilization in developed nations have shown that emergence of antiretroviral drug resistance with prolonged antiretroviral therapy is an inevitable event (Sarafianos *et al*., 2009). For example, by the year 2000, median of 28.5% of the prevalent HIV infections in San Francisco, where HAART started in 1996, were ARV-resistant (Blower *et al*., 2001). In another study in the United States, genotypic drug resistance assay revealed that the first three years of HAART treatment in the country witnessed a significant prevalence rate of ARV drug resistance among HIV-infected adults experiencing virological failure. The study showed that prevalence of any genotypic resistance to the three drug classes was estimated at 13%, while resistance to two classes and any resistance was 48% and 76% respectively (Richman *et al*., 2004). Similarly, a study in the UK found prevalence of multiclass resistance up to 15% and any single resistance up to 77% if single point estimate is considered (UK Collaborative Group on HIV Drug Resistance, 2005**).**

The concern over emergence of drug resistant HIV variants among treated individuals is that, apart from reducing treatment options and consequent death of the patients harboring the drug resistant variants (Pujades-Rodrıguez *et al*., 2010), the increasing prevalence of resistance to antiretroviral drugs has been associated with increased transmission of resistant viruses to newly infected individuals (Castelbranco *et al*., 2010). In this regard, it is reported that widespread use of ART in countries where HAART has been in use for longer time has resulted in an increased prevalence of transmitted drug resistant variants, ranging from 10% to 20% among drug naive

patients (Lihana *et al*., 2009b; Wheeler *et al*., 2010). Even in Africa, where HAART was introduced relatively recently, one study reported that resistance prevalence for drug naïve and drug-experienced persons were 5.2% and 16.0% respectively (Hamers *et al*., 2010).

No class of antiretroviral drugs has been spared from development of resistant HIV variants. In its 2009 update, the International AIDS Society-USA has recognized over 110 major and minor resistance mutations at over 70 positions in all classes of ARV drugs (Johnson *et al*., 2009). HIV Drug Resistance Databases of Stanford University, on the other hand, has listed over 90 mutations of all classes of drugs at over 57 sites in its November 2009 updates (http://hivdb.stanford.edu/pages/download/resistanceMutations_handout.pdf. accessed on January 04, 2011). Similarly, WHO has updated its list of Surveillance Drug Resistance Mutations (SRDM) in 2009, and has identified 93 mutations at 43 sites of the three classes of ARVs (Protease Inhibitors, Nucleoside/nucleotide Reverse Transcriptase Inhibitors, and Non-Nucleoside Reverse Transcriptase Inhibitor) made available to resource-constraint countries through universal access (Bennett *et al*., 2009).

Some of the NRTI mutations are capable of cross resistance, affecting more than one drugs in the class (E.g. 69 Insertion complex that affects all NRTIs currently approved by the US FDA [M41L, A62V, 69Inset, K70R, L210W, T215Y/F, K219Q/E], Q151 complex that affects all NRTIs currently approved by the US FDA except tenofovir [A62V, V75I, F77L, F116Y, D151M], and Thymidine analogue-Associated Mutations that affects all NRTIs currently approved by the US FDA [M41L, D67N, K70R, L210W, T215Y/F, K219Q/E]) (Cases-Gonzalez *et al*., 2007; Johnson *et al*., 2009). NNRTI resistance mutations are classified into four categories: primary NNRTI resistance mutations, secondary mutations, minor non-polymorphic mutations,

and polymorphic accessory mutations (Shafer and Schapiro, 2008). While primary mutations (K103N/S, V106A/M, Y181C/I/V, Y188L/C/H, and G190A/S/E) are among the first to develop during NNRTI therapy and cause high level resistance to one or more NNRTIs, secondary resistance mutations (L100I, K101P, P225H, F227L, M230L, and K238T; V179F, F227C, L234I, and L318F) are clinically relevant in selecting drug resistant NNRTIs when they occur, as often do, in combination with primary mutations (Shafer and Schapiro, 2008). On the other hand, minor non-polymorphic (such as A98G, K101E, V108I, and V179D/E) alone or in combination, and polymorphic NNRTI mutations (such as K101Q, I135T/M, V179I, and L283I) respectively cause consistent but low-level reductions in NNRTI susceptibility and modulate the effects of other NNRTI resistance mutations (Shafer and Schapiro, 2008). Recently, it has been demonstrated that cross resistance between two classes of ARV inhibitors (namely NRTI and NNRTI) occurs if mutations (like D549N, Q475A, and Y501A) occur at the connection subdomain (cn) and RNase H domain (rh) of HIV-1 reverse transcriptase (RT) (Nikolenko *et al*., 2010).

With the help of mutations selected by the drugs, the mutant variants alter their parts targeted by the drugs and circumvent the effect of antiretroviral drugs. They do this using various intriguing molecular mechanisms. For example, because NRTIs and NNRTIs target different binding sites of same polymerase moiety of reverse transcriptase and are using distinct inhibition mechanisms, the mechanism of resistance is specific for each class of inhibitor and leads to the selection of completely different sets of resistance mutations: NRTI-associated mutations are broadly distributed in the neighborhood of the nucleotide substrate binding active site but NNRTI-resistance mutations are concentrated in the hydrophobic allosteric site (Maga *et al*., 2010). Resistance to NRTIs is brought about by two mechanisms: by improving the enzyme's ability to

discriminate against and, hence, decrease incorporation of the inhibitor analogs; and by increasing the ability of the enzyme to remove 3'-terminal inhibitors from blocked DNA primers through phosphorylation reaction mediated by ATP or pyrophosphate (Cases-Gonzalez *et al*., 2007; Matamoros *et al*., 2009). Thymidine analog mutations and 69-Insertion-associated mutations work due to enhanced level excision activity of reverse transcriptase (Cases-Gonzalez *et al*., 2007); whereas M184V, non-thymidine analog-associated mutations such as K65R and L74V, and the multinucleoside resistance mutation Q151M act by decreasing NRTI incorporation (Clavel and Hance, 2004). Moreover, high-level cross-resistance to NNRTIs is conferred if any of the recognized resistance mutations occurs within the allosteric sites (Shafer and Shapiro, 2008). The dual class resistance conferred by D549N, Q475A, and Y501A mutants occurs as the result of reduced RNase H cleavage, which provides more time for the NNRTI to dissociate from the reverse transcriptase, and hence assist the enzyme to resume DNA synthesis (Nikolenko *et al*., 2010).

Resistance mutations to protease inhibitors result from amino acid substitutions that emerge either inside the substrate-binding domain of the enzyme or at distant sites (particularly gag cleavage site mutations), which directly or indirectly modify the number and the nature of points of contact between the inhibitors and the protease, thereby reducing their affinity for the enzyme (Clavel and Hance, 2004). In connection with integrase inhibitors (both FDA approved [raltegravir] and on clinical trial [elvitegravir]), most mutations that are located close to INI binding pocket decrease susceptibility by themselves while others compensate for the decreased fitness associated with other INI resistance mutations (Jegede *et al*., 2008). Resistance mutations in this class include major mutations N155H, Q148H/R/K, E92Q, the polymorphic mutations

L74M and G163R, T97A and V151I; and the non-polymorphic mutations L74R, E138A/K, Y143R/C/H, N155S, H183P, Y226D/F/H, S230R, and D232N (Canducci *et al*., 2009; Shafer and Shapiro, 2008). Likewise, the fusion inhibitor Enfuvirtide also selects for drug-associated resistance mutations both at its binding site HR1 (codons 36-45: G36**D/E/**V/S, I37V, V38**E/A/**M/G, Q40H, N42T, N43D/K/S, L44M, L45M), which offers direct resistance, and at HR2 region corresponding to the peptide sequence of enfuvirtide (N126K, N137K, and S138A), which improve fitness in combination with specific mutations within the 36-45 codones (Shafer and Shapiro, 2008).

1.4. Genomic Diversity of HIV

One of the most important features of HIV is its extensive genetic diversity and non-random global distribution of the various types, groups, and subtypes. The presence of two HIV types, HIV-1 and HIV-2, was known since the early period of the epidemic (Clavel *et al.*, 1986). Genomic analyses have shown that HIV-1 and HIV-2 share only 40% similarity in their RNA sequence (Levy, 2009). Both viral types have variations within themselves, and hence are categorized into groups, subtypes (or clades), sub-subtypes, circulating recombinant forms (CRFs), Unique recombinant forms (URFs) and geographically distinct lineages (Figure 1.4.1.1) (Robertson *et al*., 2000). Grouping of strains into these categories requires fulfilling certain criteria (Robertson *et al*., 2000), as depicted in Table 1.4.1.1. In general, inter-group sequence difference is over 35%, whereas inter-subtype difference can reach between 25 to 35%, while variation within subtypes ranges 15-20% depending on region of the genome examined (Taylor *et al*., 2008).

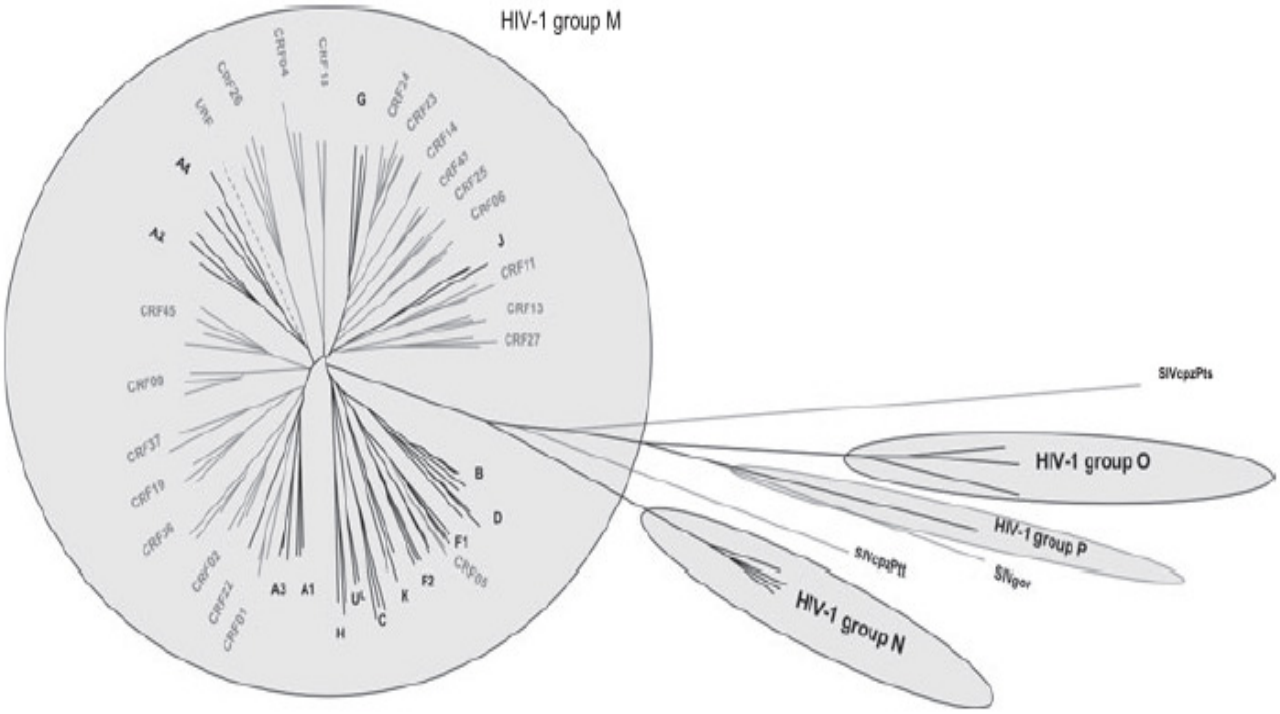


Figure 1.4.1.1 Phylogenetic tree of near full-length sequences representing the genetic diversity and linage of HIV.
(Adapted from Peeters *et al*., 2010)

HIV type 2 consists of eight different groups (A-H) without breaking on to subtypes, group A and B being epidemiologically most important (Levy, 2009). However, HIV type 1 is by far more diverse than HIV type 2, having four groups (M, N, O, and the recently discovered P) (Peeters *et al*., 2010; Simon *et al*., 1998). Of the HIV-1 groups, M is responsible for overwhelmingly majority of global infections (Osmanov *et al*., 2002). It is the most diversified of all HIV viruses, having nine subtypes including their various geographical lineages, A-D, F, H, and J (Robertson *et al*., 2000; Taylor *et al*., 2008); the original subtypes E, I and G have been identified as recombinant forms (Abecasis *et al*., 2007;Anderson *et al*., 2000; Gao *et al.*, 1998) and thus far 48 recombinant forms (losalamose HIV sequence Database, available at: http://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html; updated on September 29, 2010 and accessed on

December 28, 2010). Moreover, viral variants called 'quasispecies' from a single infected individual can differ up to 10% in their genomic sequence (Hemelaar *et al*., 2006).

Table 1.4.1.1 Phylogenetic classifications of HIV-1
(Adapted from Taylor *et al*., 2008).

Classification	Definition	Examples
Subtypes or clades	Genetically related HIV-1 strains that are essentially phylogenetically equidistant, generating a starlike, rather than a treelike, phylogeny	Subtypes A, B, C, D, F, G, H, J, and K are currently known; A through D are highly prevalent, others have low prevalence and limited geographic distributions
Sub-subtypes	Distinct lineages within a subtype; genetic distance between sub-subtypes is smaller than that between subtypes	Subtypes A and F are subdivided into sub-subtypes A1 through A4 and F1 and F2, respectively; mostly these circulate in Central and West Africa
Intersubtype recombinant forms	Mosaic strains with segments from two or more subtypes alternating across the genome	Common in mixed-subtype epidemics; thought to result from infection of a person with more than one HIV-1 subtype
Circulating recombinant forms	Specific recombinant forms that are spreading in a population; new forms are defined when three people without direct epidemiologic linkage are found to be infected; the assigned name reflects sequence of discovery and subtype composition, with "cpx" indicating forms containing three or more subtypes	Currently, 43 forms are described; CRF01_AE and CRF02_AG are found principally in Southeast Asia and West Africa, respectively; others have more limited distributions
Unique recombinant forms	Intersubtype recombinant forms recovered from only a single person	Hundreds of forms have been described on the basis of partial or complete genome sequences; their potential for epidemic spread is unknown
Geographically distinct lineages	Lineages, often country-specific, that are distinguishable phylogenetically; unlike sub-subtypes, they are not phylogenetically equidistant within subtypes	Thai B, Indian C, West vs. East African D, and Former Soviet Union A (FSU-A)

1.4.1 The Non-Random Global Distribution of HIV Types, Groups and Subtypes

The prevalence of the subtypes among infected individuals is disproportional and dynamically changing with time (Figure 1.4.1.2). Subtype C is the most predominating subtype of all, causing almost half of all global HIV infections followed by subtypes A and B in that oreder (Geretti, 2006; Hemelaar *et al*, 2006; and 2010; Osmanov *et al*., 2002). One of the interesting features of global subtype prevalence is its consistent relationship with the virus's mode of transmission.

Except for direct inoculation through parentheral route, subtype B is mostly prevalent among Intravenus Drug Users and homosexual men; subtype G is most prevalent among those infected nosocomially; whereas the other subtypes are prevalent among heterosexuals and bisexuals (Geretti, 2006). Moreover, the prevalence of recombinant forms is also increasing among infected individuals regardless of the routes of transmission. It is estimated that more than 10% of HIV-1 infections occur with circulating recombinant forms (CRFs) and unique recombination forms (URFs) in areas where multiple subtypes co-circulate (Peeters, 2000).

The other striking feature of HIV's diversification is that the global distribution of the various subtypes and recombinant forms is not random. Up until recently, the industrialized countries (North America, Western Europe, Australia, and Japan) have been dominated by subtype B; Latin America and the Caribbean by B and F; Asia and Eastern Europe by A, B, and C; central Asia and Pacific by A, B, and E; and South East Asia by B and E (Osmanov *et* al., 2002). However, the case in Africa has been different: virtually all subtypes and recombinant forms have been available in Africa from the outset of the epidemic, although regional differences were observed within the continent (Hemelaar *et al*, 2006).

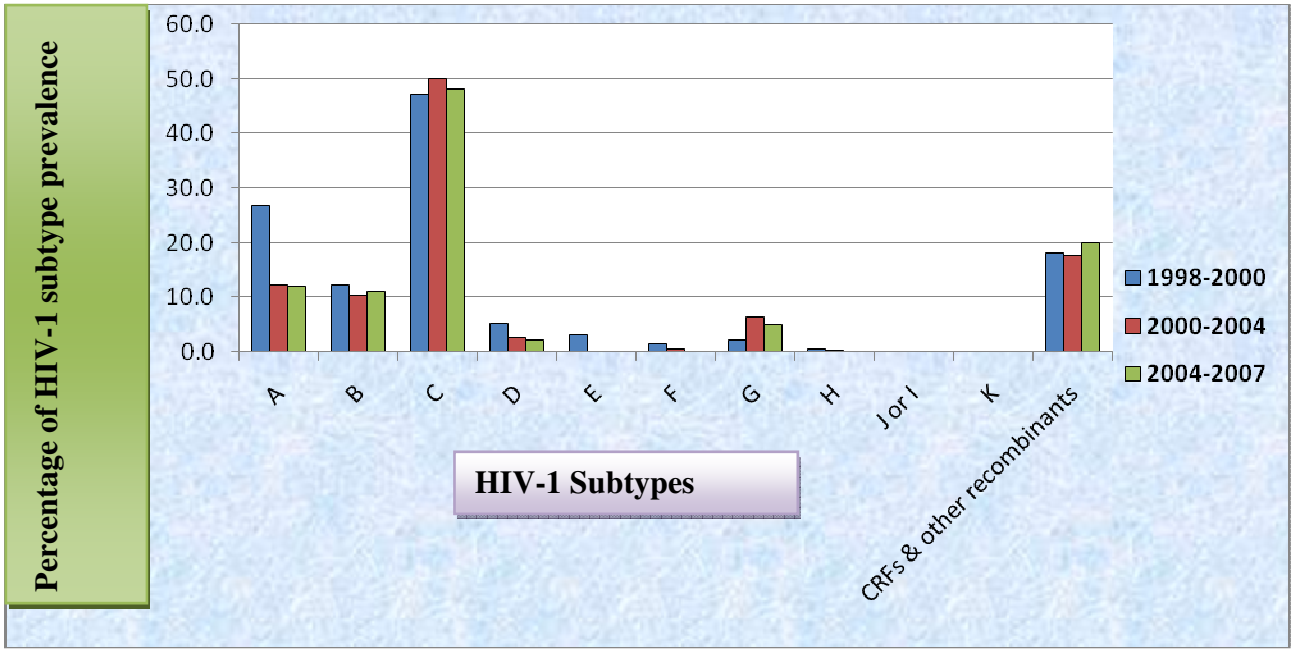


Figure 1.4.1.2 HIV-1 global subtype prevalence by years
Source of data for the figure: the 1998-2000 data, Osmanov *et al*., 2002, the 2000-2004 data, Hemelaar et al, 2006; and the 2004-2007 data, Hemelaar *et al*, 2010). Data for subtype E were available only for 1998-2000 perhaps because of the finding afterward that it is a recombinant between A and E. For subtypes F, H, J and K in the 2004-2007 study, their combined prevalence accounted to below 1%, and this estimate is not shown in this figure.

The current trend of emergence of new strains and the changed global distribution of subtypes show how quickly the virus is evolving and disseminating in the whole world (Figure 1.4.1.3). Most recent data from Hemelaar *et al.* (2010) show that the prevalence of CRFs (particularly that of CRF01_A/E and CRF02_AG) is increasing, while the the pure subtypes, except subtype C, are showing a decreasing trend, indicating importance of dual infection and recombination, (Hemelaar *et al,* 2010). As usual, subtype distribution in Africa is regional. While southern Africa is predominated by subtype C, Eastern Africa is the home for A, C and D; D is also the dominant subtype in northern Africa, while subtypes A, G, and the recombinant form CRF02_AG are the ones dominating in western Africa (Hemelaar *et al*., 2006). Central Africa, where the prevalence of HIV is only 5%, is characterized by the greatest diversity of HIV

subtypes: F, G, H, J, K, and CRF01_AE recombinant forms being the most notable (Djoko *et al*., 2010; Hemelaar *et al*., 2006; McCutchan, 2006).

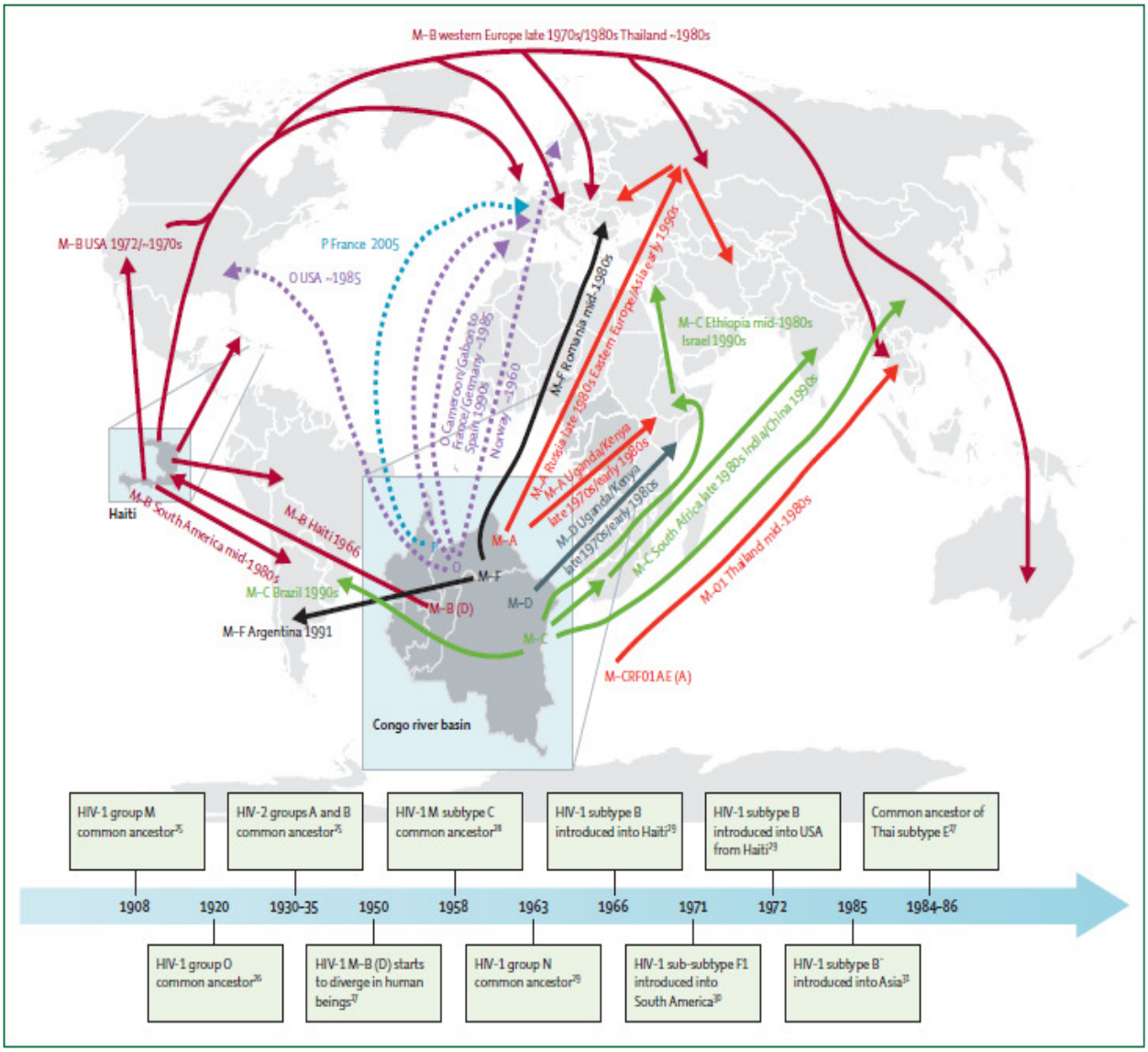


Figure 1.4.1.3 Estimated time line of global evolution and spread of HIV types, groups, and subtypes (Adapted from: Tebit and Arts, 2011). Enlarged parts of map show the main disease epicentres. The time line indicates the key events in the evolution of HIV-1 groups M, N, and O and of HIV-2. CRF=circulating recombinant form.

The fact that the recombinant forms are on the way of taking the upper hand in the global prevalence and widespread in distribution, implies the importance recombination in the viral evolution. This is resulted due to several intrinsic mechanisms employed by the virus. One is due to the inherently error-prone nature of the enzyme reverse transcriptase, which lacks proofreading activity that should ensure faithful transcription of the viral RNA genome into DNA (Holguın *et al*., 2008a). As the result of transcriptional error committed by reverse transcriptase, the new DNA transcript is conferred a mutation rate of approximately 3.4×10^{-5} mutations per base pair per replication cycle (Taylor *et al*., 2008). The effect of this change is aggravated by the high replication capability of HIV virus, 10^{10} virions per day, giving rise to millions of variants within an infected person (Simon and Ho, 2003). Viral diversity is further amplified when high mutation and replication rates are combined with recombination events, which take place as the result of a person's co-infection by two or more separate clades that are multiplying in the same cell (McCutchan, 2006; Robertson *et al*., 2000). In addition, selective pressure emanating from host immune responses (both humoral and cellular) and complex population dynamics have greatly contributed in intensifying the emergence and dispersal of new clades (Pond and Smith, 2009).

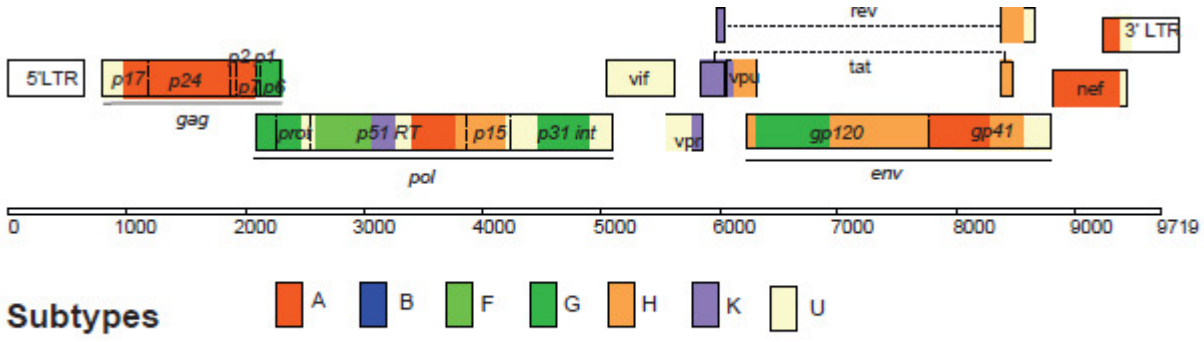


Figure 1.4.1.4 Recombinant virus CRF18_cpx: contains genetic region derived from 7 different HIV-1 subtypes
(Adapted from Levy, 2009)

In this way, several interclade (including second and third generation) recombination events took place in an increased frequency in the past. For example, the following recombinants were reported from various regions of the world: from Ghana, CRF06_cpx, sub-subtype A3, CRF09_cpx and subtypes G and D, and triple recombinants of CRF02_AG/A3/CRF06_cpx and CRF02_AG/A3/CRF09_cpx (Delgado et al., 2008); from Mali, CRF06_cpx and CRF09_cpx (Imamichi *et al*., 2009); from Cuba, CRF18_cpx (Thomson *et al*., 2005); from Argentina and Uruguay, CRF12_BF and CRF38_BF (Bello *et al*., 2010). Even intergroup recombinantion between M/O was observed very recently (Vessierea *et al*., 2010)]. So far, no evidence is obtained on natural recombination between HIV-1 and HIV-2, although this possibility was experimentally shown to take place in dually infected cells (Motomura *et al*., 2008). A very recent study has identified the first recombinant strain within HIV-2, and was coined as HIV-2CRF01_AB because its genome at the (larger portion of) 5' belongs to HIV-2 group B while the remaining 3' end belongs to HIV-2group A (Ibe *et al*., 2010). Figure 1.4.1.4 shows a genetic map of the recombinant virus CRF18_cpx that contains genomic coposition derived from 7 different HIV-1 subtypes.

1.4.2 Biological Significance of HIV Genetic Diversity

The biological implications of this formidably high level genetic variability of HIV-1 are well illustrated by their effects on HIV pathogenesis, transmission, diagnosis, treatment, and vaccine development, as have been substantiated by several investigations. Because of the relevance of viral diversity to the present study, the biological importance of HIV subtypes will be discussed in fairly greater detail in the coming subsections.

1.4.2.1 Impacts of HIV diversity on the pathogenicity of HIV infection

Disease progression in HIV-infected individuals is a multi-factorial process influenced by both virological and host factors, and hence varies from person to person. Among the virologcal factors attributed to play a great role in the determination of the clinical course of HIV infection is its extensive genetic diversity, which occurs virtually at all regions of the genome, and consequently its subtype variations. A number of cross sectional and few cohort studies have been conducted to investigate the role that subtype variations might play in the clinical course of HIV infection. For instance, it was determined early in the epidemic that HIV-2 infection is less severe and progresses more slowly with slower CD4 decline than HIV-1 infection (Alaeus *et al*., 1999; Matheron *et al*., 2003).

However, given the multitude of factors operating during HIV infection, it is extremely difficult to find consistent associations between HIV-1 subtype and correlates of transmission and pathogenesis. Not surprisingly though, the earliest studies reported conflicting results in this regard. Even among the subtypes in the group M, Hu *et al*. (1999) found a more rapid decline of CD4-T cell count in Caucasian French patients who were infected with subtype E than those infected with A, B, C, or D. Similarly, Kanki *et al.* (1999) followed the introduction and spread of HIV-1 subtypes A, C, D, and G in Senegal for a certain period and monitored the AIDS-free survival curves; the observation showed that individuals infected with a non-A subtype were 8 times more likely to develop AIDS than were those infected with subtype A. In the contrary, based on subtyping using *env* gene, Alaeus *et al*. (1999) found no difference in the rate of CD4 cell decline, clinical progression, or plasma HIV-1 RNA level between individuals infected with subtypes A, B, C or D over a mean observation period of 44 months.

However, contrary to those reports done early in the epidemic, several more recent studies have shown the importance of subtype differences in determining rate and severity of disease progression. The general agreement between these studies is that subtype D of HIV-1 is the most virulent strain of all, followed by recombinant ones with respect to rate of disease progression to AIDS defining illnesses, rate of CD4 cell decline, and increased plasma RNA load (Baeten *et al*., 2007; Kiwanuka *et al*., 2008; Senkaali *et al.* 2004). On the other end of this comparison is subtype A, which was reported to be approximately eight times less virulent than the subtypes B, C, D, G, various recombinant forms such as CRF01_AE, CRF02_AG and others (Kaleebu *et al*. 2002). Similarly, when subtypes D and A were compared, subtype D infection was associated with a 12-fold higher risk of death than subtype A infection, in spite of similar plasma HIV-1 loads (Baeten *et al*., 2007). In a comparative study of Env subtype B versus non-B subtypes (A, C, F-K, and recombinants like AC, AE, AG, BF, and DF) among Haitians, Canadians and Africans, Keller *et al*. (2009) found that Africans infected with non-B HIV clade had slower rates of disease progression compared with both Haitians and Canadians, with both groups being infected by the clade B virus.

1.4.2.1.1 Pathogenescity differences due to subtype differences in co-receptor usage

It can generally be understood that all the aforementioned differences in disease progression between Env subtypes could be associated with envelop phenotypes (syncytium-inducing Vs non-syncytium-inducing) of the infecting strain (Senkaali *et al*., 2004). For example, a study by Keller *et al*. (2009) has shown that individuals infected with HIV-1 variants with rapid/high (syncytium-inducing (SI) phenotype) progress more rapidly to AIDS than those who carry slow/low (non-syncytium-inducing (NSI)) HIV-1 variants. Such difference in envelope phenotype is largely an expression of differences in the use of the two co-receptors (CCR5 or

CXCR4). The use of CCR5 is linked to non-syncytium-inducing HIV strains (M-tropic, or R5 viruses), while CXCR4 use is linked with syncytium-inducing ones (T-tropic, or X4 viruses) (Tscherning *et al*., 1998). Thus, CCR5-tropic viruses are less pathogenic than CXCR4-tropic viruses; several investigators agree that the switch from CCR5 to CXCR4 is reported to cause faster CD4 cell decline and progression to AIDS (Abraha *et al*., 2009; Esbjörnsson *et al*., 2010). However, the generalization that R5 variants are non-syncytium-inducing viruses and hence less damaging holds true for only those in the early stage of primary infection; at late stage, these viruses acquire an increase in CCR5 affinity, with concomitant biological effects including heightened cythpatic activity against CD4+ T-cells, increased replication capacity, increased Env-mediated membrane fusion capacity, increased fusion kinetics, and increased bystander apoptosis of CD4+T-cells (Gorry *et al*., 2004). These authors argue that R5 M-Tropic HIV-1 variants are intrinsically cytopathic, but exhibit pathogenic effects that are distinct from those of X4 strains.

Discrepancies in the levels of M-tropism among primary R5 and X4 isolates could be due to differences in Env-receptor affinity at the different stages of the infection (Gorry *et al*., 2004). All HIV-1 subtypes can use both receptors at one time or another during the course of an individual's infection, but subtype D can most frequently use both co-receptors (dual tropic so called R5X4) right from the outset of the infection, unlike the other subtypes which begin CXCR4 usage only at late stage of disease progression (Tayor *et al*, 2008). But the percentage of CXCR4-using viruses in subtype C is lower even in patients with advanced AIDS than those in subtype B, (Abebe *et al*., 1999; Zhang *et al*., 2006), in which an estimated 50% of patients chronically subtype B infected individuals harbor X4 viruses (Esbjörnsson *et al*., 2010; Patel *et al*., 2008). Nevertheless, recent findings are indicating that the prevalence of X4 subtype C variants is also

increasing from time to time, especially at late-stage of diseases progression (Batra *et al*., 2000; Cilliers *et al*., 2003; Zhang *et al.,* 2010 and references therein), although some conflicting reports are seen in the literature where only 15% of X4 variants were observed among subtype C strains while 60-77% was observed for all other major subtypes (Esbjornsson *et al*., 2010).

It is generally accepted that amino acid composition of selected envelope regions determine co-receptor choice, which include variable regions of gp120, particularly the V3 loop, (Gorry *et al*., 2004; Hoffman *et al*., 2002; Mild *et al.*, 2010; Senkaali *et al*., 2004) to lesser extent V1/V2 (Gorry *et al*., 2004; Mild *et al.*, 2010), V4 (Zhang *et al*., 2010), and even gp41 (Huang *et al*., 2008). The specific amino acids most significantly involved here are those positive charged amino acids at either position 11 or 25 in the V3 region, which are strongly associated with CXCR4 use (Patel *et* al., 2008; Pastore *et al*., 2006). In addition, a net positive change in the V2 region and changes at position 440 in the C4 region of gp120 adjacent to the co-receptor binding site have also been linked to receptor choice (Gorry *et al*. 2004; Hoffman *et al*., 2002).

Differences in genetic composition at the V3 loop between subtype B and C have been documented recently. For example, Patel *et al*. (2008) have experimentally determined that the major structural differences between subtype B and C reside in the stem and turn regions of V3 loop as these two regions contain the majority of polymorphism, although more minor differences exist in the base residues 1 to 8 and 25 to 35. A very intriguing observation regarding emergence of dual tropic variants from previously R5 using strains was reported to have occurred among a subtype C-infected Zimbabwean woman (White *et al*., 2010). The unique feature of these variants is that the conversion from R5 to R5X4 occurred as a result of insertion of five amino acids within the V3 loop between amino acids 23 and 24 just after the crown region, which

are thought by the investigators to have originated from the human host genome rather than duplication within HIV-1 (White *et al*., 2010).

1.4.2.1.2 Pathogenescity differences associated with subtype-specific variation in structural and regulatory genes

Differences in pathogenecity between the various HIV-1 subtypes are resulted due not only to the diversity in the envelope region, but also due to the diversity in other structural, auxiliary and regulatory genes as well. The diversities in these regions are caused by differences in sequences of the Long Terminal Repeat (LTR), in response to transcriptional factors, and in genomic sequences for Nef, Tat, Rev, and VPU, which can, in one way or another, influence kinetics of viral replication, rate of transmission and ultimately disease progression (Gretti, 2006). The important HIV gene expression regulator sequences are located in the LTR, which contains abundant transcription binding sites that provide numerous opportunities for cellular and viral transcription factors to interact and regulate LTR activity under the specific settings of the infected cell (Krebs *et al*., 2001). LTR serves as a central point to converge and integrate the various transcriptional factors that serve as end-point effectors in signal transduction cascades and other regulatory activities; the activities of these factors regulate viral gene expression, production of progeny viruses, and spread of the infection (Krebs *et al*., 2001; Naghavi *et al*., 1999). In addition, LTR also regulates expression of viral gene products like Tat, Vpr, and gp120, which have crucial role on cellular functions, host cell viability, and intercellular communication through soluble mediators such as cytokines (Krebs *et al*., 2001).

Like the case in other genomic regions, there are subtype-specific sequence differences in the LTR region, with accompanied biological consequences. For instance, investigators have found that subtypes C, D and CRF01_AE all contained subtype-specific sequences in the negative

regulatory elements (NRE) (Naghavi *et al*., 1999). Turk *et al*. (2006) have also shown that the LTR-Tat complex derived from BF recombinants has higher transcriptional activity when compared to its subtype B counterpart. A difference in the number of certain sequences in the LTR can also have a profound effect in biological properties between subtypes. For example, it has been shown that the CRF01_AE LTR contains a single NF-kB site while LTRs from most other HIV-1 subtypes contain two or three NF-kB sites, indicating the possibility that LTR activities induced by NF-kB may differ between the CRF01_AE LTR and other LTRs (Rodenburg *et al*., 2001). The biological effect of this difference in the number of NF-kB on transient expression and viral promoter activity has been well-demonstrated. Transient expression analyses with subtype C LTR, which contains three NF-kB binding sites, has shown an increased activation as compared to LTRs containing only one or two NF-kB binding sites (Rodenburg *et al*., 2001). Similarly, measurement of a differential response to tumor necrosis factor alpha treatment and the induction level that followed among the different subtypes showed a positive correlation with the number of NF-kB sites in the respective LTRs (Jeeninga *et al*., 2000; Montano *et al*., 2000). Moreover, other investigators also demonstrated that subtype C viruses exhibited increased p24 (CA) levels and thus higher replication than viruses of other subtypes (Naghavi *et al*., 1999). Taken together, these experimental results and sequence analyses suggest that genetic diversity of HIV-1 LTR may result in differences regarding replication rate, scope of global dissemination and severity of disease progression.

Another gene that encodes for a regulatory protein is *tat*. This gene encodes for the protein Tat, which has several potent AIDS pathogenic effects. For example, Tat has been shown to induce neurotoxicity that would contribute to the development of HIV-associated dementia (King *et al*. 2006; Li *et al*., 2010). Apart from its central role in the activation of HIV gene expression,

several other biological effects of Tat have been documented including dysregulation of cytokines expression in immune cells, induction of apoptosis in neuronal cells, and interaction with other cellular factors to favor the servival of the infecting HIV-1; these processes develop conducive conditions within the infected individual for opportunistic infections (Campbell *et al*., 2010; Li *et al*., 2010).

Although all types, groups and subtypes of HIV viruses can produce Tat, the biological effects of Tat from these viruses are not of equal magnitude. For example, Desfosses *et al*. (2005) examined subtype-specific interactions between Tat, transactivation-responsive (TAR) element, and P-TEFb proteins and how these interactions may modulate the efficiency of HIV-1 transcription. The study revealed that Tat proteins derived from HIV-1 subtypes C and E were strong transactivators of LTR activity; Tat E also had a longer half-life than the other Tat proteins and interacted more efficiently with the stem-loop TAR element (Desfosses *et al*., 2005. Moreover chimeric Tat proteins harboring the Tat E activation domain were strong transactivators of LTR expression (Desfosses *et al*., 2005). Similarly, other studies have found that Tat proteins isolated from different HIV-1 subtypes including B and C exhibit differential effects on cytokine induction and other biological activities (Li *et al*., 2005; Yim *et al*., 2009). This conclusion is supported by experimental observations that Tat expressed in HIV subtype C was shown to be a more potent inducer of anti-inflammatory cytokines, such as IL-4 and IL-10 in primary human blood monocytes, compared to the subtype B counterpart (Gandhi *et al*., 2009), and reduce the production of TNF-α, CCL2 and decrease intracellular calcium influx (Campbell *et al*., 2010). However, subtype B Tat was demonstrated to be a potent neuropathogenic agent compared to the subtype C counterpart, through up-regulation of the rate-limiting enzyme of the kynurenine pathway (Campbell *et al*., 2010). In addition, Tat from subtype B could induce the

CXCR4 expression on resting CD4 cells and enhance the infectivity of HIV by using CXCR4 as a co-receptor compared to that of the subtype C Tat (Campbell *et al*., 2010).

Still a different regulatory gene exerting significant impact on HIV disease progression through its protein (Nef) is *nef*. Apart from being highly immunogenic in itself and so under host immune selective pressure, Nef is capable of modulating cell activation and signal transduction pathways in a way that cellular mechanisms are in favor of viral replication (Turk *et al*., 2009). Just like Vpu protein(Hussain *et al*., 2008), Nef is also involved in the down-modulation of cell surface expression of mature major histocompatibility complex class II (MHC-II), while up-regulating surface expression of the invariant chain (Ii) associated with immature MHC-II, a condition that leads to impaired CD4-T-helper-cell responses found in HIV-1-infected patients with progressive disease (Schindler *et al*., 2003; Stumptner-Cuvelette *et al*., 2001).

Sequence analysis studies on *nef* gene among various HIV-1 group M subtypes (including A, B, C, D, F, G, H, J, CRF02_AG, B/G/B, CRF02_AG/H, and D/G) have revealed that there is a significant amino acid variability between the variants (Parreira *et al*. 2005; Walker *et al*., 2007). This variability was found to be linked to the rate of disease progression. For example, Walker *et al*. (2007) compared *nef* sequences from individuals infected with HIV-1 subtype C viruses showing slow and rapid disease progression, and found that a good number of sequences from slow progressing variants had insertions and deletions exclusively in the N-terminal variable region, indicating the role of sequence diversity to alter the pathogenic potential even within the same subtype. A recent investigation by Turk *et al*. (2009) also underlined the significance of subtype difference in the variability of Nef biological properties: subtype C-derived Nef protein

was found to have slightly diminished capacity to downregulate MHC-I expression as compared to Nef derived from subtypes B, F, and CRF12-BF.

Taken together the literature has indicated unequivocally that the enormous genetic variability seen among HIV types, groups, and subtypes have resulted in enourmous differences in the pathogenicity and rate of disease progression seen in individuals infected by various subtypes. It should be noted that subtype D and CRFs, particularly CRF01_AE and CRF02_AG are more virulent with faster disease progression than the other subtypes.

1.4.2.2 Impacts of genetic diversity on drug resistance, virological and immunological responses to ARV treatment

CD4+ T-cell count, apart from being a marker of HIV disease progression and a criterion for initiation of antiretroviral therapy (ART), is used to asses success of antiretroviral therapy along with quantification of plasma HIV RNA (viral load). In the cases where treatment failure is observed by measuring these markers, HIV drug resistance assays are performed. However, there are considerable differences between the various subtypes with respect to their impact on virological (plasma viral load) and immunological (CD4+ T cell count) responses and drug resistance mutations.

1.4.2.2.1 HIV diversity and CD4+ T cell count

A recently conducted direct comparative study of CD4+ T cell count decline between subtype B infected and non-B infected drug-naïve individuals has shown that annual rate of absolute CD4+ T cell count and percentage CD4+ T cell decline was greater in subtype B infected individuals than in non-B infected individuals (Keller *et al*., 2009). Another comparative study conducted on the average annual rate of CD4+ T cell count among treatment-naïve Ugandans revealed that faster decline was observed among persons infected by subtype D than subtypes A,

recombinants, and multiple subtypes (Kiwanuka *et al*., 2010). This could be because of the deferential ability of subtype D to trigger apoptosis on activated CD4+ T cells, as has been demonstrated experimentally (Bousheri *et al*., 2009). A supportive evidence to the above observation of subtype D's severe impact on CD4+ count came from a study on individuals on ARV treatmen, which has reported that no significant differences in rate of CD4+ cell decline was detected in the initial response to highly active antiretroviral therapy for subtypes B, A, C and CRF02-AG; but a statistically significant four-fold faster rate of CD4+ decline (after adjustment for gender, ethnicity and baseline CD4+ count) was observed for subtype D (Easterbrook *et al*., 2010). However, report from study by Geretti *et al*. (2009) on patients under ARV treatment indicated absence of difference in CD4+ cell count recovery between subtypes A, B, C, D, and recombinant AG, although persons with subtype B infection showed higher baseline CD4+ cell counts and could maintain the advantage throughout therapy.

1.4.2.2.2 HIV diversity and plasma viral RNA level

Like the cases of CD4+ T cell count and responses to HAART, studies on the impact of HIV-1 subtype diversity on the level of viral RNA have produced contradictory reports. For example, Geretti *et al*. (2009) reported that, while viral load suppression on ARV treated patients occurred more rapidly in those infected with subtype C and subtype A than B, the hazard of virologic rebound was similar in all subtypes; and concluded that patients infected with prevalent non-B subtypes were as likely to achieve viral load suppression as persons infected with subtype B. Similarly, Easterbrook *et al*. (2010) reported no statistically significant differences regarding subsequent rate of virological rebound for subtypes B, A, C and CRF02-AG, with the exception of subtype D which showed a higher rate of virological rebound at six months compared with subtypes B, A and C. Conversely, a prospective cohort study in Thailand has documented that

median RNA viral levels at the earliest time within 3 months of seroconversion were more than three times higher for persons infected with subtype E than subtype B, although this difference decreased over time, being similar at 12, 18, and 24 months following seroconversion (Hu *et al*. 2001). Support to this observation comes from a study that assessed ARV-treatment response of African patients who harbored non-B subtypes against the response of European patients who harbored predominantly subtype B. In this study there was evidence of an increase in viral load after 9 months for the African group, resulting in a widening viral load gap between the two cohorts (Frater *et al*., 2002).

1.4.2.3. HIV diversity and ARV drug resistance mutation patterns

All ARV drugs of all classes have been exclusively developed in resource rich nations through molecular dynamics and rational drug design using template target proteins of subtype B (Soares, 2008). Moreover, data used to define drug resistance mutations have been generated almost entirely from subtype B HIV-1, despite the global dominance of non-B subtypes (Martinez-Cajas *et al*., 2008). Cognizant of the clinical implication of any difference in drug resistance mutations between subtype B and non-B, several studies have been carried out around the world, which unfortunately have produced conflicting results. While some studies have indicated that most patients infected with all HIV variants respond well to antiretroviral (ARV) therapy, and some of those failing therapy develop drug resistance mutations similar to those seen in subtype B (Frater *et al*., 2001; Garrido *et al*., 2008), others argue that there are emerging differences that have been identified in some subtypes (Abecasis *et al*., 2006; Grossman *et al*., 2001; Kantor, 2006). Nonetheless, given the inherent differences between subtypes in codon sequences at all regions including those at positions associated with drug resistance mutations, viruses of different subtypes are likely to encode different amino acid substitutions at these positions. It is thus

logical to hypothesize that genetic diversity might influence the type of emerging resistance mutations, the degree of cross-resistance to ARVs within a drug class, as well as the rate of emergence of resistance (Martinez-Cajas *et al.,* 2009).

A number of HIV-1 drug resistance pattern studies focusing on non-B subtypes in patients receiving antiretroviral therapy have indicated that presence of polymorphisms in these subtypes before commencement of therapy may provide a background for the emergence of subtype-specific pathways to acquired/major resistance (Tylor *et al*., 2008). In this regard, extensive natural codon polymorphisms have been well documented from nearly every subtype, each with its own signature codons at these sites. The extent of this diversity is so significant that up to 53% of protease and 48% of reverse transcriptase positions have been found polymorphic before treatment (Kantor, 2006). Literature review by Kanto and Katzenstein (2003) reveals that out of the 99 protease amino acids, 30 positions in subtype A (30%), 31 in B (31%), 26 in C (26%), 34 in D (34%), 33 in F (33%), 28 in G (28%), 14 in CRF01_AE (14%) and 35 in CRF021_AG (35%) are polymorphic in sequences from drug naïve persons. Similarly, of the first 240 RT amino acids, 41 positions in subtype A (17%), 45 in B (19%), 56 in C (23%), 43 in D (18%), 35 in F (15%), 32 in G (13%), 46 in CRF01_AE (19%) and 56 in CRF021_AG (23%) are polymorphic in sequences from drug naïve persons (Kanto and Katzenstein, 2003).

Some of these amino acid substitutions occur at high rates in non-B viruses at positions known to be associated with drug resistance in subtype B; these include protease positions 10, 20, 36, 63, 71, 77 and 93, and RT positions 69, 75, 98, 106, 118 and 179 (Kantor and Katzenstein, 2004). Moreover, systematic literature reviews by Martinez-Cajas *et al.* (2009 and 2008) have identified

patterns of specific polymorphisms at drug resistance sites among the various subtypes: (i) the non-nucleoside reverse transcriptase inhibitor resistance mutation, V106M, has been seen in subtype C and CRF01_AE, but not in subtype B, (ii) the protease inhibitor mutations L89I/V have been reported in C, F and G subtypes, but not in B, (iii) a nelfinavir selected non-D30N containing pathway predominats in CRF01_AE and CRF02_AG, while the emergence of D30N is favored in subtypes B and D, (iv) studies on thymidine analog-treated subtype C infections from South Africa, Botswana and Malawi have reported a higher frequency of the K65R resistance mutation than that typically seen with subtype B. Additionally, some substitutions that seem to impact non-B viruses differentially are: reverse transcriptase mutations G196E, A98G/S, and V75M; and protease mutations M89I/V and I93L (Martinez-Cajas *et al.,* 2009).

Most of these naturally occurring polymorphisms have been considered to be secondary drug resistance mutations, which compensate for replicative fitness compromised due to the presence of primary mutation in the presence of ARV drugs (Chaplin *et al*., 2010). Besides influencing the baseline susceptibility to ARVs, natural polymorphisms also impact on the genetic barrier for acquisition of ARV resistance, the establishment of different pathways to resistance acquisition, the phenotypic modulation of acquired resistance, and ultimately on the clinical response to current and future ARVs (Martínez-Cajas *et al.*, 2008; Soares MA. 2008). For instance, HIV type 2 and group O strains of HIV-1 possess intrinsic resistance to nonnucleoside reverse-transcriptase inhibitors (Taylor *et al*., 2008) due to the presence of mutations at positions linked to NNRTI resistance (181I, 188L and 190A), which are commonly found in drug-naïve individuals (Colson *et al*., 2004). Moreover, *in vitro* phenotypic drug resistance assays have shown that some secondary protease mutations alone confer reduced susceptibility to protease inhibitors without the occurrence of accompanied primary drug resistance mutations (Clemente *et al*., 2006; Parkin *et*

al., 2005). This kind of resistance conferred by secondary mutations has been observed from atazanavir among subtype C isolates (Fleury *et al*., 2006), and ritonavir**,** amprenavir and nelfinavir resistance in subtype G (Holguin *et al*., 2004).

Similarly, reduced susceptibility to the nucleotide/nucleoside reverse transcriptase inhibitors were detected for abacavir in some CRF02_AG isolates, which was related to the mutations D123N and I135T in the RT; and for zidovudine, didanosine and lamivudine in subtype D (Palmer *et al*., 1998). A subtype D isolate from Ugandan patients showed high-level resistance to the non-nucleotide inhibitors nevirapine and delavirdine as well as low level cross-resistance to efavirenz because of I135L and/or V245T mutations (Gao *et al*., 2004). In another study, the effect of the protease mutations K20R and M36I, both highly polymorphic in non-B subtypes, on viral replicative capacity and drug susceptibility was investigated using clones with non-B protease region but having subtype B backbone (Holguin *et al*., 2006). The study made clear that pretherapy clones harboring M36I replicated more rapidly than wild type or K20R/M36I double clones, showing its ability to provide replicative advantage to the non-B subtypes. Further more, mutated clones, which were all sensitive prior to drug exposure, had better replicative capacity under drug pressure with reduced susceptibility to some protease inhibitors, suggesting that additional compensating protease substitutions might occur in response to exposure to protease inhibitors (Holguin *et al*., 2006).

Studies on differences between subtypes in regards to acquired drug resistance mutation pattern from ARV-treated individuals have shown mixed results. Whereas the distribution and frequency of most mutations leading to resistance appear to be similar among subtypes (Chaplin *et al*., 2010), certain mutations seem to occur more frequently in a certain subtype than others (Tylor *et*

al., 2008). For example, two NRTI mutations (M41L and L210W) and three NNRTI mutations (V90I, A98G, V106I) were found in a recent study to be significantly associated with an array of non-B West African subtypes (Chaplin *et al*., 2010). Comparison of acquired resistance mutations with the function of subtypes revealed that CRF02_AG was less likely to have the M41L mutation compared to G, CRF06_cpx, A, and other subtypes or recombinant sequences (Chaplin *et al*., 2010). The same authors further detected that subtype A patients showed a 42.5-fold increased risk for the L210W mutation; and subtype G patients had an increased risk for A98G and V106I, while subtype CRF02_AG patients had an increased risk for V90I but a decreased risk for A98G.

The interesting phenomenon in regards to impact of subtype diversity on acquired drug resistance is that subtypes can behave differently in their susceptibility to ARV drugs even when they harbor the same resistance mutations at that particular site. For instance the major protease inhibitor mutations M46V, I54V, V82A and L90M confer increased resistance levels to nelfinavir and lopinavir in subtype C compared to subtype B (Gonzalez *et al*., 2006). In addition, the role of protease mutation I93L in subtype B and C is contradictory: while it is a minor drug-resistance mutation in B (Johnson *et al*., 2008a), its presence in subtype C renders the virus hypersusceptibility to lopinavir (Gonzalez *et al*., 2003). Similarly, comparison of the susceptibility of subtype B and C HIV-1 integrase enzymes harboring the mutations E92Q and N155H has shown that subtype C integrase enzymes bearing the combined resistance mutations E92Q/N155H were approximately 10-fold more susceptible to each of raltegravir and elvitegravir, than were subtype B recombinant integrase containing the same mutations (Bar-Magen *et al.*, 2010).

Conversely, resistance against the same drug can be conferred by unrelated mutations (alternative pathways). Drug-resistance to the protease inhibitor nelfinavir is conferred by two different routes: in subtype B the mutation D30N is preferentially selected in the presence of the drug, although the mutation L90M is also encountered, but in the case of non-B subtype resistance to this drug is conferred exclusively by L90M (Johnson *et al*., 2008a; Soars, 2008). In addition, D30N of B confers resistance to nelfinavir only, while L90M of non-B confers resistance to most protease inhibitors currently in clinical use (Johnson *et al*., 2008a; Soars, 2008). In the same way, tipranavir resistance in subtype G is conferred by mutations I82T and I82M, but the same resistance in CRF01_AE is acquired as the result of the mutation I82F (Abecasis *et al*., 2006; Clemente *et al*., 2006).

An important concept that must be borne in mind in relation to the influence of HIV genetic background is that rate of development of ARV resistance associated mutations depends on many factors, like the maintenance of protein function, replicative fitness advantage in the presence of drug, and ease of the mutational change (Pond and smith, 2009). With respect to maintenance of protein function and replicative fitness advantage after mutation, many studies have investigated the target enzymes under experimental conditions. In this regard, it was experimentally shown that protease from subtype A and C are inhibited by indinavir, ritonavir, saquinavir, and nelfinavir with 2.5–7-fold and 2–4.5-fold respectively weaker than that from subtype B (Velazquez-Campoy *et al*., 2001). The same investigators have found that the C subtype protease has highest efficacy (4–11 fold higher than the B subtype), while the A subtype protease exhibits values ranging between 1.5 and 5, suggesting a higher biochemical fitness of the A and C proteases in the presence of existing inhibitors. Whereas the protease mutations V82 and I84 have notable effects on the catalytic performance of non-B subtypes, characterization of these

mutations in a South African HIV-1 subtype C (C-SA) has indicated that V82A single mutation and V82/I84V double mutations had no significant impact on the proteolytic functioning of the enzyme, but the binding affinities of and inhibition by saquinavir, ritonavir, indinavir, and nelfinavir were weaker for each variant than for the wild-type protease, with the double mutant exhibiting the most dramatic change (Mosebi *et al*., 2008).

The ease with which the mutation occurs is also important factor to determine the rate of occurrence of a certain drug resistance mutations. For example, it has been noted that a nonnucleoside reverse transcriptase resistance associated mutation V106M occurs in subtype C following a single nucleotide change only, whereas the mutation in subtype B is achieved after substitutions of at least two nucleotides (Campbell, 2006). Similalrly, the protease mutation V82S is achieved in subtype B after substitution of GTC to ACC requiring the replacement of first two nucleotides GT with AC; whereas the same mutation in G and CRF14 is obtatined by a single nucleotide change from the naturally polymorphic ATC to AGC (Pond and Smith 2009). Many more of such substitution requirements have been documented by investigagtors and are available in the literature.

1.4.2.4 Inpacts of HIV diversity on the performance of serological diagnosis, viral load assays, and ARV drug resistance tests

1.4.2.4.1 HIV diversity and HIV screening/diagnosis

Initial HIV screening and diagnostic tests that use ELISA antibody, Western blot assays, and rapid antibody tests can all detect either one or both types of HIV (Koch *et al*., 2001). US Center for Disease Control and Prevention (CDC) affirms that current HIV-1 EIAs can accurately identify infections with nearly all non-B subtypes and many infections with group O HIV subtypes (Urnovitz *et al*., 1997). Similarly, rapid HIV-1 tests are claimed to be capable of

detecting all the major subtypes of group M (Urnovitz *et al*., 1997). The immunoassays designed to detect antibodies specific for HIV-1 are those targeting antibodies mounted against the env gp41 immunodominant region (Brennan *et al*., 2006). Considering the occurrence of extensive natural polymorphisms in all regions of HIV-1, it is possible to encounter variants that have modified or that have eliminate key epitopes targeted by these assays as the result of which reduced sensitivity or total absence of detection of the antibodies might happen during HIV screening.

In fac, a number of cases have been documented where sensitivity of the commonly used ELISA tests for detection of seroconversion have been reduced or lost, especially for non-B subtypes including groups other than M (Lee *et al*., 2007; Brennan *et al*., 2006; Simon *et al*., 1998). For instance, assay evaluation studies have found that some commercial assays were unable to detect group O infections (Loussert-Ajaka *et al*., 1994; Gurtler *et al*., 1995). Moreover, a group of investigators have shown that a fourth generation assay that was supposed to detect antibodies of all subtypes failed to detect a subtype B isolate having a natural polymorphism at the immunodominant region, indicating that even variability within the subtype for which the assay was developed can hinder sensitivity of the assays (Gaudy *et al*., 2004). This problem is more pronounced during running tests for treatment monitoring: viral load tests for quantification of the virus in the blood, and genotypic drug resistance assay for monitoring emergence of drug resistance mutations against the drugs on treatment.

1.4.2.2.2 HIV diversity and plasma HIV viral load tests

The various viral load assays that have been in use were developed and optimized for monitoring the subtype most prevalent in resource-rich countries, that is subtype B. They use mainly

sequence-based detection mechanism, which can be influenced by the tremendous sequence variabilities seen in HIV. Observational and experimental studies have indicated that this is indeed a problem when the assay are to be used in monitoring plasma RNA level from non-B subtypes, due to either failure to quantify/underquantify (Baldrich *et al*., 2001; Damond *et al*., 2007; Drexler *et al*., 2007; Geelen *et al*., 2003; Rouet *et al*., 2007; von Truchsess *et al*., 2006) or discrepancies of test results between various assays, particularly tests on recombinant strains (Holguı´n *et al*., 2008b; von Truchsess *et al*., 2006). For example, Jenny-Avital and Beatrice (2001) have reported cases where a US FDA-approved RT-PCR-based quantifying commercial assay (Roche Amplicor MONITOR version 1.0) repeatedly showed undetectable or low plasma viral RNA in samples containing subtypes A and E, while a branched-chain deoxyribonucleic acid assay (bDNA Quantiplex 3.0, Bayer Diagnostics) was found to be substantially higher. Similarly, earlier viral load test comparison study using the Roche HIV monitor assay and the Organon Nucleic Acid Sequence-Based Amplification (NASBA) HIV-1 RNA quantitative assay showed failure of the two assays to detect subtype A for 56% and 44% of the samples, respectively (Alaeus *et al*., 1997). Same descripancy was observed when performance of NucliSens EasyQ was compared to Amplicor HIV-1 Monitor Assay version 1.5 in 79 samples of subtype C-infected patients originating from Ethiopia; the disparity between the results of the two viral load assays was highly significant in subtype C samples such that in the vast majority, higher values of viral load were obtained by the Amplicor assay, while no differences between the two assays were found in subtype B samples (Gottesman *et al*., 2006).

In another comparative evaluation of two commercial assays (Amplicor v1.5; Roche Diagnostic Systems, Alameda, CA or Versant v3.0; Bayer Diagnostics, Emeryville, CA) and one in-house-built assay (Agence Nationale de Recherches sur le SIDA second-generation (G2) real-time

reverse transcriptase polymerase chain reaction (RT-PCR) test), approximately 30% of specimens harboring non-B subtypes were underquantified by at least 0.51 log10 in Amplicor v1.5 versus 5% underquantified in G2 real-time PCR (Rouet *et al*., 2007). It was further observed that discrepant results were also obtained even with subtype B samples (14% underquantified by Amplicor v1.5 vs. 7% by G2 realtime PCR). Similar percentages were observed when comparing results obtained with the G2 real-time PCR assay to those obtained using the Versant assay Rouet *et al*., 2007). Swanson *et al*. (2001) evaluated three commercial viral load assays and found that quantification of subtypes in group M was highly correlated in all the three assays but only one of them could quantify strains from group O. Similarly, another evaluation study of four commercial assays has shown that only one of the four assays mentioned in he previous study was able to quantify group O strains, while group M strains were quantified by all the assays tested, albeit at varied degrees (Swanson *et al*., 2005).

Recognizing the problems associated with inconsistency of viral load results in samples harboring non-B subtypes, commercial assays have developed tests with improved sensitivity and reproducibility across all subtypes, taking viral genetic diversity into consideration during their development (Reveiwed in de Mendoza and Soriano, 2009). Accordingly, four commercial viral-load assays are available with better performance record compared to their earlier versions: COBAS® TaqMan® (Roche; Pleasanton, CA, USA); NucliSENS EasyQ® HIV-1 v2.0 (bioMérieux; Boxtel, Netherlands); RealTime™ HIV-1 m2000rt (Abbott; North Chicago, IL, USA); and VERSANT® HIV-1 RNA 1.0 kPCR (Siemens; Berkeley, CA, USA). But their high cost and their requirement for a basic laboratory infrastructure that is not usually available in many resource-poor countries are thought to be hindrance to their wide use (de Mendoza and Soriano, 2009).

1.4.2.3.3 HIV diversity and ARV drug resistance assays and interpretation problems

Prior to the introduction of integrase inhibitors, most genotypic drug resistance assays detected HIV drug resistance by amplifying and sequencing the protease (PR) and reverse transcriptase (RT) regions of the *pol* gene. For this purpose, the assays used primers that target conserved sequences in these regions. Two FDA approved commercial assays are currently in use: ViroSeq™ HIV-1 Genotyping Version 2.7 (Applied Biosystems, Foster City, CA, USA) and TruGene (Siemens, Deerfield, IL USA) (Wallis *et al*., 2010). However, for the same reason that the viral load assays failed, the performance of genotyping assays has also been challenged by the diversity of HIV-1 subtypes at the protease and reverse transcriptase regions. Several evaluation studies on the performances of these assays, particularly their ability to successfully genotype subtypes other than B, have produced variable reports. While some studies reported that the two assays could genotype non-B subtypes as well as the B, (Eshleman *et al*., 2005b;Jagodzinski *et al*., 2003), others indicated that the two could not produce useable sequences for genotyping with equal accuracy for non-B subtypes as it is for B, and there were discrepancies in genotyping between those successfully sequenced samples, although the descripancies appeared to be random and did not affect interpretation of the major resistance codons. (Beddows *et al*., 2003). Similar dissatisfactory results were reported on the different subtypes in which the performance of the entire set of primers in the ViroSeq showed a significant decrease in the positive results for subtypes A, G and the recombinants, whereas a tendency to less positive results could be detected for subtypes CRF12_BF, D, H and J (Maes *et al*., 2004).

In order to minimize the influence of genetic variability at the target regions of drug resistance genotyping, a number of research groups have turned to developing their own in-house genotyping assays. For example, Wallis *et al*. (2010) developed and validated a new in-house

assay that targeted the unique subtype C circulating in South Africa by comparing sequence data and drug resistance profiles generated by the assay from 90 patient and 10 external quality control samples to data generated by the ViroSeq^{TM} HIV-1 Genotyping kit. The investigators found that the in-house assay was more efficient in amplifying all 100 samples, compared to 91 samples amplified by using Viroseq. Moreover, they found the in-house sequences to be 99.2% homologous to the ViroSeq sequences, and identical drug resistance mutation profiles were observed in 96 of samples tested (Wallis *et al*., 2010).

The inability of various commercial assays to correctly detect genotypic drug resistance mutations is not the only problem associated with identification of drug resistance among the different subtypes. After running the genotypic drug resistance assay, the sequences containing resistance mutations have to be interpreted. This may be done manually but the manual interpretation is difficult because a large number of protease and reverse transcriptase drug resistance mutations interact and emerge in complex patterns. To circumvent this problem several rules-based algorithms have been developed to assess HIV-1 susceptibility to available antiretroviral drugs. However, the rules are mostly derived from data on subtype B infected patients without considering inter-subtype genetic diversity (Soares, 2008). For this reason, several studies have found marked discordance between the interpretation algorithms, the clinical implication of which is tremendous. Kijak *et al*. (2003) assessed the concordance on the interpretation of HIV-1 drug resisance genotypic data by Stanford University database (SU), TreueGene (Visible Genetics (VG), and Virtual Phenotype (Virco, Belgium (VP) and found complete concordance between the three algorithms on only 13.7%, with major discrepancies (high level resistance interpretation by one algorithm and sensitive interpretation by another)

being observed between VG and VP in over 10% of the cases for didanosin, zalcitabine, stavudine and abacavir.

Similarly, Ravela *et al*. (2003) compared the concordance of four publicly available resistance interpretation algorithms: Agence Nationale de Recherches sur le SIDA (ANRS); HIV RT and Protease Sequence Database (HIVDB) (http://hivdb.stanford.edu); Rega Institute version 5.5 (Rega-5.5); and Visible Genetics version 6 (Toronto, Ontario, Canada) (VGI-6). These investigators reported that out of 30,675 interpretations they assessed, 4.4% were completely discordant, with at least one algorithm assigning a sensitive and another a resistance; 29.2% were partially discordant, with at least one algorithm assigning a sensitive and another an intermediate, or at least one algorithm assigning as an intermeiate and another as a resistance; and 66.4% displayed complete concordance, with all four algorithms assigning the same interpretation. Furthermore, in another recently conducted study of interpretation algorithms, results from Stanford University HIV drug resistance database (SHDB) were compared to the results obtained from Geno2-Pheno (G2P) and DR_Seqan (DS) on protease and reverse transcriptase sequences of non-B subtypes submitted to the Stanford database (Kandathil *et al*., 2009). This study reveals that concordance between SHDB and G2P was 85.6% while that between SHDB and DC was 37%.

1.4.2.4 Impacts of HIV diversity on transmissibility, replication rate and viral fitness

Incriminating a specific factor for observed differences in transmission rate HIV among HIV-infected individuals is difficult because several determinants interplay during the transmission process, including the prevalence and distribution of different risk behaviors in the population, the prevalence of co-factors for the sexual transmission of HIV, as well as non-immune and

immune host factors (Hu *et al*., 1999). Nonetheless, a number of prospective epidemiological studies have shown the importance of subtypes in transmissibility differences. For example, earlier studies have already recognized that HIV-2 is less transmissible both sexually and vertically compared to HIV-1 (Andreasson *et al*., 1993; Kanki *et al*., 1994). A prospective study by Hawes *et al*. (2008), in which cervico-vaginal specimens were taken at 3-day intervals over a 6-week period, has demonstrated significantly lower rate of viral shedding and lower levels of HIV-2 RNA in the female genital tract than HIV-1 RNA. However, the authors stressed that shedding correlated with plasma viral load irrespective of virus type.

Among the HIV-1 group M subtypes, subtype C has been found by many investigators to be more transmissible heterosexually and vertically than subtypes A and B (Graham, 2007; Iversen *et al*., 2005; Renjifo *et al*., 2001; Renjufi *et al*., 2004). Using mucosal shedding of HIV as surrogate marker of infectivity, John-Stewart *et al*. (2005) have identified differences in the shading of HIV-1 between subtypes A, C, and D among Kenyan pregnant women. The investigators observed that pregnant women infected with subtype C were significantly more likely to shed HIV-1–infected vaginal cells than were those infected with subtype A or D, before and after adjusting for age, CD4 cell count, and plasma HIV-1 RNA load. The same pattern was also observed regarding shedding in breast milk (John-Stewart *et al*., 2005). However, when mother-to-child transmission was investigated directly from LTR-sequenced HIV-1 from Tanzanian women who transmitted HIV to their children, transmission rate was higher in the order of A, A/C recombinant and C, although the difference was not statistically significant, but no transmission of subtype D (Tapia *et al*., 2003). Yet, a contrary result was documented among Kenyan women who transmitted HIV vertically to their new borne, in which subtype D (as analyzed by p24 gag and gp41) was found more transmissible than subtype A (Yang *et al*., 2003).

When heterosexual transmissibility of subtype A and D were compared, after adjusting for index HIV-positive partners' age, viral load, stage of disease, genital ulcer disease, and HIV-negative partners' genital ulcer disease and nonuse of condoms, subtype A viruses were found associated with a higher rate of transmission than subtype D (Kiwanuka *et al*., 2009).

The fact that subtype C is in ever increasing prevalence even in places where other subtypes, such as A and D, have been predominating prompted Walter *et al*. (2009) to hypothesize that there could be biological factors favoring transmission of subtype C. With the aim of proving evidence for this speculation, these researchers underwent experimental investigations on chimeric viruses using V1–V3 envelope fragments from a subtype-A/C dually infected woman with preferential genital replication of subtype C. They evaluated viral adaptation, spread and cell fusion ability *in vitro* using peripheral blood mononuclear cells and HeLa-CD4-CCR5 cell lines. They also performed structural modeling using a crystal structure of gp120-CD4-X5. These experiments identified two envelope motifs, compact V1–V2 loops and V3-316T, which are found at high frequency throughout subtype-C evolution and affect gp120 interactions with CD4 and CCR5, respectively, indicating that these conserved two motifs might be important in the transmission and spread of subtype C. Moreover, in support of the notion that subtype C is more favored for ease of transmission comes from the observation and experimental findings that the viruses involved in new infections are R5 phenotypes, the phenotype that persist throughout the course HIV-1 infection in the larger population of subtype C (Montano *et al*., 2000).

Among other things, route of transmission also seems to affect transmissibility of subtype. The observation that HIV-1 *env* subtype E and C strains from Thai and Indian heterosexuals, respectively, infected Langerhans' cells more efficiently than *env* subtype B strains from US

homosexual men suggested that Langerhans' cell tropism of subtypes E and C resulted in more efficient vaginal transmission, which helped explain the explosive heterosexual subtype E and C epidemics seen in Thailand, India, and South Africa (Mastro *et al*., 1997; van Harmelen *et al*., 1997). On the other hand, subtype B strains, which predominate among homosexual men and IDU in industrialized Western countries, have been found to be transmitted less efficiently vaginally, while retaining the potential for efficient transmission via blood, including transmission resulting from microscopic abrasions presumed to occur during receptive anal intercourse (Mastro *et al*., 1997).

Another important determinant for rate of HIV spread (transmission) and predominance of specific subtypes is increased viral fitness. Viral fitness refers to an organism's replicative adaptability in a given environment (van Opijnen and Berkhout, 2005). HIV-1 variants with high levels of fitness have a selective advantage over other less-fit variants in clinical infections and transmission, suggesting that reduced replication efficiency is one property of variants with reduced disease progression (Dykes and Demeter, 2007). While differences in transmission fitness between the same or different subtypes has been shown largely due to genotypic differences at viral envelop (gp120), particularly at V1-V3 (Marozsan *et al*., 2005; Saunders *et al*., 2005; Troyer *et al*., 2005) viral fitness in many other aspects of HIV strains results from the interplay between the environment and all genes and processes that have a role in the viral life cycle (van Opijnen and Berkhout, 2005). In relation to this, comparative studies were conducted by different groups of investigators to see the impacts of subtypes on the relative replication fitness (*ex vivo*) of various HIV isolates of the major (M) group HIV-1 subtypes A, B, C, D, and CRF01_AE; and group O. These studies have shown that all group M isolates were more fit than group O isolates, suggesting that reduced replicative and transmission fitness may be contributing

to the low prevalence and limited geographical spread of HIV-2 and group O HIV-1 in the human population (Abraha *et al*., 2009; Arien *et al*., 2005). Among the subtypes within group M, subtype C isolates were found to be less fit than A, B, D, and CRF01_AE subtypes (Abraha *et al*., 2009; Ball *et al*., 2003; Gretti, 2006). It was speculated that weakness in C could be related to properties residing in gp120, in particular to a reduced avidity for binding the CD4/CCR5 receptors on host cells (Geretti, 2006). It was further suggested that the efficiency of host cell entry may have a significant impact on relative fitness of the infecting virus (Ball *et al*., 2003; Geretti 2006).

1.4.2.5 Impacts of HIV diversity on host immune response/evasion and vaccine efficacy

Development of vaccine for AIDS faces several challenges including practical, political, economic, social, and ethical problems, of which the most fundamental challenges now reside at the level of the basic biology of HIV-1 infection and pathogenesis (Garber *et al*., 2004). Genetic variability of the etiologic viruses between the different subtypes is considered by many as very critical, although some researchers disagree to the opinion that subtype diversity affects vaccine efficacy. For instance, in support of the latter group of researchers, observations on immune responses mounted by naturally HIV-1 infected individuals have documented more broadly cross-reactive cytotoxic T-cell responses between different subtypes, although intrasubtype responses were often stronger and more frequent than intersubtype reactivities (Cao *et al*., 2000; Currier *et al*., 2003; Hemelaar *et al.,* 2006; Slobod *et al*., 2005). These and other researchers argued that human neutralizing antibodies produced against subtype B could neutralize other primary HIV-1 subtypes including A, C, D, CRF01_AE, and F, suggetsting that clades, as defined by genetic sequence, do not correspond to immunotypes, as defined by mutually exclusive immune responses (Ferrantelli *et al*., 2004; Moore *et al*., 1996; Nyambi *et al*., 2000; Slobod *et al*., 2005). The latter authors

speculated that this cross-clade responsiveness is explained by the fact that the B- and T-cells recognize precise epitopes rather than the overall sequence similarity of viruses (Slobod *et al*., 2005).

However, the reason for this lack of or little relationship found in these earlier studies between HIV-1 clades and neutralization serotypes could be that the previous studies may have been confounded by inadvertent use of inter-clade recombinant viruses as many of the viruses used were not characterized by full-length genome sequencing or even complete envelope (env) gene sequencing (Brown *et al*., 2008). With improved *in vitro* and *in vivo* investiagational protocols, people are producing evidences, since recently, that contradict the opinion that immunologic response to HIV has no relaltionship with specific clades. For instance, a number of studies have identified two classes of neutralizing antibodies specific for the HIV-1 Env protein that are eventually elicited in most infected individuals: namely, strain-restricted and broadly cross-reactive antibodies (Li *et al*., 2009).

The strain-restricted antibodies appear early after infection and are generally directed against linear determinants within the gp120 variable regions, including the second (V2) and the third (V3) regions (or loops) (Li *et al*., 2009). On the other hand, five broadly neutralizing antibodies isolated from HIV-infected individuals have been identified (McMichael, 2006). The two most broadly neutralizing antibodies, 2F5 and 4E10, are directed against the membrane-proximal external region (MPER) of the HIV gp41 (McMichael, 2006). The other two broadly neutralizing antibodies, b12 and 2G12, are directed against the gp120 envelope glycoprotein, while the 5th one is a CD4-induced and directed against co-receptor binding site (Li *et al*., 2009; McMichael,

2006). The b12 antibody recognizes an epitope that overlaps with the CD4 binding site (CD4bs); and 2G12 recognizes a cluster of glycans on the outer domain of gp120 (Li *et al*., 2009).

Additionally, an experimental ivestigagtion conducted to understand the influence of clade on virus neutralization (neutralization sensitivity and antibody breadth) using pooled polyclonal antibodies from well-characterized infections with respect to the infecting viral clade and a large, full-length sequenced panel of viruses, has revealed that pairing of the antibody pool against homologous clade viruses generated the highest geometric mean neutralizing antibody titer in 4 out of 6 clades tested, and that neutralization patterns showed numerous examples of reciprocal cross-recognition between antibody and viruses of specific clade pairs (Brown *et al*., 2008). Futhremore, the researches noted that cross-clade neutralization was more limited, with fewer distinct cross-clade relationships evident, particularly the clade C antibody pool being broadly crossreactive having neutralized the greatest number of viruses in the assays (Brown *et al*., 2008).

Thus, it may be reasonable to believe the existence of subtype-specific epitopes among the various HIV strains. This is supported by various investigators. For example, Gaschen *et al*. (2002) compared estimates of codon-specific ratios of nonsynonymous to synonymous substitution rates (dN/dS) in subtype B and C envelope genes in an attempt to explore the hypothesis that subtype-specific patterns exist in the exposure of antigenic domains that are able to elicit antibody responses strong enough to drive escape mutation. They reported that high rates of diversifying selection were identified in different regions of the envelope protein (Env) in the two lineages in the Env V3 to C4 region: the V3 loop is less variable in the C subtype than in other subtypes, and as expected, the density of sites in the V3 loop with dN/dS was higher in subtype B than in C; this pattern was reversed, however, in the region just proximal to the V3

loop, where multiple sites show an excess of nonsynonymous substitutions in the C but not in the B (Gaschen *et al*., 2002).

Given the enormous worldwide diversity of HIV-1 group M variants, whose gene products at amino acide level diverge between 15% in a relatively conserved region (Gag) and 35% in highly variable region (Env), the design and choice of antigen identities for inclusion in candidate AIDS vaccines is therefore an overwhelming challenge (Garber *et al*., 2004; Gaschen *et al*., 2002). This can be easily understood by looking at the continual changes of influenza vaccine strains which occurs because of a less than 2% amino acid changes between a vaccine strain and circulating strain in a new epidemic (Fischer *et al*., 2007). This global diversity of influenza in any given year, which necessitates a change in vaccine strain, is comparable to the diversity of HIV within a single infected individual at one time point (Korber *et al*., 2001). The influenza vaccine case becomes trivial when one considers the challenges posed by HIV epitope variability among HIV clades. For example, investigators experimentally detected a decline within 391 to 772 days in the ability of neutralizing antibodies from acutely infected patient to neutralize formerly susceptible viruses, suggesting that changes in the viral envelope had occurred within these time intervals that rendered the successive viruses resistant to neutralization (Wei *et al*., 2003). It was also shown that single amino acid changes can allow an epitope to escape T-cell surveillance (Fischer *et al*., 2007). Since many T-cell epitopes differ between HIV-1 strains at one or more positions, potential responses to any single vaccine antigen are limited (Fischer *et al*., 2007).

Now that the need to consider HIV diversity in the development of AIDS vaccine is well-taken by the scientific community and vaccine developers, the big question is how to approach it. A successful vaccine is anticipated to cause such desirable immune responses with sufficient

breadth and depth to cover extensive immunologic diversity and mutational patterns of the virus to exert antiviral activity against a substantial breadth of primary isolate viruses (Barouch and Korber, 2010). It is also desirable that a successful vaccine elicit responses against conserved regions of the viral proteome in which mutations would severely compromise the viability of the virus; and simultaneously, it must not elicit responses against variable elements of the virus that can mutate while retaining function, and that can absorb much of the adaptive host immune response (Rolland *et al*., 2007). To this end, several approaches have been proposed. One approach is that a single clade may induce sufficiently cross-reactive T-cell responses to protect against other variants of both the same and heterologous clades. The choice of a natural isolate can be based on having the closest sequence to all others, or picking a strain derived from acute infection (Letourneau *et al*., 2007). However, research results are available that challenge this approach (Lee *et al*., 2004; McMichael, 2006). For example, Larke *et al*. (2007) indicate that single-clade A, B and C vaccines used alone induce T cells with only limited recognition of variants other than those present in the vaccine; and that the epitope hierarchy varied according to the immunizing clade. Besides, the fact that HIV-1 variants were observed to have escaped from existing T cell responses in infected individuals by single mutations in epitopes (Fischer *et al*., 2007; McMichael, 2006) makes the use of a single natural isolate for a vaccine as a high risk for protecting against a different clade and even against many variants of the same clade.

A second approach considers deriving vaccine immunogens from‘centralized’ sequences, which employ consensus/average, or centre-of-the-tree or sequences or extrapolated amino acids to a common clade or group ancestor (Gashen *et al*., 2002; Korber *et al*., 2009). However, there is a fear that this strategy may be stretched too far for optimal coverage of CD8 + T cell epitope variants of the whole group M (Letourneau *et al*., 2007). A third one uses vaccines capable of

delivering a cocktail of immunogens derived from different clades (McMichael, 2006). This approach is not without its problem, since possible immune interference, such as epitope antagonism between different but closely related peptide sequences in the vaccine may limite responses to some epitopes (Larke *et al*., 2007). A fourth approach focuses on the use of computational methods for assembling a polyvalent vaccine candidate that optimizes the coverage of T cell epitopes composed of mosaic immunogenes. In this method, it is beilieved that mosaics provide diversity coverage comparable to that afforded by thousands of separate peptides, but, because the fragments of natural proteins are compressed into a small number of native-like proteins, they are hopped to be good for vaccines (Fischer *et al*., 2007).

Yet a new approach is proposed by Letourneau *et al*. (2007), wherein a novel T cell immunogen, designated HIV_{CONSV}, was designed by assembling the 14 most conserved regions of the HIV-1 proteome into one chimaeric protein. The authors explain that each segment is a consensus sequence from one of the four major HIV-1 clades A, B, C and D, which alternate to ensure equal clade coverage. They claimed that this vaccine approach provides an attractive and testable alternative for overcoming the HIV-1 variability, while focusing T cell responses on regions of the virus that are less likely to mutate and escape (Letourneau *et al*., 2007). In conclusion, an HIV-1 vaccine, whether composed of natural or artificial immunogenes, needs to elicit substantially potent and cross-reactive immune response taking HIV-1 diversity into considerarion. Each approach has its potential merit and drawbacks, which will be known with certatinity after evaluation in clinical studies.

1.5. HIV/AIDS in Ethiopia

1.5.1 Epidemiology of HIV in Ethiopia

The earliest time point for detection of HIV among Ethiopians was 1984, which was detected when a serological survey was conducted on 1754 people in various parts of the country from 1982-1987 (Tsega *et al*., 1988). Two studies working on estimating the timing of the first introduction and subsequent evolution of HIV in Ethiopia have reported discordant results: while the earlier estimate by Abebe *et al*. (2001) claimed HIV-1 to have entered into Ethiopia around 1983, a recent estimate by Tully and Wood pushed it further back into late 1960s or early 1970s (Tully and Wood, 2010).

Since the first report in 1984 of HIV's presence in Ethiopia, its prevalence has seen three patterns. Initially, the prevalence rate was slow (about 0.5%) until 1985 (Abebe *et al*., 2001). From this time on, however, the epidemic spread rapidly and expanded its sphere among various risk groups. In 1988, its prevalence rate among long-distance track drivers and commercial sex workers along the main trading road was 13% and 17% respectively (Mehret *et al*., 1990a; 1900b). The prevalence among women attending antenatal-care clinics in Addis Ababa was 4.6% in 1989 but rose to 15% in 2001 (Tsegaye *et al*., 2002), while the prevalence in Afar Region was documented as 19.7% (Assefa *et al*., 2003). Within short time period, the prevalence shot up to an alarmingly high rate, particularly in Addis Ababa: beginning middle of 1990s until the last 1990s it was estimated to be 10-27% in pregnant women, at least 45% in commercial sex workers and 7% among blood donors (Abebe *et al*., 1997; Hussein *et al*., 2000).

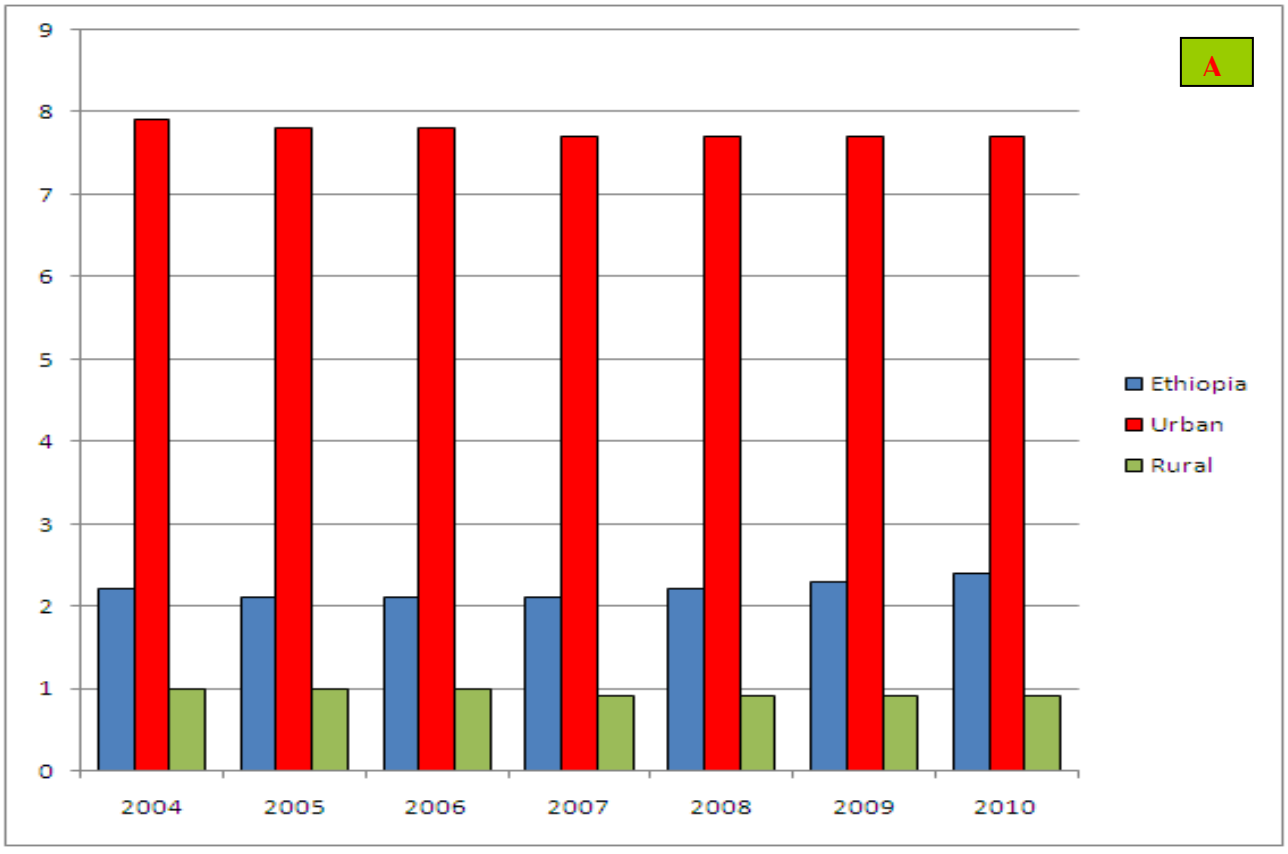
Same high prevalence rate was documented among a cohort of 1,679 factory workers with baseline prevalence of 9.4% among males and 12.4% among females (de Wit *et al*., 2002;

Mekonnen *et al*., 2005a). A report from a retrospective study of HIV-1 prevalence among 63,869 visa applicants between the years 1993 to 2001 has found an overall period prevalence of 9.5%, but ranged from 6.8% in 1993 to 10.4% in 1997, with stabilized prevalence at around 11% after 1997 (Tegbaru *et al*., 2004). Analysis of sera collected in 1994 for determination of the rate of acquisition of antibodies against measles, rubella, and hepatitis B have shown the prevalence of 6.0% for men and 6.9% for female with peak prevalence in the 25–29 year age group of 16.3 and 11.8% respectively (Fontanet *et al*., 1998), indicating that 1990s were the years during which HIV prevalence was the highest in Ethiopia.

The third pattern appeared in the first half of 2000s, in which the prevalence has shown sign of decline in the urban and rural populations (Solomon and Murray, 2001). This was supported by a review from Kebede *et al*. (2000), which examined published and unpublished reports and surveillance data from records of governmental and non-governmental institutions. In addition, a recent report from MOH/HAPCO (2006) has shown that out of the 25 urban sites that had HIV prevalence data for 2002, 2003, and 2005, 12 had significant declines while 10 had either non-significant increase (9/10) or no change (1/10) from the previous trends. The situation in rural sites were different: of the 7 sites with available data of for comparable years, only 2 had significant decline while 4 of them had increasing prevalence with 1 site statistically significant increase. However, the overall trend of the 7 rural sites showed a statistically significant declining trend (MOH/HAPCO, 2006).

The current status HIV prevalence in Ethiopia is that it is still one of the countries that are hit-hardest by HIV/AIDS epidemic. It is estimated that the national prevalence of HIV in 2007 was 2.1% (MOH/FHAPCO, 2007a). The most recent estimate by UNAIDS/WHO reports that the

overall prevalence rate in 2008 among adults was 1.4%, (UNAIDS/WHO, 2009), which is a substantial decline compared to what has been reported in the year 2005 in which antenatal care surveillance data indicated an adjusted prevalence of 3.5% (MOH/HAPCO, 2006), and contrary to what has been predicted by MOH/FHAPCO (2007a) report as depicted in (Figures 1.5.1.1A and B) Several factors could have played in the reportedly declined HIV prevalence rate in Ethiopia, including improved ANC-based sentinel surveillance system with increased rural representation where prevalence is lower; impact of mortality of HIV infected individuals; and HIV/AIDS control and mitigation efforts including those that lead to behavior change (MOH/HAPCO, 2006). Moreover, it has been predicted that AIDS deaths in both rural and urban areas are expected to decline from 2006 onwards, though more pronounced in urban areas, mainly due to the expanding ART program (MOH/HAPCO, 2006).



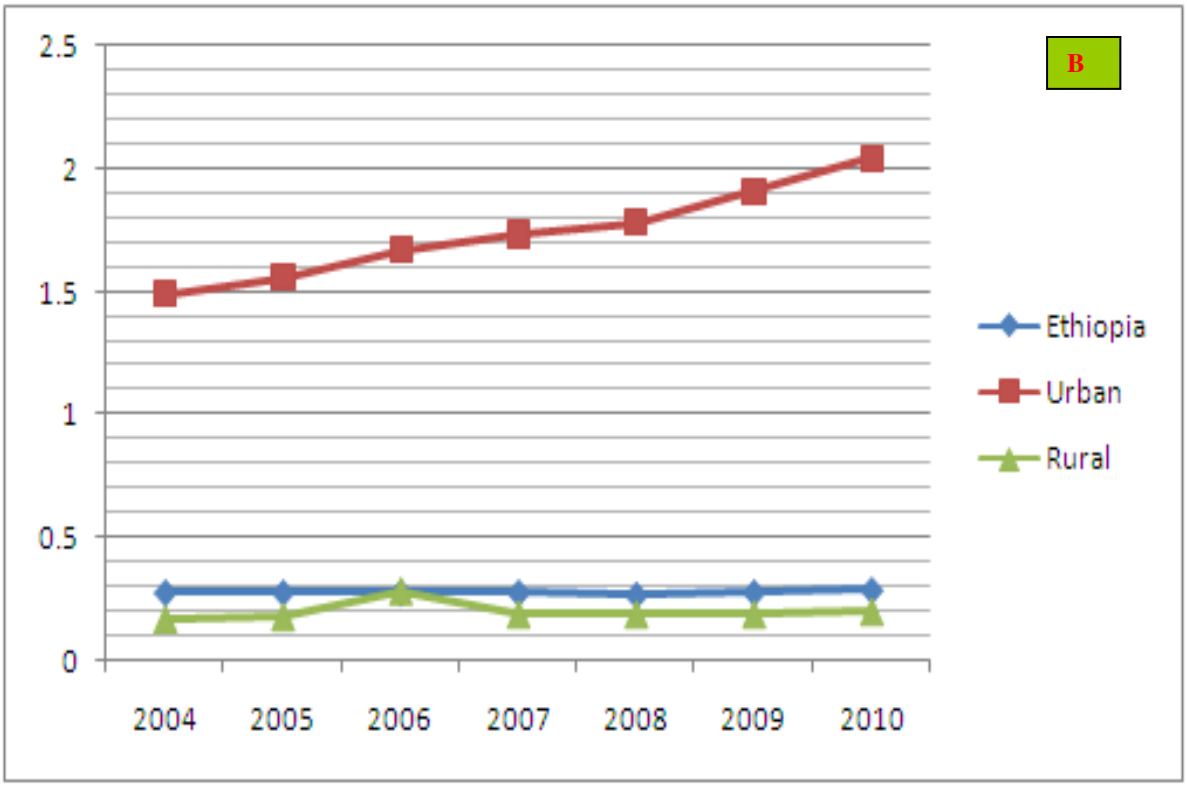


Figure 1.5.1.1 Estimated HIV rates (in per cent) of prevalence (A) and incidence (B) during 2004-2010 (Data from MOH/FHAPCO, 2007a)

1.5.2 ARV treatment in Ethiopia

Various intervention activities have been in place to curb the expansion of HIV-1 epidemic in Ethiopia since the National HIV policy was endorsed in 1998, one of which was ART initiative in 2003 (MOH, 2005) and further scaling up of this initiative as the result of supports from stakeholders like President's Emergency Plan for AIDS Relief (PEPFAR), Global Fund, and the World Health Organizations. Prior to the launching of the 2003 ART initiative, antiretroviral drugs had been in use since 2000 by few people who could purchase the drugs from the black market (MOH, 2005). From 2003 to 2005, five ARVs were distributed to selected Urban Dwellers Association Pharmacies and Ethiopian Red Cross Pharmacies and made available upon

prescription from ART-certified physicians: three nucleoside analog reverse transcriptase inhibitors (zidovudine, lamivudine and stavudine) and two non-nucleoside analog reverse transcriptase (nevirapine and efavirenz). During this period, about 5000 registered patients started the medication, 90% of whom were paying US$30-90 a months while the rest 10% were treated free of charge if they could produce written support from local administration (*kebele*) (MOH, 2005).

With the launching of ART scale-up program in January 2005, however, the drugs became freely available to the needy patients, and the number of beneficiaries has been ever increasing since then. At the end of September 2004, only 9,500 AIDS patients could get access to ART (PEPFAR, available at http://www.avert.org; accessed on 1 September 2006); but this figure rose drastically in the subsequent years: 13,100 in 2005 (MOH, 2005), 46,000 in 2006 (MOH/HAPCO, 2006), and so on as shown in Figure 1.5.2.1. The latest report from MOH/HAPCO (2010) indicates that about 180,000 people are currently on ART in 532 health facilities throughout the country. Diversity of available ARV drugs has also increased from five in 2000 to about ten in various combinations in the current regimens (MOH/HAPCO, 20007b). An updated guideline for recommended ART regimens in Ethiopia has been provided by MOH/HAPCO in 2007, and is summarized in Tables 1.5.2.1 and 1.5.2.2.

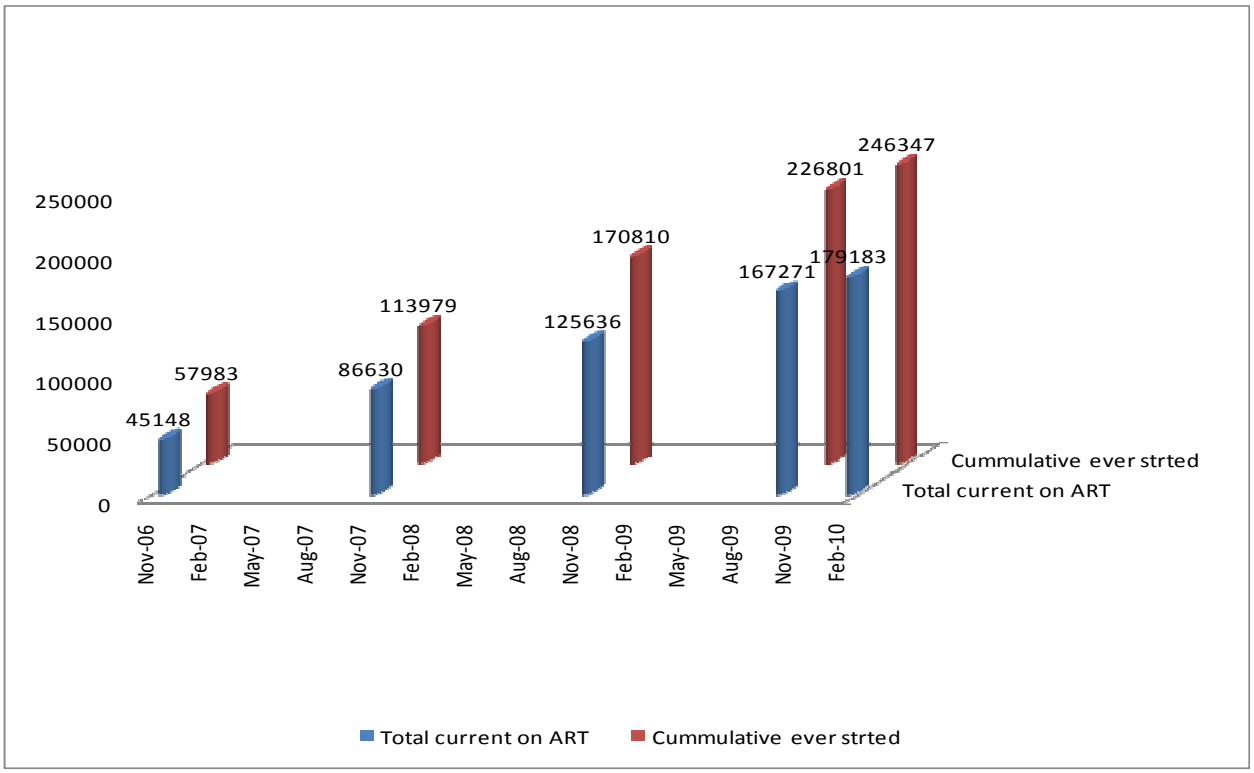


Figure 1.5.2.1 Number of people on ART since November 2006 to February 2010
(Data obtained from MOH/HAPCO ART monthly updates. Available at http://www.etharc.org/resources/download/)

Evidences show that the scale-up of ART in Ethiopia has greatly improved the longevity of HIV-infected persons in the country. Retrospective and prospective studies on mortality impact of AIDS in Addis Ababa showed that prior to the introduction of ART in 2000, mortality had increased by at least 5.0% since 1984 due to HIV; the same study claims that the probability of Ethiopians dying of HIV/AIDS before the age of 60 was more than 17.5% (Sanders *et al*., 2003a). However, recent mortality estimate from verbal autopsy interview surveillance of burials at all cemeteries in Addis Ababa has found a decline in AIDS-related mortality between 2001 and 2005 by 21.9% and 9.3% for men and women respectively; and by 38.2% for men and 42.9% women between 2005 and 2007, which would have been between 56.8 and 63.3% in the absence of ART (Reniers *et al*., 2009).

Table 1.5.2.1 Recommended First-Line ART Regimens for Adolescents and adults in Ethiopia (MOH/FHAPCO, 2007b)

Recommended ARV Regimens for Adults and Adolescents: One of the following should be used unless there are contraindications:
Preferred : • TDF+FTC+EFV = triple FDC • ZDV+3TC+EFV = combivir + EFV • ZDV+3TC+NVP = triple FDC
Alternatives: • D4T/3TC/EFV = double FDC (d4T/3TC) + EFV • TDF/3TC/NVP • D4T/3TC/NVP = triple FDC • ABC/3TC/EFV • ABC/3TC/NVP • ABC/3TC/ZDV = combivir + ABC

Table 1.5.2.2 Recommended Second-Line ARV Regimens for Adolescents and adults in Ethiopia (MOH/FHAPCO, 2007b)

First-line Regimen	Second-line Regimen (*during treatment failure*)
TDF+FTC or 3TC +EFV or NVP	ZDV ±3TC +LPV/r or ATV/r Or ZDV+ABC+LPV/r or ATV/r
ZDV or d4T+3TC+EFV or NVP	TDF+3TC±ZDV+LPV/r or ATV/r Or ABC + ddI[a] +LPV/r[b] or ATV/r
ABC + 3TC + ZDV	EFV or NVP + LPV/r or ATV/r

[a] Didanosine alone must be taken on an empty stomach, at least one hour before or at least 2 hours after (<50% absorbed after) a meal. Tablets should be dissolved in at least 30 ml of water; no other liquids may be used to dissolve the tablets. The enteric coated version will not need to be dissolved.
[b] LPV/r use the heat stable tablet (200/50 mg).

1.5.3 HIV Subtype Distribution in Ethiopia and Possible Gap of Information on Ethiopian Subtype Variation

Several types of HIV-related researches have been conducted since the first report of HIV detection in Ethiopia in 1984 (Tsega *et al*., 1988). Currently, a huge amount of published and unpublished HIV-related research reports are being produced throughout the country, which can be categorized into researches on basic biomedical, epidemiological, risk factors, risk determinants, clinical, impacts, intervention, and monitoring and evaluation (Mulatu *et al*., 2005). The researches conducted so far have been reviewed and updated on annual basis since 2003

under the title “Bibliography on HIV/AIDS in Ethiopia and Ethiopians in the Diaspora”, and are published in the Ethiopian Journal of Health Development (Mulatu *et al*., 2005; Convers *et al*., 2006; Mulatu *et al*., 2007; Kloos *et al*., 2008; Converse *et al*., 2009).

However, the pioneering research works relevant to HIV-1 subtyping in Ethiopia were laid down when the Ethio-Netherlands HIV/AIDS Research Projects (ENARP) became operational in 1994 and continued producing excellent quality research outputs before it ended in 2002 (reviewed in Sanders *et al*., 2003b). A number of important scientific findings about HIV/AIDS in Ethiopia have been contributed by the project. One aspect in this regards is related to the unique nature of the immune system of Ethiopians: (1) CD4+ T cell counts in HIV-negative Ethiopians is lower (median 749/μL) than their Dutch counter parts; (2) in HIV-positive Ethiopians after seroconversion, and at stages 1 and 2 of the WHO staging system, CD4+ T cell count values were 150–200 cells/μL lower than what has been observed among their counterparts in other countries, including the Netherlands; (3) the immune system of Ethiopians is in a state of permanent activation (Sanders *et al*., 2003b).

In addition, it was learned that the presence of low CD4+ T cell count among HIV-positive Ethiopians was not found to enhance rate of CD4+ cell depletion nor decrease their life expectancy more than any HIV-positive population in other countries (Mekonnen *et al*., 2005b). The ENARP projects were also able to develop NASBA molecular beacons for a *gag*-based (Ayele *et al*., 2004) and an *env*-based assays that could be used to discriminate between subclusters C and C’ Ethiopian isolates (Ayele *et al*., 2005). Many more findings from earlier studies have contributed to the diagnosis and monitoring of disease progression among HIV-infected individuals. For example, by the end of 1980s it was already determined that primer

pairs in the *gag*, *pol*, and *env* regions that could amplify HIV-1 DNA in peripheral mononuclear blood cell samples from HIV-1-seropositive individuals in Sweden (probably subtype B) were unable to amplify HIV-1 DNA obtained from Ethiopian samples (most probably subtype C) (Ayehunie *et al*., 1990), showing the poor performance of early primers to broadly amplify all HIV subtypes that were distributed among various populations.

The other important findings from earlier research works, which are more relevant to this study, relate with the genetic characteristic of HIV isolates circulating among HIV-infected Ethiopians. A number studies conducted to determine subtyping of Ethiopian HIV-1 isolates have repeatedly found that subtype C and its variant or subcluster C' are the predominant HIV-1 isolates in Ethiopia (Abebe *et al*., 1997; Abebe *et al* 2000; de Wit *et al*., 2002; Harris *et al*., 2003; Hussien *et al*., 2000; Salminen *et al*., 1996) and remained unchanged for at least five years (1988-1993) (Sherefa *et al*., 1994). However, subtypes A (Abebe *et al*., 1997; Abebe *et al* 2000; Adal *et al*., 2005), D (Abebe *et al* 2000; Hussien *et al*., 2000), and A/C recombinant (Sherefa *et al*., 1998) have also been identified from different localities, albeit few in number. Most of these subtyping works were done based on sequences from the *env* (C2V3 or gp41) and/or *gag* regions (Abebe *et al*., 1997; Abebe *et al* 2000; Ayehunie *et al*., 1993; de Wit *et al*., 2002; Hussien *et al*., 2000).

Only few full or near full length HIV-1 genome sequencing of Ethiopian isolates have been performed that identified the genetic characteristics of the isolates (Harris *et al*., 2003; Salminen *et al*., 1996; Sherefa *et al*., 1998). Scattered sequencing has also been conducted on other subgenomic regions such as partial gag/pol (Sönnerborg and Sherfa, 1997) and LTR (De Baar *et al*. 2003). Thus, it is clear that not enough genomic characterization of Ethiopian HIV-1 isolates has been done, when one considers the duration of the epidemic's existence in the country.

Therefore there is a need for more works on genomic characterization of HIV-1 isolates from Ethiopian patients, for which reason this research project was carried out.

CHAPTER 2: THE PROJECT

2.1 The Rationale

Given the biological significances (such as level of virus load, disease progression, chemokine receptor use, vertical transmission rate, transcriptional activation rate and consequently replication rate, clinical management, vaccine design, and drug resistance pattern) associated with genomic differences among the different HIV-1 subtypes, it is important for any country to understand the genomic characteristics of the serotypes circulating among the population. Although several important studies were carried out in Ethiopia to determine HIV-1 subtypes circulating in the country, most of the studies focused on characterization of V3 loop and gp41 of the *env*, p7 and p17 of gag region (Abebe *et al*., 1997; Abebe *et al* 2000; Ayehunie *et al*., 1993; de Wit *et al*., 2002; Hussien *et al*., 2000; Sherefa *et al*., 1997; Sönnerborg and Sherefa, 1997).

However, such a rely on limited region of the genome can be potentially misleading, especially at this time that HIV-1 circulating recombinant forms (CRFs) are overtaking the upper hand from the pure forms. There are a number of instances that showed the inadequacy of using a portion of the genome (in contrast to full-length) in characterizing HIV subtypes circulating in a given geographic area. For example, a full-length sequence analysis performed by Carr *et al*. (1996) showed that an isolate from Thailand, formerly designated as subtype E by envelope sequence, was actually a recombinant form between subtype A (in the entire *gag/pol* region and most accessory genes) and mixture of subtype E and subtype A in the envelope sequences. Similarly, the clade which was thought to belong to subtype C based on sequencing of C2V3 region turned out to have mosaic genomic structures of subtypes C (in the envelope region) and B' Thai in the rest of the genome by full-length genomic analysis (Su *et al*., 2000). Even more intriguing is the finding that the clade known today as subtype K based on the result of a near-full-length genome

sequencing study, was formerly designated as divergent form of subtype F as identified by envelope sequences (Triques *et al*., 2000).

Because only few full length (Harris *et al*., 2003; Salminen *et al*., 1996; Sherefa *et al*., 1998) or regions other than *env* (De Baar *et al*. 2003; Sönnerborg and Sherfa, 1997) have been used in the previous genotype determination studies of Ethiopian HIV-1 isolates, it is possible to assume that the widely accepted subtype C and its close relative C' as the predominant subtype circulating in the country might have suffered from same limitations mentioned above. In other words, it is possible that the presence of recombinant forms or other subtypes could have been missed by the inherent deficiency of partial sequencing of *env* and/or *gag* used in the previous studies. After all, subtypes A and D were identified from isolates originating from Addis Ababa and Dessie by those sequencing works targeting even the env region (Abebe *et al*., 1997; Abebe *et al* 2000; Adal *et al*., 2005; Hussien *et al*., 2000). In addition, from the few full genomes analyzed, one study has already found a recombinant A/C subtype from a serum sample collected early in the epidemic: the genomic composition of this isolate was such that the *gag* and *pol* genes and the major parts of the *env* genes scored best with subtype A while the region flanking the V3 loop, and the V3 domain itself, matched best with subtype C (Sherefa *et al*., 1998). Were this subtyping done on V3 loop alone, as has been often the case, this isolate would have been grouped under subtype C, implying that subtypes and recombinant forms other than identified so far could be found in Ethiopia if careful sequence studies were conducted on samples collected from several regions of the country.

Moreover, the fact that relatively longer time elapsed since publications of the few full genome analysis of Ethiopian HIV-1 subtype C (Harris *et al*., 2003; Salminen *et al*., 1996; Sherefa *et al*.,

1998) indicates that there is an enormous likelihood for emergence of new variants, as the result of rapid mutation and recombination events that might have taken place since then. Furthermore, given the high prevalence rate of non-C HIV-1 subtypes in the neighboring countries (Gray *et al*., 2009; Janssen *et al*., 1997) coupled with the current trend of international travel and immigration (Gray *et al*., 2009), it is reasonable to presume that these non-C subtypes might have already been introduced into Ethiopia, and that conducting analysis of isolates at genomic regions other than *env* may reveal a new HIV subtype profile in the country.

Another source of concern with regard to emergence of new subtypes (or CRFs) in this country is the recently introduced Highly Active Antiretroviral Therapy (HAART). Although currently used antiretroviral drug combinations are capable of maintaining low-to-undetectable plasma HIV-RNA level, which is very essential to prevent progression to AIDS, they do not eliminate or even completely suppress HIV-1 replication in all tissue compartments (Barbour *et al*., 2002; Coombs *et al*., 2003; Quiñones-Mateu and Arts, 2001). Because of the selective pressure exerted on drug-sensitive strains during this therapy, variants that naturally contain genes for drug resistance will gain extra advantages of selective survival and transmission (Ait-Khaled *et al*., 2003; Moutouch *et al*., 1996). In this regard, data have already shown that HIV-1 strains from all types, subtypes and geographic regions contain natural polymorphisms in the protease and reverse transcriptase genes, whose biological significance as yet not determined for the majority of them (Alexander *et al*., 1999; Gibb *et al*., 2002; Gordon *et al*., 2003). Since HAART allows people infected with HIV to live longer and feel healthy, there is a chance that patients on HAART resume their normal sexual life. Concomitant to this, the likelihood of transmitting drug resistant strains will be enhanced if unsafe sex is practiced. This in turn could be expected to increases the chance of

superinfection and recombination both in newly infected and doubly infected formerly seropositive individuals. Thus, one of the aims of carrying out this PhD dissertation project was to fill the aforementioned gaps of information about Ethiopian HIV-1 subtypes and possible CRFs by performing genome analyses at regions of protease and reverse transcriptase, the regions which were poorly represented in the previous subtype determination studies.

Another gap of information relates to HIV-1 drug resistance profile of Ethiopian isolates. Most of the current knowledge of HIV-1 drug susceptibility and resistance and interpretations of genotypic changes in HIV-1 reverse transcriptase and protease are based on data obtained from HIV-1 subtype B viruses, whose prevalence is restricted mostly to the Western world. In the rest of the world, where non-subtype B viruses are dominating, a high frequency of natural polymorphisms associated with resistance to protease inhibitors and reverse transcriptase inhibitors have been observed (Grossman *et al*., 2001; Kantor and Katzenstein, 2003). The presence of these unique mutations and of secondary mutations not found in subtype B may potentially be associated with the development of different resistance profile in non-B subtypes during therapy (Martinez-Cajas *et al*., 2009). In this regard, certain questions are unanswered, remaining open for investigations: Are the current drug regimens, widely used in subtype B infection equally effective in other HIV subtypes? Will increased access to antiretroviral therapies in developing countries, such as Ethiopia, result in new patterns of antiretroviral drug resistance? Can the currently used genotypic drug resistance assays and interpretations, which were developed for subtype B, detect all the mutations associated with phenotypic drug resistance in non-subtype B? While the answers for these questions can be furnished from large-scale clinical trials and monitoring drug responses through surrogate indicators such as viral load and

CD4 count, performing simultaneous genotypic, phenotypic and fitness tests will give valuable information regarding the relevance of the unique polymorphic amino acid substitutions to treatment-related drug resistance in non-subtype B strains.

It is generally believed that genotypic drug resistance analysis of HIV predicts patients' response to antiretroviral therapy (Haubrich and Demeter, 2001) and that understanding the basis of HIV drug resistance represents a key requirement for individualized HIV patient care. These facts along with the wider accessibility of HAART in Ethiopia indicate the urgent need for understanding the drug-susceptibility/resistance pattern of HIV-1 isolates from the country. Hence, this PhD project was also undertaken to determine prevalence and rate of genotypic drug resistance profile of HIV-1 isolates to antiretroviral drugs against protease and reverse transcriptase on specimens collected from antiretroviral-naïve and antiretroviral drug-experienced adult Ethiopians.

Worldwide observations have shown that the continuous use of ART is accompanied by emergence and expansion of drug resistant HIV-1 strains against ARV drugs in use in the region. As Ethiopia is one of the benefactor countries of the expanded access to ART, preparations have to be made by health care providing bodies for the inevitable emergence of drug resistance, through implementation of genotypic drug resistance monitoring system. However, due to high costs and stringent requirements for storage and transport of plasma specimens, drug resistance monitoring is not done as part of clinical practice in Ethiopia. One way of alleviating the stringent requirements for storage and transport of plasma is through use of field-friendly sample collection methods such as Dried Blood Spot (DBS) (Bertagnolio *et al*., 2007). Several

investigations have already tested and validated the specificity and sensitivity of DBS for viral load (Ayele *et al*., 2007) and genotypic drug resistance monitoring (Hamers *et al*., 2009) under their own specific settings, and nearly all of them found it satisfactory for the purpose (Viljoen *et al*., 2010). However, the performance DBS in detecting drug resistance mutations in Ethiopia is not known. Therefore the other aim of this PhD dissertation project was to validate the suitability of utilizing DBS for HIV-1 drug resistance testing under the Ethiopian settings.

2.2. Objectives

2.2.1 General Objectives

To characterize the genetic diversity of HIV-1 subtypes and determine antiretroviral drug resistance profile of HIV-1 isolates from HIV-infected Ethiopians at various stages of the infection.

2.2.2 Specific Objectives

1. Determine prevalent HIV-1 subtypes and circulating recombinant forms (CRFs) from isolates of Ethiopian patients using sequence analyses from protease and reverse transcriptase genomic regions;
2. Determine genomic differences between recently infected ARV drug-naive, chronically infected ARV drug-naïve but treatment eligible, and chronically infected ARV drug-experienced patients of HIV-1 isolates from Ethiopian HIV-infected individuals;
3. Determine profiles of genotypic drug resistance mutations at reverse transcriptase and protease regions from isolates of antiretroviral drug-naïve persons who were recently infected or chronically infected but treatment eligible; and antiretroviral drug-experienced chronically infected patients.

4. Evaluate the Dried Blood Spot (DBS) technology for use in HIV drug resistance testing by using double whole blood and plasma samples.
5. Compare the performance of In-house-brewed primers and associated reagents against those from a commercial kit for detecting drug resistance assays targeting protease and reverse transcriptase regions.

2.3. Study Methods

2.3.1 Study Population and Design

Three different HIV-infected populations were targeted for this study. These included recently infected ARV drug-naive, chronically infected ARV drug-naïve but treatment eligible, and chronically infected heavily treated patients. Because of the differences between these populations, three different study designs were utilized.

2.3.1.1 Recently HIV-infected ARV drug-naïve study population

Selection of study population for the purpose of detecting transmitted HIV drug resistance mutations from recently HIV-infected individuals was accomplished in accordance to the guideline formulated by WHO, so-called Truncated Sequential Sampling (TSS) for HIV Drug Resistance Threshold Survey (HIVDR-TS), as a novel proxy method to track the emergence and transmission of HIV drug resistant strains in countries scaling up ART (Bertagnolio and Sutherland, 2005; Myatt and Bennett, 2008). Briefly, the protocol recommended eligibility for HIVDR transmission surveys be confined to individuals under 25 years of age and, in the case of women, those with no previous pregnancies; and sampling $\leq$50 specimens from individuals consecutively diagnosed with HIV and likely to be recently infected in a specific city or geographic area within a country (Myatt and Bennett, 2008). Using this method, it is possible to classify transmitted drug resistance into three; namely, low prevalence ($\leq$5%), moderate

prevalence (5-15%), and high prevalence (≥15%) (Myatt and Bennett, 2008). In this study, therefore, consideration was given to HIV-positive pregnant women who participated in the 2005 ANC-based HIV surveillance program in the seven health centers in Addis Ababa selected as sentinel centers by the Ministry of Health, and who met all the inclusion criteria indicated in the guidelines provided by WHO (Myatt and Bennett, 2008) and the National Sentinel Surveillance Program (MOH, 2004). The inclusion criteria for the study were: (i) being pregnant woman who visited either of the seven identified health centers for regular prenatal follow-up during the working days between 6 April to 8 August 2005; (ii) donating blood for syphilis screening (RPR testing) as part of routine antenatal care; (iii) age below 25; and (iv) current pregnancy being the first ever. The seven antenatal care centers used during the study period were Akaki clinic, Gulele clinic, Higher 23 clinic, Kazanchis clinic, Teklehymanot clinic, Armed Forces Hospital, and Federal Police Hospital.

2.3.1.2 Chronically HIV-infected ARV drug-naïve study population

For this component of the study, the source population was from those HIV-infected individuals who were on medical follow-up to start ART at St. Paul's General Specialized Hospitals in Addis Ababa. Selection for inclusion of study participants depended on age (≥18 years), ARV treatment status (ARV drug-naïve), level of HIV infection (chronically infected and ARV treatment eligible), and willingness to take part in the study. All patients meeting the stated criteria and who visited the hospital for medical follow-up purpose between April and July 2005 were eligible for the study. One hundred and four eligible individuals were included in the study within the specified time frame.

2.3.1.3 Chronically HIV-infected heavily ARV treated study population

The purpose of this component of the study was to identify prevalent ARV drug resistance mutations from as many treatment failing patients as the resource made available for this study allowed. Therefore, the source population was all HIV-infected patients who were already on ART for quite some time. Earlier reports from the United States have shown a 63% virological failure rate among treated individuals (Richman *et al*., 2004). However, due to the short duration since ART initiation in Ethiopia at the time of genotyping, about 20% virological failure rate was arbitrarily taken in the calculation using the formula presented below:

$$\frac{n = (Z_{1-\alpha/2})^2 PQ}{D^2}$$

Where $Z_{1-\alpha/2} = 1.96$; P = Prevalence of virological failure at the rate of 20%; Q = 1 – P; and D = Margin of error

n= $[(1.96)^2 \times 0.2 \times 0.8]/(0.05*0.05) = 246$

With the assumption of 20% amplification and genotyping failure (Beddows *et al*., 2003), the required sample size was calculated to be 295. However, due to resource constraints, the number of recruited HIV-infected heavily pretreated patients suspected of being treatment failing were 266 only. Standard criteria for selecting treatment failing individuals were adopted from the 2008 guidelines developed by the US Department of Health and Human Services (DHHS, 2008). Treatment failure is defined as suboptimal response to ARV therapy, which is expressed by immunological failure, virological failure, and clinical deterioration. Immunological failure can be defined as a failure to achieve and maintain an adequate CD4+ T-cell response despite virologic suppression, which may be expressed in patients who fail to increase CD4+ T-cell counts above a specific threshold (e.g. >350 or 500 cells/μL) over a specific period of time (e.g. 4–7 years) or failure to increase CD4+ T-cell counts above pre-therapy levels by a certain

threshold (e.g. >50 or 100 cells/μL) over a given time period. Virologic failure is defined as the inability to achieve or maintain suppression of viral replication to levels below the limit of detection (<50 copies/ml) and may manifest as: incomplete virologic response, where two consecutive HIV RNA >400 copies/ml after 24 weeks or >50 copies/mL by 48 weeks are detected in a treatment-naïve patient who is initiating therapy; or virologic rebound, where after virologic suppression, repeated detection of HIV RNA is observed above the assay limit of detection (e.g., 50 copies/ml). By clinical progression it means occurrence of HIV-related clinical events that are progressing towards AIDS-defining illnesses (DHHS, 2008).

Accordingly, the criteria set for consideration to enlist participants in this study included: HIV-infected patients being on ARV treatment for at least 24 weeks (6 months), and if RNA load test result was available virologic failure (repeated HIV RNA >400 copies/ml; or >50 copies/ml by 48 weeks; or repeated detection of HIV RNA after virologic suppression), and/or immunologic failure (failure to increase CD4 count by 25-50 cells/ μL above the baseline over the first year of therapy; or a decrease to below the baseline CD4 count on therapy), and/or clinical progression (occurrence or recurrence of HIV-related events), detectable viral load result from the blood sample drawn at the time of this study, and being willing and providing written consent to participate in the study. Assessment of these parameters along with collection of WHO clinical staging and treatment history, and selection of eligible participants were made by the health care providers from each patient's clinical history in the data log. Two hundred and sixty-six patients were screened initially to be included in the study. All the participants were taken from ART clinics of Zewditu Memorial Hospital and Tikur Anbessa Specialized Teaching Hospital.

2.3.2 Ethical Issues

For the purpose of addressing ethical issues, recently infected and chronically infected study participants were treated differently.

2.3.2.1 Ethical issues for recently infected study participants

The National HIV Sentinel Surveillance Guidelines (MOH, 2004) was followed in selecting sample collection approaches from recently infected study subjects. ANC-based HIV surveillance used leftover blood from routine syphilis testing and routinely collected information. Since this study was undertaken as part of the ANC-based HIV surveillance, only left-over blood after syphilis testing was taken, which was then labeled with a surveillance code number and tested for HIV in unlinked and anonymous fashion without any personal identifiers. No consent was sought, and no feedback of the HIV test result was provided to ANC client (MOH, 2004). However, women were informed about the availability of voluntary HIV counseling and testing (VCT) and prevention of mother to child HIV transmission (PMTCT) services available in their area. All participants who tested positive for syphilis were treated at their respective health centers. Besides, the project secured official approval from the National Ethical Review Committee before its commencement.

2.3.2.2 Ethical issues for chronically infected study participants

In this case, Linked Confidential testing with informed written consent was employed to collect blood specimen from chronically infected patients. Eligible patients were provided with consent form and asked to read and understand the content, which was prepared in Amharic and English (Annex I). The content of the consent contained complete information on the objectives of the study, the advantages and risks associated with taking part in the study, and the measures employed to maintain confidentiality of the information collected from them. When patients were found volunteer on their own will to participate, they were asked to sign on it. After obtaining the

written consent from the patient, specimen was collected by trained laboratory technologist. The specimen was then labeled with a code that could be linked to personal identifying information. Ethical clearance was obtained twice (once during the commencement of the project, and second time during renewal of the project's lifespan) from Institutional Review Board and the National Ethical Review Committee.

2.3.3. Specimen Collection and Storage

2.3.3.1 Collection, storage, and sero-screening of specimens from recently infected individuals

About 5 ml blood sample was drawn from each study participant by venipuncture in anticoagulant-free vacutainer tube. Serum was separated from whole blood within an hour time after blood draw by centrifugation at 400-800g for 10 minutes. One-milliliter leftover serum was aliquotted with micropipette or disposable Pasteur Pipette from the serum collected for RPR test into a cryotube and taken to EHNRI. Transportation of specimen from site of collection to EHNRI was carried out on the same day in cold boxes with freezer packs. All specimens were aliquotted further into two cryotubes of 0.5-ml, registered electronically, and then stored at -80^0C until used during genotyping. The eligible specimens were arranged consecutively according to date of blood draw, and transported frozen in dry ice for genotyping to Central Virology Laboratory (Tel-Hashomer, Israel). Sero-screening was performed at Addis Ababa Regional laboratory following the national algorithms as stated below (MOH/HAPCO, 2006). All specimens were tested with Vironostika® EIA (BioMerieux bv, Boxtel, The Netherlands) and the test results were recorded on provided data collection sheets. First tests from some sites were also performed at National HIV Reference Laboratory, EHNRI. Ten percent of randomly selected HIV negative and all HIV-positive specimens were re-tested using Enzygnost EIA (Dede Behring Marburg GmbH, Germany) at EHNRI for quality control purposes; Murex Antibody test (Abott

Murex, UK) was used as a tie breaker. HIV reactive specimens were re-tested at EHNRI. If confirmed, they were classified HIV positive; all other specimens were classified as HIV negative (MOH/HAPCO, 2006).

2.3.3.2 Collection and storage of specimens from chronically infected individuals

Sample collection from both drug-naïve and drug-experienced chronically infected patients was done following same procedure, except for preparation of dried blood spot (DBS) samples for the drug-naïve chronically infected group. For samples destined for genotyping and viral RNA quantification purposes, up to 10 ml whole blood was collected from each study participant by venipuncture in EDTA-treated vacutainer tube. In parallel, a 3-ml whole blood was drawn into EDTA-treated vacutainer tube from each patient for CD4+ T cell count and other hematological tests, and the tests were run on the same day. Plasma was then separated from blood cells within two hours in the laboratory where blood draw was performed by centrifugation at the speed of 1200g (3000 rpm) for 10 to 15 minutes in accordance with the recommendation of WHO (2009). The plasma was then aliquotted into 3 nunc tubes of 1.5-ml capacity, and transported on the same day to EHNRI, where it was stored at -80^0C until required for genotyping and other downstream applications. Those plasma samples which were required to be taken to Atlanta for actual genotyping purposes were packed while frozen and transported along with dry ice. Immediately on arrival (within 24 hours after packaging) they were unpacked and re-stored in ultra-deep freezer at -80^0C.

During collection of baseline blood sample from each chronically infected drug-naïve treatment eligible patients, a portion of whole blood was used to the preparation of DBS using the procedure described in Hallack *et al*. (2008). Fifty microliter EDTA-anticoagulated whole blood

was pipetted onto 12-mm premarked circles on 903 filter paper cards (Schleicher and Schuell, Keene, NH). Five such spots were made for samples collected from each patient. The cards were then left in a vertical position at room temperature for up to 24 hours to ensure their drying. Individual cards with the five spots were then packaged in a gas-impermeable bag Bitran® zipper-lock bags (Fisher Scientific Company, Pittsburgh, PA) containing a silica gel pack and humidity indicator (Mini Pax Sorbent; Multisorb Technologies, Buffalo, NY). The bags were stored at -20^{0}C until used in the analyses.

2.3.4. Specimen Processing and Performing Various Assays

2.3.4.1 Viral load assays and CD4 + T cells count

2.3.4.1.1 CD4+ T Cell Count

No attempt of CD4+ T cell count was made for recently infected drug-naive participant group. For both chronically infected drug-naïve and -experienced patient groups, CD4+ T cell count was done using FACScan Flow Cytometery (Becton Dickinson Immunocytometery System, San Jose, CA., USA) and FACSCount CD4/3 SW Version 1.0 08/05 flow cytometry system (Becton Dickinson Immunocytometry Systems [BD], San Jose, Calif.) at EHNRI and sites of blood draw, respectively, following the manufacturers' instructions.

2.3.4.1.2 Viral load determination for serum samples collected from recently infected participants

Viral load was determined using in-house real-time One-step RT-PCR Master Mix reagents (P/N 4309 169; Applied Biosystems, Roche Branchburg, New Jersey, USA) according to Burgard *et al.* (2000) on ABI Prism 7000 Sequence Detection System (Applied Biosystems, a division of Perkin -Elmer). The quality of the extraction was evaluated in terms of the DNA values of a single copy gene (RNaseP) (TaqMan RNase P Control reagents, P/N 4316844, Applied Biosystems, Foster City, CA 9404, USA) in the serum using a validated in-house realtime PCR

assay of the laboratory. Only samples containing the equivalent of >100 copies/ml RNA by the in-house real-time RT-PCR were subjected to genotyping.

2.3.4.1.3 Viral load assay for samples collected from chronically infected drug-naïve patients

For samples collected from chronically infected drug-naïve patients, plasma viral load was determined using Nuclisens, Easy Q HIV-1 assay (NASBA Diagnostics, BioMeriex, Botex) according to the protocol recommended by the manufacturer. Briefly, the input RNA was extracted using boom method (Boom *et al*., 1990). Amplification and quantification of the target RNA was then carried out using Easy Q HIV-1 viral load test kit, which has an internal calibrator for two different targets specific molecular beacons, one for wild type HIV-1 and one for the calibrator amplicon. The lowest detection limit of the assay was 50 copies/ml.

2.3.4.1.4 Viral load assay for samples collected from chronically infected drug-experienced patients

Viral load assay for plasma samples of chronically infected drug-experienced patients was done using RealTime HIV-1 assay, which has an internal Quality Control system, with automated Abbott *m*2000*rt*™ System (Abbott Molecular Inc., Des Plaines, IL 600018, USA) at International Clinical Laboratories (a private-owned laboratory located in Addis Ababa, the only Joint Commotion International accredited laboratory in East Africa) according to the manufacturer's instruction. However, sample preparation for the automated quantification assay was done manually following the manufacturer's instruction in the same laboratory by extracting and concentrating the target RNA molecules from 200 µL plasma to make the target accessible for amplification and to remove potential inhibitors of amplification from the extract. The upper detection limit of the assay as 10 million copies/ml while its lowest limit for the amount of input

sample used in this study was 150 copies/ml, although this could be lowered further to 40 copies/ml if the amount of input samples were at least 600 µL.

2.3.4.2 RNA Extraction, PCR Amplification, Amplicon detection, and purification

2.3.4.2.1 RNA extraction

Viral RNA was extracted from sera samples of recently infected participants using automated BioMeriex extractor (easyMAG 00102, firmware 1.0.8; BioMerieux, Lion, France). RNA extraction from chronically infected plasma samples was performed manually using QIAamp® Viral RNA Mini extraction kit (QIAGEN Group**,** QIAGEN Inc., 28159 Avenue Stanford, Valencia, CA 91355) according to the manufacturer's instruction. Briefly, a 140-µL already thawed plasma sample was added onto a 560-µl prepared Buffer AVL containing carrier RNA in a 1.5-ml microcentrifuge tube. The mixture was then incubated at room temperature for10 minutes followed by addition of 560 µL of ethanol (96–100%) and centrifuged at 6000xg (8000 rpm) for 1 minute. These processes allow lysis of the samples under highly denaturing conditions to inactivate RNases and to ensure isolation of intact viral RNA. After discarding the filtrate, buffering conditions were adjusted to provide optimum binding of the RNA to the QIAamp membrane by subsequently adding 500 µL of Buffer AW1, centrifuging at 6000xg (8000 rpm) for 1 minute, adding 500 µL of Buffer AW2, and centrifuging at full speed (20,000xg; 14,000 rpm) for 3 minutes. These procedures ensure cleaning of unwanted products and binding of the RNA to the membrane. Finally, the RNA was eluted to a clean 1.5-ml microcentrifuge tube by adding a 60 µL of Buffer AVE and incubating for 1 minute, and then centrifuging at 6000xg (8000 rpm) for 1 minutes.

Total Nucleic Acid (TNA) extraction from DBS samples was done manually using NucliSens Extraction Kit (bioMerieux, Inc., Durham, NC) as per the Standard Operation Procedure (SOP)

developed by modification of published assay (McNulty *et al*., 2007), validated by the HIV genotyping laboratory at HIV International Laboratory Branch, Division of HIV/AIDS, CDC, Atlanta, and endorsed by the CDC Atlanta on 24 March 2008 to be used for research purpose for HIV-1 Drug Resistance Genotyping Testing (unpublished document received along with the genotyping system). Briefly, a whole blood spot was excised from the DBS card with a surgical scissors and placed into NucliSens Lysis Buffer tube (Cat#: 284047 (50 x 9 ml) BioMerieux, Inc., Durham, NC). The tubes were placed on Twist Shaker (Cat#: EF4909B; FINEPCR® Twist Shakers, Daigger and Company, Inc.) for 2 hours at room temperature with shaking at 50 rpm, with periodic checking to ensure the suspension of the spots in the buffer during rotation. Then the tubes were centrifuged at 1200 rpm for 5 minutes followed by addition of 50 µL of pre-vortexed Silica suspension, vortexing, and placing back on the Twist Shaker for 10 minutes with shaking at 50 rpm to mix before transferring into new tubes. This was followed by repeated washing and centrifugation by adding Wash Buffer, 70% Ethanol, acetone, and treating the silica pellets in heat block at 56^0C for 10 minutes. Finally, 50 µL of Elution Buffer was added to the dried silica pepllet, resuspended by vortexing and incubated at 56°C for 10 minutes with mixing in a Thermomixer. It was then centrifuged at 10,000g for 2 minutes to get the nucleic acid eluted.

2.3.4.2.2 PCR amplification, amplicon detection, and purification

For samples genotyped with ViroSeq™ HIV-1 v2.0 genotyping kit (Celera Diagnostics), most of which were those samples obtained from recently infected and chronically infected drug-naïve participants, RNA reverse transcription of target regions, PCR amplification (both reverse transcription and nested PCR combined in one), and PCR product purification were done according to the protocol recommended by the manufacturer. The ViroSeq™ HIV-1 Genotyping System includes reagents and protocols for every step of genotyping from RNA extraction to

generation of a genotyping report. In the event of ViroSeq™ system failure, in-house primers were used according to Snoeck *et al.* (2005) with one-step RT-PCR kit from QIAGEN (cycling conditions: 50^0C for 30', 95^0C for 15''; 39 cycles of 94^0C for 30'', 50^0C for 30'' and 68^0C for 3'; 68^0C for 10' and 4^0C hold).

For samples genotyped by In-house system, RNA reverse transcription and nested PCR were done with SuperScript™ III one step RT-PCR systems with Platinum® *Taq* high Fidelity (Cat #: 12574-036, Invitrogen) and AmpliTaq Gold LD with 10X buffer II system (Cat#: 4338856, Applied Biosystems Applied Biosystems, 850 Lincoln Centre Drive, Foster City, California 94404) respectively following the manufacturers' protocol. Briefly, a 40-µL RT reaction mix was prepared from 25 µL 2X reaction mix, 1 µL each of the forward and reverse primers (a 2-mers forward primer [Pro Out 3Fv (F1)]: 5'- CCT CAA ATC ACT CTT TGG CAR CG-3' (spanning 2253-2275 in HXB2 genomic map; and a 23-mers reverse primer [RTgeg4R (R1)]: 5'- ATC CCT GCA TAA ATC TGA CTT GC-3' (spanning 3370-3348 in the genomic map of HXB2)), 12 µL of DEPC-treated Water, 1 µL of SS III RT-PCR enzyme, and 10 µL of denatured TNA (from DBS extract) or RNA (from plasma extract), and pre-heated at 65^0C for 10 minutes. The mixture was then loaded into the thermocycler which was pre-programmed to run in the following cycle conditions: one RT cycle consisting of 50^0C for 45', and 94^0C for 2'; forty PCR cycles of 94^0C for 15'', 55^0C for 20'' and 72^0C for 2'; one cycle of 72^0C for 10', and 4^0C hold. A 100-µL reaction mix for nested PCR was prepared from 10 µL 10X Buffer II, 8 µL $MgCl_2$, 2 µL dNTP (10 µM each), 3 µL each of the forward and reverse internal primers (a 24-mers forward primer[PAF1V (F2)]: 5'- CTT TGG CAA CGA CCC CTY GTC WCA-3' (spanning 2265-2288 in HXB2 genomic map; and a 23-mers reverse primer [215/219/3R (R2)]: 5'- CTT CTG TAT GTC ATT GAC AGT CC-3' (spanning 3326-3304 in HXB2 genome map), 69 µL of DEPC-treated Water, 1 µL Taq Gold LD, and 4 µL of RT-PCR products. The

reaction mix was then loaded into pre-warmed thermocycler and run with the following cycle program: one denaturation cycle of 94^0C for 4 minutes; forty cycles of 94^0C for 15'', 55^0C for 20'', 72^0C for 2'; and one cycle of 72^0C for 10' and 4^0C hold.

After running each reaction, presence and size of amplified products was checked through agarose gel electrophoresis following SOP at EHNRI, where a mixture of 10 μL PCR product and 2 μL loading dye was loaded into wells of 1% (w/v) agarose gel in which 5 μL Ethidium Bromide (EtBr) was added. The entire gel was covered with 1x TBE running Buffer, run for 45 minutes at 100V. The gel was then observed under UV light using imaging system (Gel Logic 200 Imaging System, Kodak Molecular Imaging Software Version 5.0; Carestream Health, Inc., 4 Science Park, West, New Heaven CT, 06511, USA), where amplification products fluoresced and the bands photographed. The amplification process was accepted when the experimental products and the positive control were within the right size range (1.06 kb) or fell just below the 1.2 kb marker, and no band from negative control was visible. Any amplification result was rejected and repeated if the positive and negative control results were valid and the patient sample did not produce a visible band. The intensity of the sample bands were compared to that of known bands on the ladder in order to choose an appropriate dilution factor for sequencing, in accordance to the recommendation of CDC's protocol (unpublished bench protocol).

When the gel reading was acceptable, purification of PCR products was conducted using QIAquick PCR Purification Kit (Cat# 28106, Qiagen, QIAGEN Inc., 28159 Avenue Stanford, Valencia, CA 91355) in accordance with the manufacturer's instruction. Briefly, one volume of PCR product was mixed in a 1.5-ml microfuge tube with 5 volumes of Buffer PBI of the kit and vortexed to mix.

This was then placed in a QIAquick spin column and centrifuged, washed with 0.75 ml buffer PE, and then eluted with 50 µL Buffer EB into a separate microfuge tube.

2.3.4.3 Cycle Sequencing

Sequencing of purified and properly diluted PCR products was done using BigDye™ Terminator Cycle Sequencing Kit v.3.1 (CAT#: 4337455, Applied Biosystems, 850 Lincoln Centre Drive, Foster City, California 94404) following the manufacturer's instruction. Both the ViroSeq™ HIV-1 v2.0 genotyping kit protocol and the In-house HIV-1 Drug Resistance Genotyping Testing protocol use same cycle conditions and same purification techniques, differing only slightly with the constituent reaction mix and primers used. In the ViroSeq™ HIV-1 v2.0 genotyping system, 12 µL pre-mixed BigDye™ terminator sequencing reagents containing either of the seven primers (A, B, C, D, F, G, and H) that are able to amplify and so sequence in both directions were mixed with 8 µL of the purified and properly diluted PCR product. In the In-house genotyping system, on the other hand, the reaction mix for each 20 µl reaction constitute 11.5 µL water, 2 µL (20 Pmol) either one of the six primers (A35: 5'-TTG GTT GCA CTT TAA ATT TTC CCA TTA GTC CTA TT-3'; AV36: 5'-CAG TAC TGG ATG TGG GTG ATG-3'; AV44: 5'-TAC TAG GTA TGG TAA ATG CAG T-3'; HIV90.R: 5'-AAT GCT TTT ATT TTT TCT TCT GTC AAT GGC-3'; PAF1V (F2): 5'- CTT TGG CAA CGA CCC CTY GTC WCA-3'; or 215/219/3R (R2): 5'- CTT CTG TAT GTC ATT GAC AGT CC-3'), 4 µL Big Dye Buffer, 2 µL Big Dye Terminator, and 0.5 µL diluted template PCR product. The 20 µL reaction mixes in both genotyping systems were mixed in a 96-well 3100 plate and loaded onto the thermocycler (GenAmp PCR System 2700 or GenAmp PCR System 9700, ABI Applied Biosystem, 850 Lincoln Centre Drive, Foster City, California 94404 USA) pre-programmed to the following cycling conditions: 25 cycles of 96^0C for 10'', 50^0C for 5'', 60^0C for 4'; and 4^0C hold. This cycle sequenced BigDye-labled DNA product was then purified using 80% isopropanol and centrifuged. Into the purified DNA fragments was added 20 µL

formamide, and loaded onto the ABI PRISM® 3100–Genetic Analyzer (ABI Applied Biosystems, 850 Lincoln Centre Drive Foster City, CA 94404 USA).

2.3.4.4 Sequence assembly and editing

Sequencing of the loaded DNA fragments with the ABI PRISM® 3100–Genetic Analyzer was carried out using the following sequencing parameters: DYE SET, Z; MOBILITY FILE, DT3100POP6 (BDV3)V1. mob; PROJECT NAME, 3100_project; RUN MODULE, StdSeq50_POPDefaultModule; and ANALYSIS MODULE, BC_3100POP6SR_SeqOffFtOff.saz. The ABI PRISM® 3100–Genetic Analyzer contains 3100 Software, called ABI PRISM® DNA Sequencing Analysis Software v3.7, which consists of, among others, ABI PRISM® 3100 Data Collection Software that, in conjunction with others, collects, processes, and stores data, automatically extracts and analyzes the data including base-calling after each run. The sequence analysis data collected by the Data Collection Software was assembled and copied to external drive and transferred to another computer work station for further analyses including base-calling, manual editing, and consensus sequence constructing using ViroSeq HIV-1 Genotyping System software (version 2.5) for samples genotyped with ViroSeq and ChromasPro version 1.42 software (Technelysium Pty. Ltd.) for samples genotyped with In-house genotyping methods.

2.3.5 Sequence Analyses

The edited sequences were converted and saved in FASTA format for further analyses using other bioinformatics software depending on the purposes of the analyses.

2.3.5.1 Identification of drug resistance mutations

Drug resistance mutations were identified by entering the FASTA-formatted sequence data into the Stanford HIV RT and Protease Sequence Database, Stanford University (http//www.hivdb.stanford.edu/hiv.). Resistance report for isolates from recently infected

participants was produced for PIs, NNRTIs and NRTIs by comparing the list of drug resistance mutations identified by the Stanford genotyping system with those on the 2009 WHO transmitted drug resistance surveillance list, which included 34 protease mutations at 15 positions for PI resistance, 40 RT mutations at 18 positions for NRTI resistance, and 19 RT mutations at 10 positions for NNRTI resistance (Benett *et al*., 2009). On the other hand, resistance reports for isolates from chronically infected drug-naïve and –experienced participants were produced by comparing the list of drug resistance mutations identified on the November 6, 2009 Stanford Drug Resistance Database update (available at: http://hivdb.stanford.edu/pages/download/) by the Stanford genotyping system, which comprised of 21 major protease mutations at 17 positions for PI resistance, 23 major RT mutations at 18 positions for NRTI resistance, and 22 major RT mutations at 15 positions for NNRTI resistance, and two other genotyping algorithms: namely, the 2009 WHO transmitted drug resistance surveillance list [comprising 34 protease mutations at 15 positions for PI resistance, 40 RT mutations at 18 positions for NRTI resistance, and 19 RT mutations at 10 positions for NNRTI resistance] (Benett *et al*., 2009), and the 2009 IAS-USA list of both major and minor drug resistance mutations [comprising 75 protease mutations at 38 positions for PI resistance, 20 RT mutations at 15 positions for NRTI resistance, and 25 RT mutations at 14 positions for NNRTI resistance] (Johonson *et al*., 2009).

2.3.5.2 Determination of performance of DBS and In-house genotyping systems

Comparison of the performance of DBS and In-house genotyping system with the conventional plasma and the commercial kit ViroSeq™ was done using the freely available BioEdit version 7.0.9.0 (Hall, 1999). These comparisons were done between data of nucleotide sequence generated from pair-wise DBS/plasma, and those generated from paired In house and commercial kit, where per cent similarities were calculated and phylogenetic trees were constructed using the

aligned nucleotide sequences by the same software. Relative performances of DBS and In-house genotyping system against the conventional plasma and the commercial kit ViroSeq™, respectively, were calculated by comparing amino acid mutations detected first separately for protease and reverse transcriptase and then pooled together. Relative sensitivity was calculated using the formula (Loong, 2003): 100 X True positive (TP) (i.e. no. of mutations detected by both test and standard assays)/ TP + False Negatives (i.e., no. mutations not detected by the test assay but detected by the standard assay). Relative positive predictive values were calculated by multiplying the ratios of number of all mutations detected by the test assay to number of all mutations detected by the standard assay by 100 (Loong, 2003). However, the relative specificity of the test assays could not be determined because the negative values (that is, the number of all mutations that would be missed by both the standard and the test assays) could not be correctly determined.

2.3.5.3 Sequence alignment and genotyping

For subtyping and determining phylogenetic relationships between the various isolates, the FASTA-formatted nucleotide and amino acid sequences were aligned using MEGA 4.0.2 (Tamura *et al*., 2007). These aligned nucleotide and amino acid sequences were used for various downstream analyses. As a sequence quality control measure to detect potential contamination during amplification and cycle sequencing, sequence products were analyzed using phylogenetic tree construction by neighbor-joining method. Sequences with 100% similarities were excluded from further analyses.

2.3.5.3.1 Subtype determination and phylogenetic analyses

Subtype assignment to the aligned sequences was performed on protease (PR) and reverse transcriptase (RT) separately, and on merged PR and RT sequences as Pol using three online

software; Stanford HIV-1 Drug Resistance Database (http://hivdb.stanford.edu/), REGA HIV subtyping tool (Leuven University, Leuven, Belgium; http://www.bioafrica.net/subtypetool/html/) jumping profile Hidden Markov Model (jpHHM) at GOBICS (http://jphmm.gobics.de/); and one offline software Phylogenetic analyses using MEGA 4.0.2 (Tamura *et al*., 2007). When discrepancies occurred between results of these four software, as is often the case when subtyping non-B clade (Holguin *et al*., 2008b; Parczewski *et al*., 2010), particularly when sign of recombination was detected, SimPlot Version 3.2 (Lole *et al*., 1999) by comparing the query sequences to the full set of the 9 reference sequences with 100 bootstrap replications.

Appropriate model for phylogenetic analyses of the current data set was sought using FindModel software available at the Los Alamos National Library (http://www.hiv.lanl.gov/), and the software suggested JC: Jukes-Cantor (model 1) to be the most appropriate for this data set, and this was used in the analyses. After gap stripping, phylogenetic analyses were made using default parameters of the MEGA 4.0.2: namely, distances between the sequences were calculated using the Kimura two-parameter model as an optimal substitution model with a transition-to-transversion ratio of 1.5; Neighbor-joining was used to create phylogenetic trees with 1000 bootstrap replications.

2.3.5.4 Determination of Sequence Diversity

Genetic diversity of isolated viral variants from the three groups of participants was tested using SNAP (Synonymous/Non-synonymous Analysis Program) and Shannon entropy test both available at the Los Alamos National Laboratory (http://www.hiv.lanl.gov/content/sequence/SNAP/SNAP.html). The principles behind the SNAP analysis is that non-synonymous mutations to a DNA sequence cause a change in the translated amino acid sequence, whereas synonymous mutations do not. The comparison between the number of non-synonymous mutations (dn), and

the number of synonymous mutations (ds), can suggest whether natural selection is acting to promote the fixation of advantageous mutations (positive selection) or to remove deleterious mutations (purifying selection). In general, when positive selection dominates, the dn/ds ratio is greater than 1, implying that diversity at the amino acid level is favored probably because of the fitness advantage provided by the mutations. On the other hand if the dn/ds ratio is less than 1, negative selection dominates it means that most amino acid changes are deleterious and, therefore, are selected against. When the positive and negative selection forces balance each other, the dn/ds ratio is close to 1 (http://www.hiv.lanl.gov/; Korber, 2000). During test of Shannon entropy, sequence variability or diversity between one group relative to the other was done by assessment of position-specific measure of variations between the different sequence groups. The tests were carried out on nucleotide and amino acid sequence alignments using the default parameters (100 randomization with replacement, and p-value of 0.05) (http://www.hiv.lanl.gov/).

2.3.6 Statistical analyses

Demographic and clinical data recorded from each patient logbook (age, gender, time of HIV diagnosis, CD4 T-cell counts, HIV viral load, WHO immune and clinical staging and treatment status and time of treatment) were compiled for each patient when appropriate. Drug resistance mutations for PI, NRTI, and NNRTI were characterized by their frequency. Pearson's correlation was calculated to find out co-variation of drug resistance mutations within and between classes of drugs, and to determine if any correlation existed between the drugs patients that patients are taking (NVP vs EFV; D4T vs AZT vs TDF) and resistance mutations selected by the drugs. Categorical data were compared using χ^2 test. Continuous data were investigated using Student t

test. Significant level of 0.05 was used in all relevant statistical calculations. All analyses were performed using SPSS version 15.0 (SPSS Inc. 1989-2006).

CHAPTER 3: RESULTS

3.1 Profiles of the Study Populations and the Data Sets

3.1.1 Recently infected drug-naïve participants

A total of 2432 samples collected for the national ANC-based HIV surveillance between 6th of April and 8th of August 2005 were considered for the HIV drug resistance threshold survey (HIVDR-TS). Sero-prevalence determination was done both at Addis Ababa Regional Laboratory and EHNRI. The tests have reported that 330 of the 2432 (13.6%) collected samples were confirmed HIV-positive. However, because of poor specimen handling and storage and ambiguity in registering demographic data at one of the sites (Federal Armed Forces Hospital), the entire samples from this site were disregarded from the analysis in this study. Therefore, only 2196 specimens were available for consideration in this study (Table 3.1.1). Three hundred of these samples were found HIV-positive. Demographic data from 11 of the 300 HIV-positive samples were incomplete, and thus only 289 (96.33%) of them were usable samples. The minimum, maximum and mean age of these participants were 16, 40, and 24.97 (with SD= 4.58), respectively. Of these samples, 131 specimens were obtained from women of age less than 25 regardless of number of pregnancies, and 112 from women with no previous pregnancy regardless of age. However, only 77 (26.64%) of them fulfilled the survey criteria (i.e., age < 25, first pregnancy, and residence in Addis Ababa). Even so, not all of the 77 eligible specimens were used in the genotyping. Preliminary RNA load determinantion showed that 14 of these 77 samples failed to contain any detectable RNA, leaving only 63 specimens for amplification (Table 3.1.2). Of the latter, 41 (65.1%) specimens were managed to get amplified and genotyped by ViroSeq™, TrueGene™, and In-house Genotyping systems, while all the rest failed.

Table 3.1.1 Demographic data and specimen conditions considered in the 2005 HIV drug resistance Threshold Survey (HIVRD-TS).

Site	No. of Samples collected	No. of HIV-positive individuals	No. of usable samples	Age range of people with usable samples		No. of previous pregnancies		No. of Eligible samples for HIVDR-TS
				< 25	≥ 25	0	≥1	
T/Haimanot Clinic	326	35	34	19	15	9	25	6
Kazanchis Clinic	349	67	67	29	38	28	39	21
Kefitegna 23 Clinic	443	47	38	18	20	18	20	13
Gulele Clinic	393	56	55	28	27	24	31	18
Akaki Clinic	435	40	40	25	15	20	20	15
Federal Police Hospital	250	55	55	11	44	13	42	4
Total				131	159	112	177	77
Grand total	2196	300	289		289		289	77

Table 3. 1.2 Measurements and tests performed to select for HIVDR-TS genotyping from eligible samples

Site	No. of Eligible samples for HIVDR-TS	RNA copies/ml				Attempted amplification	Successfully Genotyped
		>2000	1000-2000	100-1000	Undetectable		
T/Haimanot Clinic	6	3	1	1	1	5	4
Kazanchis Clinic	21	6	2	6	7	14	9
Kefitegna 23 Clinic	13	4	0	6	3	10	4
Gulele Clinic	18	9	2	6	1	17	13
Akaki Clinic	15	10	1	3	1	14	11
Federal Police Hospital	4	1	0	2	1	3	0
Total (only vertical)	77	33	6	24	14	63	41

3.1.2 Chronically infected drug-naïve participants

Total of 104 chronically infected participants who were eligible, and hence, recommended to start HAART, were recruited in this baseline study. Although paired DBS/plasma samples were collected from all of the 104 participants, only 85 (81.73%) of the samples (either DBS or plasma or both) were amplified and genotyped. Therefore only these 85 participants were considered for every analysis in this study. While 39 (45.9%) of these participants were females, 46 (54.1%) were males (Table 3.1.3). The median and mean age of the participants were 32 and 33.91 (SD=

8.70), respectively (Figure 3.1.1). Whereas a single age group (29-39) accounted for 36/85 (42.35%) of all the participants, a little over 75% of the participants were below the age of 40 years (Table 3.1.3). Regarding disease stages of the participants, recorded data showed that 28/39 (71.79%) of the women and 35/46 (76.09%) of the men participants were in the advanced WHO clinical stages (stages III or IV), with an overall 63/85 (74.12%) of the participants being in these two stages. Eleven participants from each women and men participants belonged to WHO stage I and II.

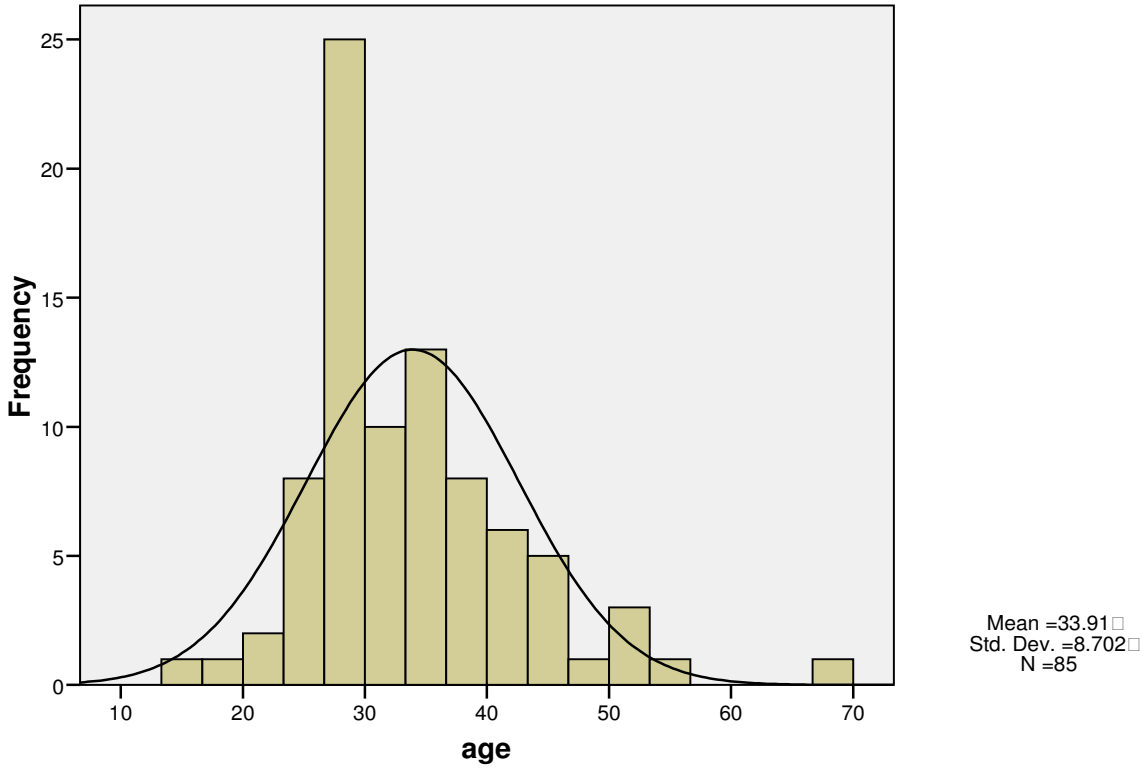


Figure 3.1.1 Age distribution of chronically infected drug naïve but treatment eligible participants

Table 3.1.3 Age distribution of chronically infected drug naïve treatment eligible participants by gender and baseline WHO stages

Gender	WHO staging		age group						Total
			<= 18	19 - 28	29 - 39	40 - 49	50 - 60	61+	
Female	I	Count		0	0	0	1		1
		% of Total		.0%	.0%	.0%	2.6%		2.6%
	II	Count		4	3	2	1		10
		% of Total		10.3%	7.7%	5.1%	2.6%		25.6%
	III	Count		9	7	5	0		21
		% of Total		23.1%	17.9%	12.8%	.0%		53.8%
	IV	Count		4	3	0	0		7
		% of Total		10.3%	7.7%	.0%	.0%		17.9%
	Total	Count		17	13	7	2		39
		% of Total		43.6%	33.3%	17.9%	5.1%		100.0%
Male	I	Count	0	0	2	1	0	0	3
		% of Total	.0%	.0%	4.3%	2.2%	.0%	.0%	6.5%
	II	Count	0	0	7	1	0	0	8
		% of Total	.0%	.0%	15.2%	2.2%	.0%	.0%	17.4%
	III	Count	1	5	9	6	2	0	23
		% of Total	2.2%	10.9%	19.6%	13.0%	4.3%	.0%	50.0%
	IV	Count	0	5	5	1	0	1	12
		% of Total	.0%	10.9%	10.9%	2.2%	.0%	2.2%	26.1%
	Total	Count	1	10	23	9	2	1	46
		% of Total	2.2%	21.7%	50.0%	19.6%	4.3%	2.2%	100.0%
Grand total		Count	1	27	36	16	4	1	85
		% of Total	1.18%	31.76%	42.35%	18.82%	4.71%	1.18%	100%

3.1.3 Chronically infected drug-experienced participants

Total of 266 participants currently on ARV treatment for at least 24 weeks were recruited in this study, except for one case in which the duration of treatment was 16 weeks. Gender distribution was 120 (45.1%) and 142 (53.4%) for females and males, respectively, with 4 (1.5%) missing data. The mean and median ages of the participants were 36 and 37.76 years (N= 255 Range=54; SD= 9.822), respectively. While the youngest participant was 19 years old, the older one was 73. The majority of the participants fell in the age range of 36-60 (50.2%) followed by 26-35 (38.0%), the two groups making 98.4% of all the participants. Of the 254 participants in this

study group for whom data on WHO clinical staging were gathered successfully, 134 (52.8%) and 80 (31.5%) were in stages III and IV, respectively (Data not shown). Therefore, participants in the two stages comprised of 214 (84.25%). With regard to duration of infection, the minimum duration was 14 months while the maximum one was 168 months (12 years), with median 52.00 months (Range= 36) mean duration of 55.38 months (N= 258; Range= 154; SD= 25.18 months), showing the degree of chronicity of the infection. One hundred and thirty-five (more than 50%) of the participants were infected for more than 4 years (Data not shown). Nearly 60% of all men and 58% of all women participants have been sero-positive for over 4 years.

Complete data regarding length of ARV treatment was available for 262/266 (98.50%) participants. The minimum duration of treatment was 4 months (about 16 weeks), whereas the longest one was 116 months (over 9-and-half years). Median treatment duration was 40.00 (Range= 112) and mean 42.09 months (N= 262; SD= 17.33). More than 250 (95.42%) of the participants were on treatment for more than a year (Table 3.1.4). Total of 86 participants (29 women and 57 men) were on treatment for more than four years. Moreover, 221 (84%) participants were on treatment for at least three years, and 158 (59.4%) of the total participants were treated for four or above years of duration. Most of the women participants (about 58%) were on treatment for 2-4 years, while most of the men (about 66%) were on over three years treatment period.

Table 3.1.4 Distribution of duration of ARV treatment by gender

Gender		Duration in months treatment						Total
		<= 12	13 - 18	19 - 24	25 - 36	37 - 48	49+	
Female	Count	2	6	12	28	39	29	116
	% within Gender	1.7%	5.2%	10.3%	24.1%	33.6%	25.0%	100.0%
	% within Duration of treatment	66.7%	46.2%	75.0%	44.4%	54.2%	33.7%	45.8%
	% of Total	.8%	2.4%	4.7%	11.1%	15.4%	11.5%	45.8%
Male	Count	1	7	4	35	33	57	137
	% within Gender	.7%	5.1%	2.9%	25.5%	24.1%	41.6%	100.0%
	% within Duration of treatment	33.3%	53.8%	25.0%	55.6%	45.8%	66.3%	54.2%
	% of Total	.4%	2.8%	1.6%	13.8%	13.0%	22.5%	54.2%
Total	Count	3	13	16	63	72	86	253
	% of Total	1.2%	5.1%	6.3%	24.9%	28.5%	34.0%	100.0%

Data on initial ART regimens were available for 261 participants, of whom 120 (almost 46%) were women while the rest 141 (about 54%) were men. One hundred and sixty-three (62.45%) started with ART having nevirapine (NVP) as background NNRTI, whereas the rest 98 (37.55%) were on Efavirenze (EFV)-containing regimen (Figure 3.1.2). Drug usage by gender showed that most women (67%) were on NVP-containing regimens as opposed to 33% on EFV-containing regimens. Similarly, 59% of men were on NVP-containing regimens as opposed to 41% EFV-containing regimens. Whereas the majority of the participants (206/261 [78.9%]) still stayed with their initial prescription, 55 (21.1%) have changed into a different combination due to several factors, poor virological response being among them.

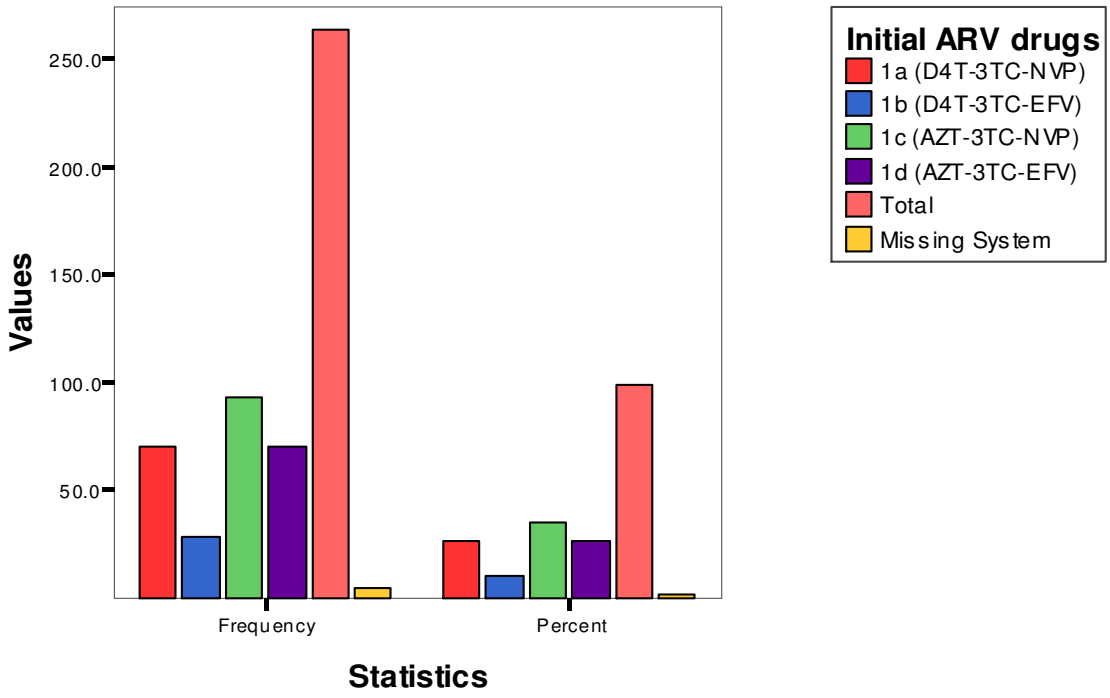


Figure 3.1.2 Distribution of 1st line ARV regimens used in the initial treatment of study participants.

Among the participants who changed regimen, 39/261 (14.94%) switched to a different 1st line regimens, while the rest 16/261 (6.13%) started 2nd line regimens including 3TC-TDF-kaletra (14/261 [5.36%]), and ABC-TDF-kaletra (2/261 [0.77%]), (Figure 3.1.3). The newly replacing 1st line drug regimens to which persons failing with initial treatment were switched included: 3TC-TDF-EFV (28/261 [10.73%]), 3TC-TDF-NVP (9/261 [3.45%]), 3TC-AZT-ABC (1/261 [0.38%]), and TDF-ddI-NFV (1/261 [0.38%]).

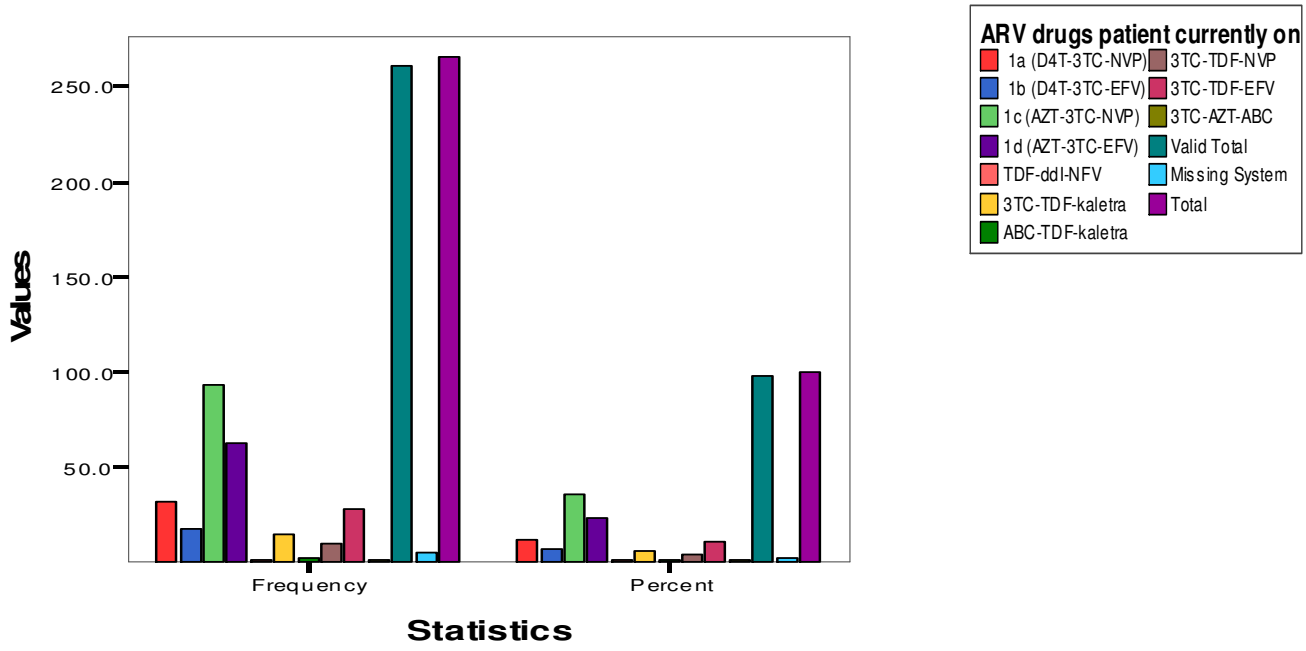


Figure 3.1.3 ARV regimens used currently in the treatment of drug-experienced study participants

3.2. Findings from Clinical Laboratory Measurements

3.2.1 Chronically infected drug-naïve participants

Baseline plasma RNA and baseline CD4+ T cell count were conducted for this group of participants. All the 85 participants with valid genotyped samples had their plasma viral RNA load determined successfully. The minimum value detected was 15,000 copies/ml while the maximum one was 2,500,000 copies/ml. Median and mean values were 230,000 (Range= 2,485,000 copies/ml) and 421,258.82 copies/ml (N=85; SD= 529944.363), respectively. The majority of the participants (59 [69.4%]) had values over 100,000 copies/ml (Table 3.2.1). Likewise, $CD4^+$ T cell counts determination was successful for 84 of the 85 samples. The minimum and maximum values were 4 and 289 cell/µL, respectively. The median and the mean values were almost similar: while the median count was 103.50 cell/µL (Range= 285 cell/µL),

the mean value was 105.95 cell/μL (N=84; SD= 66.783 cell/μL). About 23% of the participants had CD4+ count of 50 cell/μL or below (Table 3.2.2). Whereas nearly 92% of them had values below or equal to 200 cell/μL, none had count above 300 cell/μL. Analyses of paired plasma viral load and CD4+ T cell count tests have shown that 32/84 (38.10%) study participants with complete data for all three variables had viral load over 100,000 copies/ml. Over 96% (31/32) of these or over 53% (31/58) of all those belonging to the highest viral load group had CD4+ T cell count $\leq$ 200 cell/μL.

Table 3.2.1 Baseline plasma viral load measurement results for chronically infected drug-naïve participants

Groups of measurement (copies/ml)		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	10001 - 50000	13	15.3	15.3	15.3
	50001 - 100000	13	15.3	15.3	30.6
	100001+	59	69.4	69.4	100.0
	Total	85	100.0	100.0	

When gender was considered in the analysis, it was demonstrated that the highest plasma viral load was observed among women with $CD4^+$ count below or equal to 100 cells/μL and who were in the WHO stage IV (Figure 3.2.1). Among male participants, the highest plasma viral load was observed in those with $CD4^+$ counts between 101-200 cells/μL and who were in WHO stage III. In both women and men participants, all those in WHO stage I aggregated into $CD4^+$ count group 101-200 cells/μL, although the mean viral load was greater for men than women (Figure 3.2.1).

Table 3.2.2 Baseline CD4 Cell count measurements for chronically infected drug-naïve participants

CD4+ T cell count (cell/µL)		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	<= 50	19	22.4	22.6	22.6
	51 - 51	2	2.4	2.4	25.0
	52 - 101	21	24.7	25.0	50.0
	102 - 201	35	41.2	41.7	91.7
	202 - 301	7	8.2	8.3	100.0
	Total	84	98.8	100.0	
Missing	System	1	1.2		
Total		85	100.0		

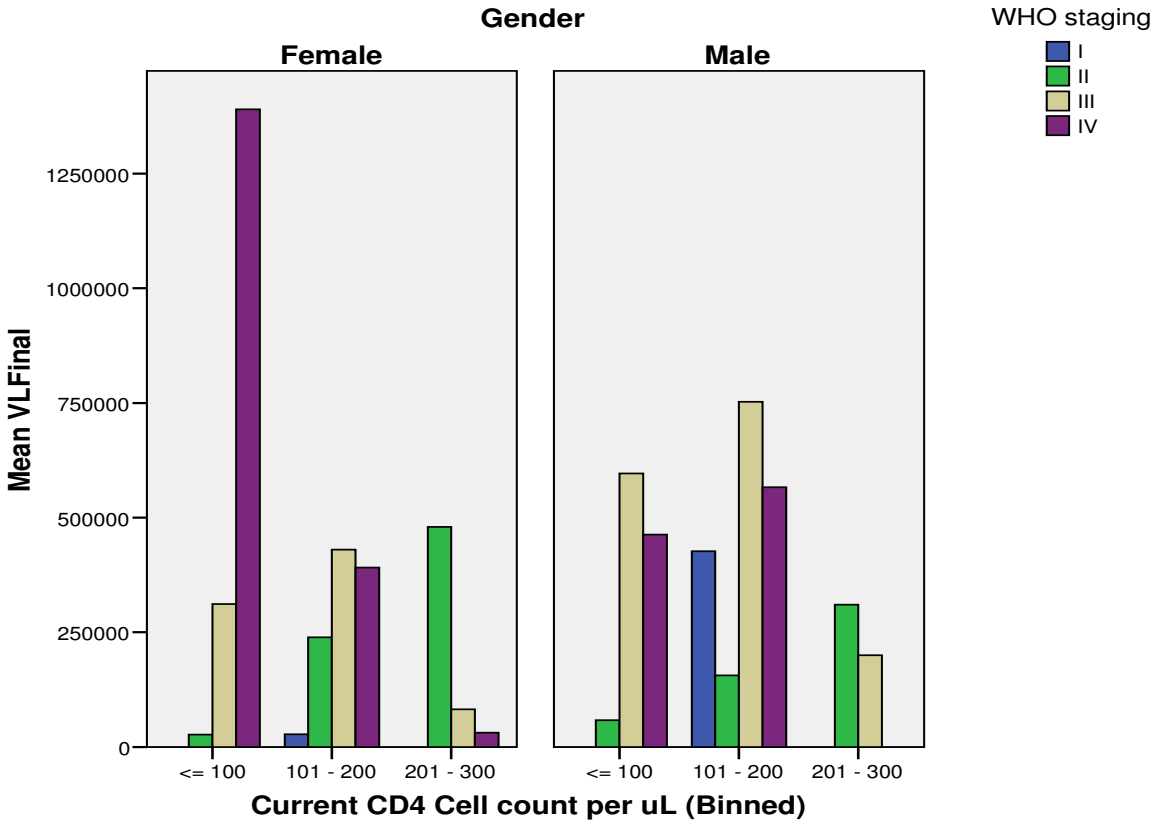


Figure 3.2.1 Distribution of plasma viral RNA by gender, CD4+ T cell count, and baseline WHO clinical staging among chronically infected drug-naïve participants

3.2.2 Chronically infected drug-experienced participants

A total of 266 drug experienced participants were considered in this study. Plasma viral RNA load was performed for all of these cases, of whom only 83 (31.2%) have shown some indication of RNA detection, with over 150 copies/ml. The minimum and maximum values were 158 and 6,408,783 RNA copies/ml, respectively. The median viral load was 14,294.50 (Range= 6,408,625 copies/ml). Over 55% the participants with detectable plasma viral RNA harbored over 10,000 copies/ml (Table 3.2.3). The most frequent (over 27%) RNA group was the one with the highest viral load (over 100,000 copies/ml) followed by the group with 1000 to 10,000 copies/ml (over 25%), whereas the least frequent one (12.0%) belonged to the group having next of the highest viral load group (between 50,000 and 100,000 copies/ml). In regard to virological failure in relation with length of treatment, over 88% of all virologically failing population or 64/252 (25.40%) of the whole study population were on treatment for over 24 months.

Table 3.2.3 Current plasma RNA load (copies per ml) of chronically infected drug-experienced study participants

Plasma viral RNA load groupings (copies/ml)		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	151 - 1001	16	19.3	19.3	19.3
	1002 - 10001	21	25.3	25.3	44.6
	10002 - 50001	13	15.7	15.7	60.2
	50002 - 100001	10	12.0	12.0	72.3
	100002+	23	27.7	27.7	100.0
	Total	83	100.0	100.0	

Failure rates of 1st line ARV drug regimens used in the initial treatment as measured by number (%) of participants with detectable plasma RNA against number of participants who were using the regimen were computed from 261 study participants (Table 3.2.4). The highest rate of failure was observed among those treated with 1b (D4T-3TC-EFV) (14/28 [50.0%]); followed by 1a

(D4T-3TC-NVP) (25/70 [35.7%]); 1d (AZT-3TC-EFV) (22/70 [31.4%]); and 1c (AZT-3TC-NVP) (20/93 [21.5%]).

Table 3.2.4 Virological treatment response by initial drug regimens among drug-experienced participants

Virologic response	Measurement	Initial ARV drugs*				Total
		1a	1b	1c	1d	
Undetectable plasma RNA	Count	45	14	73	48	180
	% within Plasma Viral detection	25.0%	7.8%	40.6%	26.7%	100.0%
	% within Initial ARV drugs	64.3%	50.0%	78.5%	68.6%	69.0%
Detectable plasma RNA	Count	25	14	20	22	81
	% within Plasma Viral detection	30.9%	17.3%	24.7%	27.2%	100.0%
	% within Initial ARV drugs	35.7%	50.0%	21.5%	31.4%	31.0%
Total	Count	70	28	93	70	261
	% within Plasma Viral detection	26.8%	10.7%	35.6%	26.8%	100.0%
	% within Initial ARV drugs	100.0%	100.0%	100.0%	100.0%	100.0%

*1a= D4T-3TC-NVP; 1b= D4T-3TC-EFV; 1c= AZT-3TC-NVP; 1d= AZT-3TC-EFV

Results for current $CD4^+$ T cell count was available for 207 (77.8%) of the 266 participants (Table 9.2.5). The minimum and maximum counts were 9 and 848 cells/μL, while the median and mean counts were 214 cells/μL (Range= 839) and 237.77 cells/μL (N= 207; SD= 146.274), respectively. The majority of these participants (138 [~67%]) had $CD4^+$ count between 100 and 350 cell/μL. Whereas 31 (15%) participants fell in the critically lowest $CD4^+$ count limit ($\leq$100 cells/μL), only 12 (5.8%) fell on immunologically safe range ($\geq$500 cells/μL).

Table 3.2.5 Current CD4 Cell count per μL for chronically infected drug-experienced participants

Current CD4 cell count groupings (cells/μL)		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	<= 50	8	3.0	3.9	3.9
	51 - 101	23	8.6	11.1	15.0
	102 - 201	64	24.1	30.9	45.9
	202 - 351	74	27.8	35.7	81.6
	352 - 501	26	9.8	12.6	94.2
	502+	12	4.5	5.8	100.0
	Total	207	77.8	100.0	
Missing	System	59	22.2		
Total		266	100.0		

Of the 83 samples with detectable plasma viral load, it was possible to produce CD4+ cell count results for 74 only. For ease of analysis, both plasma viral RNA and CD4+ cell count results were grouped into five and six categories, respectively. Analyses of paired plasma RNA-CD4+ count tests have shown that three $CD4^+$ count groups (namely 51-101, 102-201, and 201-351 cells/μL) contained all the five plasma viral load groups; viral load categories 1002-10001, 10002-50001, and 50002-100000 copies/ml were represented by more than 20% in each of the three CD4+ groups mentioned. The most frequent (over 50%) of plasma viral load range (10,000-50,00 copies/ml) fell within CD4+ count group of 202-351 cells/μL. Surprisingly, the smallest possible detectable viral load group (150-1001 copies/ml) was represented in over 10% of the smallest possible CD4+ group ($\leq$ 50 cells/μL). Likewise, the largest possible viral load group (i.e., over 100,000 copies/ml) was represented in more than 10% of the highest CD4 count group attained in these participants ($\geq$ 502 cells/μL). The great majority of the lowest viral load counts (150-1001 copies/ml) were found in persons with CD4+ count of $\leq$ 350 cells/μL. Figure 3.2.2 summarizes percentage of plasma viral load measurements as the function of CD4+ cell counts.

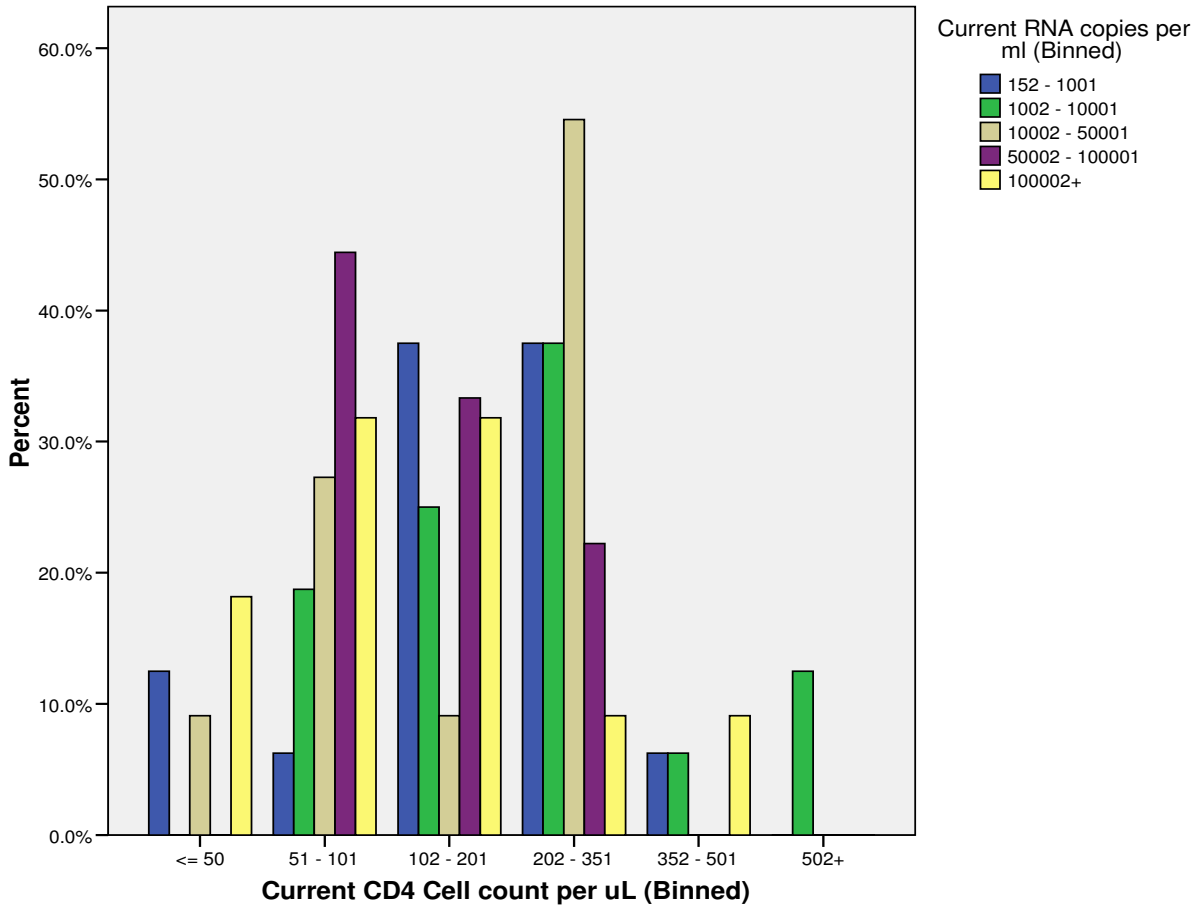


Figure 3.2.2 Distribution of ranges of $CD4^+$ cell count (cell/μL) among chronically infected drug-experienced participants with detectable plasma viral RNA

3.3 Genotyping

3.3.1 Performance of in-house-brewed primers for amplification and sequencing HIV-1 partial Pol region

A total of 52 paired samples (40 plasma from chronically infected drug-naive and 12 sera from recently infected drug-naïve samples) were amplified and sequenced at partial Pol region using either the commercial ViroSeq™ V2.0 genotyping system and home-brewed drug resistance genotyping system donated from the CDC (Atlanta). The In-house system used here was an adopted protocol from a published work by McNulty *et al*. (2007), the primers, reagents and operational procedures of which were optimized to amplify and genotype DBS samples from HIV-1 subtype B, although some of the primers were developed and validated earlier to have the

capacity to amplify and genotype several subtypes other than B (Snoek *et al*., 2005). All plasma/sera samples (100%) for which ViroSeq^{TM}-genotyped counterpart samples were available were successfully amplified by this In-house genotyping system (Figure 3.3.1). Although the system was to have a sensitivity of amplifying RNA samples with viral load of 1000 copies/ml (McNulty *et al*., 2007), it was possible to amplify samples below this limit; as low as 490 copies/ml and as high as 6.5 million RNA copies/ml were amplified using samples from drug-experienced patients in another part of this study. Moreover, all, except one, of the sequencing primers used were able to sequence their respective region very well. The exception in this regard was HIV90V, which repeatedly produced sequences of inferior quality.

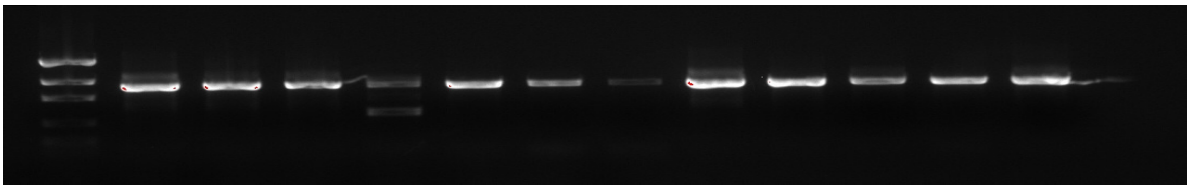


Figure 3.3.1 RT-PCR amplification products of HIV-1 plasma vral RNA using Inhouse genotyping system.
The bands at the left flank are DBNA size makers, while the rest are from the samples.

Panel A

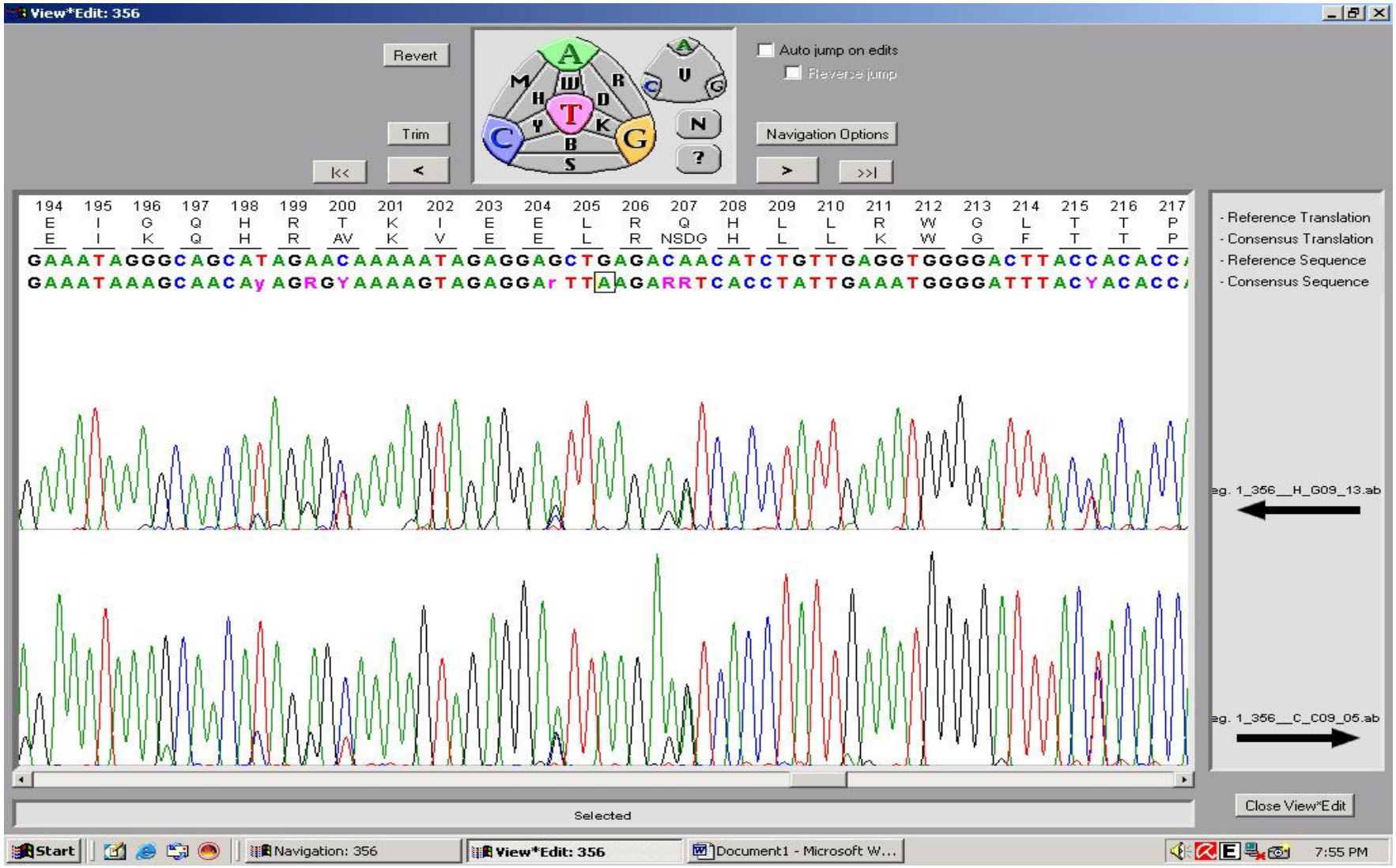

View*Edit: 356
Revert
Trim
<<
<
Auto jump on edits
Reverse jump
Navigation Options
>
>>|
- Reference Translation
- Consensus Translation
- Reference Sequence
- Consensus Sequence
eg. 1_356__H_G09_13.ab
eg. 1_356__C_C09_05.ab
Selected
Close View*Edit
Start
Navigation: 356
View*Edit: 356
Document1 - Microsoft W...
7:55 PM

Figure 3.3.2 Screenshots from contig Chromatograms constructed from sequences produced by ViroSeq assay and genotyping software (Panel A) and In-house assay produced by ChromasPro software (Panel B)

Originated from same patient (baseline 356). Codone 207 in the ViroSeq (Panel A) is represented by nucleotides 892-894 in the In-house assay (Panel B). The codon represented here is Q207 from the RT; four mixed quasispeicies have been detected in both assays containing D/G/N/S. Similarly, codon 200 in the ViroSeq and nucleotides 870-872 in the In house assay contained mixturres of viruses that have mutated from amino acid T to A and V.

After editing the sequences produced by the two systems manually (Figure 3.3.2A and B), FASTA-formatted consensus sequence of the six primer sequences genotyped by the In-house system was paired with its counterpart genotyped by ViroSeq™ system. Pairwise alignment by ClustalW multiple alignment program in BioEdit Version **7.0.9.0** software and computation of mean nucleotide similarity has shown that the mean concordance between the two systems at the level of Pol nucleotide sequence was 98.72%, the minimum and maximum concordance being 96.63% and 99.88%, respectively (Table 3.3.1). Separate protease and reverse transcriptase gene analysis has also shown almost similar results. Average protease nucleotide sequences, concordance was 98.37% while the minimum and maximum similarities were 95.60% and 100%, respectively. Similarly, average concordance for reverse transcriptase was 98.70%, whereas the minimum and maximum values were 96.33% and 99.88%, respectively.

Phylogenetic analyses with Neighbor-Joining/UPGMA method version 3.6a2.1 of BioEdit version 7.0.9.0 have shown that all the paired sequences branched together (Figures 3.3.3A and B). In addition, the performance of In house genotyping system relative to that of the ViroSeq™ V2.0 genotyping at the amino acid level showed that In-house system had relative sensitivities of 92.16%, 95.5%, and 94.26% for the regions of protease, reverse transcriptase and pooled protease and reverse transcriptase, respectively (Table 3.3.2). Relative positive predictive values were also 99.41%, 99.51%, and 99.47% for the regions of protease, reverse transcriptase and pooled protease and reverse transcriptase, respectively.

Table 3.3.1: List of ViroSeq Vs In house nucleotide sequence similarities as determined by BioEdit Version 7.0.9.0

S No.	Sample ID	VL copies/ml	Pol sequence % similarity	Protease sequence % similarity	Reverse transcriptase % similarity
1	Baseline 102	31,000	97.84	98.53	97.64
2	Baseline 115	290,000	96.99	97.01	96.59
3	Baseline 154	180,000	99.52	99.63	99.08
4	Baseline 155	34,000	99.52	99.23	99.61
5	Baseline 156	89,000	99.04	99.27	99.74
6	Baseline 157	150,000	98.20	97.44	98.07
7	Baseline 165	200,000	98.56	98.90	98.08
8	Baseline 175	310,000	98.20	99.63	97.93
9	Baseline 183	290,000	97.96	99.27	97.44
10	Baseline 190	310,000	98.80	98.69	99.08
11	Baseline 197	310,000	98.08	96.34	98.43
12	Baseline 200	290,000	97.48	95.60	97.66
13	Baseline 203	2,400,000	99.88	99.23	99.61
14	Baseline 204	110,000	98.20	96.34	98.69
15	Baseline 205	2,50,000	98.44	97.10	99.09
16	Baseline 211	1,600,000	98.92	98.19	99.08
17	Baseline 214	81,000	98.80	98.17	98.95
18	Baseline 220	740,000	99.40	98.90	99.61
19	Baseline 226	130,000	97.96	96.84	98.04
20	Baseline 227	110,000	98.80	99.27	98.56
21	Baseline 230	22000	99.16	98.90	99.21
22	Baseline 237	200,000	99.40	99.27	96.33
23	Baseline 241	150,000	99.28	99.27	99.48
24	Baseline 268	200,000	97.96	96.70	98.04
25	Baseline 270	200,000	99.04	99.27	98.69
26	Baseline 275	270,000	97.24	96.34	97.17
27	Baseline 281	1,100,000	97.84	97.80	97.25
28	Baseline 286	270,000	97.84	97.45	97.50
29	Baseline 289	490,000	99.16	98.90	99.21
30	Baseline 291	790,000	98.92	97.47	99.34
31	Baseline 312	54,000	98.80	98.90	98.69
32	Baseline 332	18,000	99.76	99.23	99.74
33	Baseline 334	1,000,000	99.04	98.17	99.83
34	Baseline 339	91,000	99.76	98.53	99.74
35	Baseline 342	450,000	99.64	99.63	99.35
36	Baseline 345	43,000	99.40	98.17	99.34
37	Baseline 347	490,000	99.88	99.63	99.87
38	Baseline 356	23,000	97.00	97.07	97.23
39	Baseline 363	98,000	98.92	98.90	98.94
40	Baseline 364	1,300,000	99.40	98.54	99.72

Table 3.3.1 … Continued

S No.	Sample ID	Pol sequence % similarity	Protease sequence % similarity	Reverse transcriptase % similarity
41	serum 97-58-004	96.63	96.19	97.44
42	serum 97-58-356	98.68	97.44	99.04
43	serum 97-59-035	98.92	98.53	98.88
44	serum 97-59-064	98.20	98.53	98.24
45	serum 97-60-108	98.20	98.90	98.08
46	serum 97-60-208	98.56	97.44	98.72
47	serum 97-60-375	99.16	98.17	99.72
48	serum 97-61-073	99.64	100.00	99.52
49	serum 97-61-099	99.52	99.27	99.52
50	serum 97-61-201	99.16	99.63	98.85
51	serum 97-61-315	99.40	99.63	99.36
52	serum 97-61-326	99.52	99.63	99.36
	Count	52.00	52.00	52.00
	Average	**98.72**	**98.37**	**98.70**
	Maximum	99.88	100.00	99.88
	Minimum	96.63	95.60	96.33

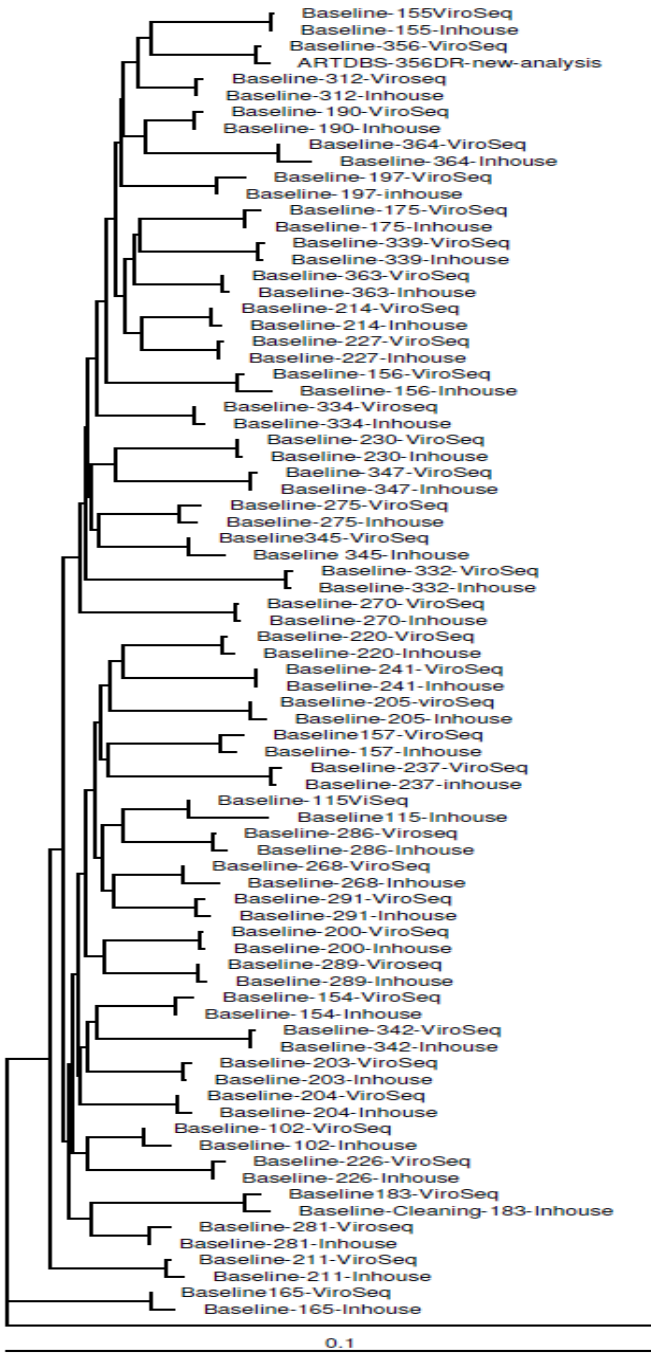


Figure 3.3.3A: Phylogenetic tree of paired plasma samples from chronically infected drug-naïve participants genotyped by ViroSeqTM V2.0 and Home-brewed genotyping systems. CPZ.CM.1998.CAM5 was used as an out-group but cropped out from the figure to spare enough space to enlarge the figure.

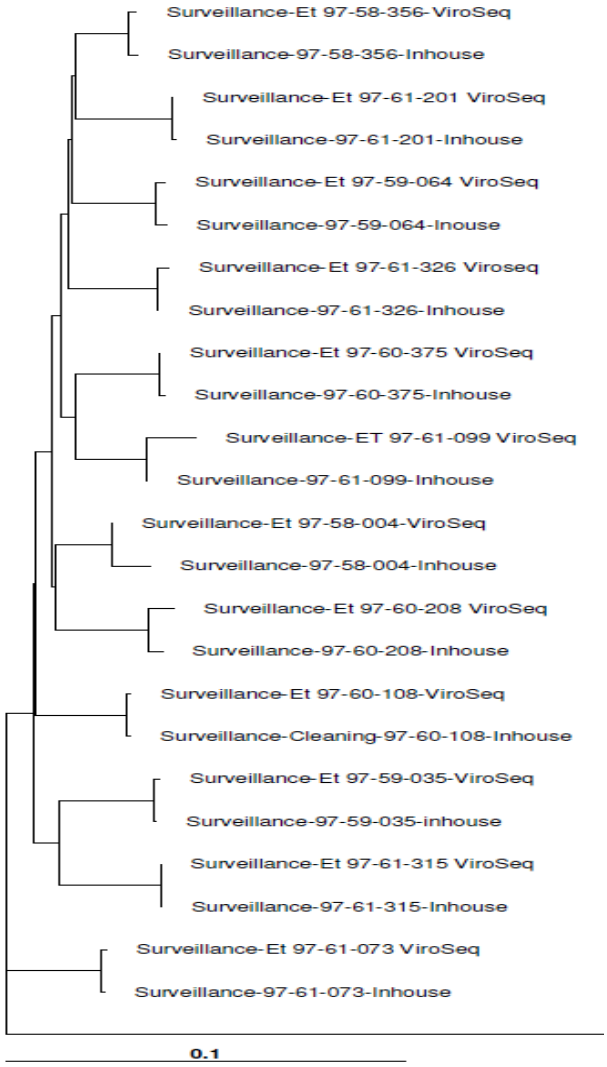


Figure 3.3.3B: Phylogenetic tree of paired sera samples from recently infected drug-naïve participants genotyped by ViroSeqTM V2.0 and Home-brewed genotyping systems. CPZ.CM.1998.CAM5 was used as an out-group but cropped out from the figure to spare enough space to enlarge the figure.

Table 3.3.2 Performance of In-house genotyping system relative to the performance of ViroSeq™ V.2.0

ViroSeq V.2.0 genotyping system		Performance of In-house genotyping system					
		Mutation at PR		Mutations at RT sites		Pooled mutations from both PR and RT regions	
		No. of mutations detected by In-house system	No. of mutations not detected by In-house system	No. of mutations detected by In-house system	No. of mutations not detected by In-house system	No. of mutations detected by In-house system	No. of mutations not detected by In-house system
No. of mutations detected by ViroSeq		470	40	777	36	1247	76
No. of mutations not detected by ViroSeq		37	NA*	32	NA*	69	NA*
Total		507	NA*	809	NA*	1316	NA*
Performance of In-house system relative to ViroSeq system	Relative sensitivity	92.16%		95.57%		94.26%	
	Positive predictive value	99.41%		99.51%		99.47%	

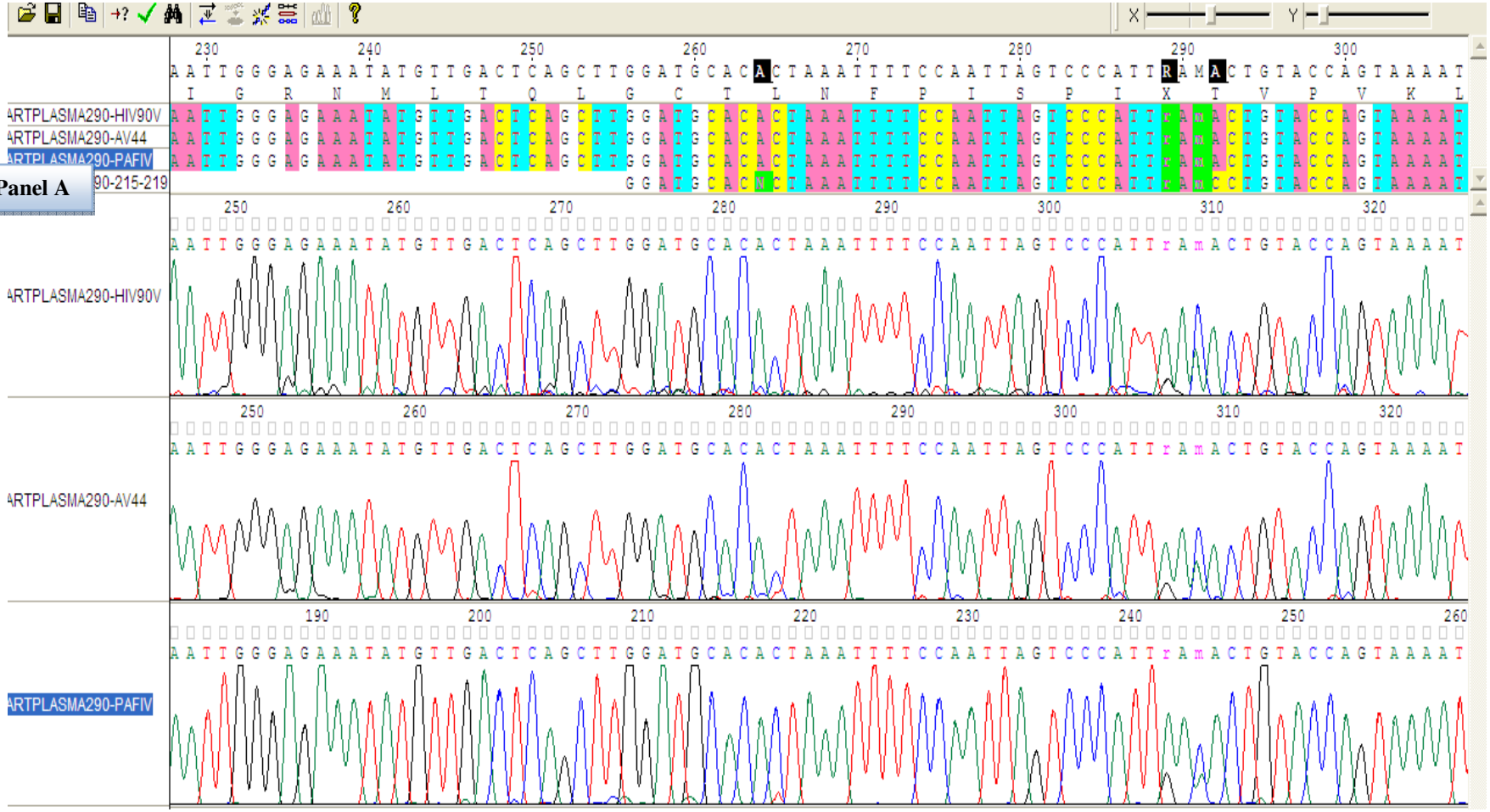
*NA= Not available

3.3.2 Performance of DBS for detection of HIV-1 drug resistance mutations at partial Pol region

A total of 104 paired plasma/DBS samples were collected at baseline from chronically infected drug-naïve participants. However, only 35 DBS samples with their plasma counterparts were available for analysis. The rest were unavailable either because the DBS samples were lost, or plasma counterpart lost or poorly stored. All the DBS samples with available plasma samples were successfully amplified using the In-house genotyping system (Figure 3.3.4). The minimum viral RNA obtained for the analysis was 15,000 copies/ml, while the maximum was 1,600,000 copies/ml.

Figure 3.3.4 RT-PCR Amplification products of HIV-1 vral RNA from DBS samples.
The bands at the left flank are DNA size makers, while the rest are from the samples.

All RT-PCR amplified samples from the DBS were successfully sequenced using the six primers from the In-house genotyping system with good quality sequences. However, primer HIV90V was poor in providing clear chromatogram sequence products, the same way it behaved while sequencing plasma samples. Figure 3.3.5A and B show a screenshot of a section from sequence chromatogram produced during sequencing of one of the DBS samples.


Panel A
ARTPLASMA290-HIV90V
ARTPLASMA290-AV44
ARTPLASMA290-PAFIV

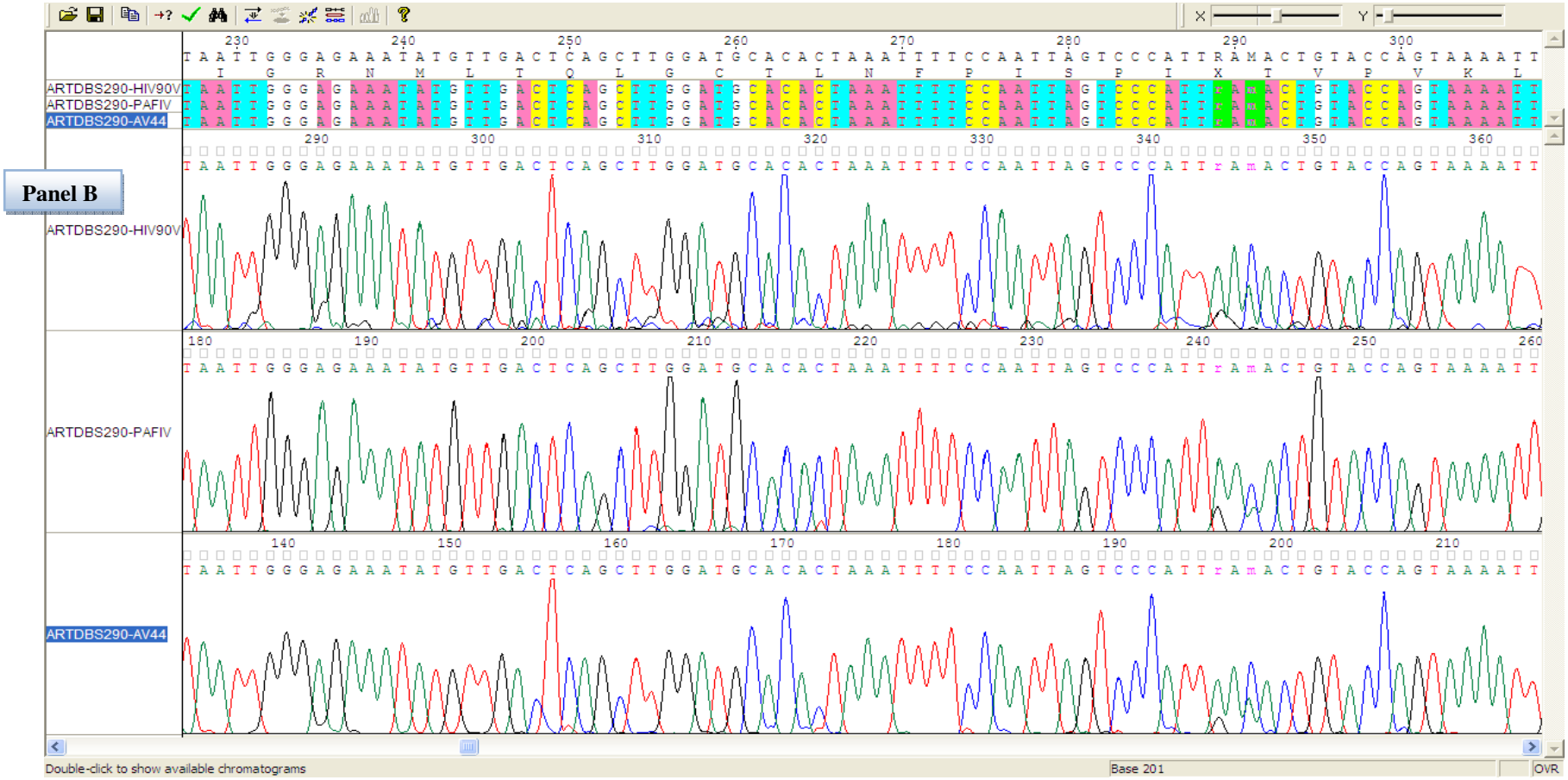


Figure 3.3.5 Screenshots from contig Chromatograms constructed from sequences of plasma sample (Panel A) and DBS sample (Panel B) Originated from same patient (baseline 290) and produced by ChromasPro software for manual editing. Nucleotides 289-292 in the contig detected from both the plasma and DBS samples were manualy edited at different times and found to represent the 6th codon in the RT which showed here presence of 3 mixed virus quasispecies having the wild type K, and mutatnts with E and V.

After editing the sequences manually, FASTA-formatted consensus of the six primer sequences from the DBS sample was paired with its plasma counterpart. Percent nucleotide sequence concordance between sequences of DBS and plasma samples for protease, reverse transcriptase and pooled Pol regions were determined by pairwise alignment of the corresponding plasma and DBS sequences using ClustalW program from BioEdit software version 7.0.9.0. These analyses have shown that the minimum, maximum and mean concordance between the two corresponding samples at the protease region were 98.82%, 100%, and 96.07%, respectively (Table 3.3.3). These values for reverse transcriptase region were 98.84%, 100%, and 96.34%, respectively. Pooling the two regions together, the minimum, maximum, and mean values obtained were 98.27%, 100%, 96.07%, respectively. Phylogenetic analysis of plasma/DBS pairwise aligned corresponding sequences using the BioEdit version 7.0.9.0 has resulted in all the paired sequences branching together (Figure 3.3.6).

Comparison of the performance of DBS sample relative to that of the conventional plasma sample at the amino acid level showed that DBS had relative sensitivities of 97.08%, 96.63%, and 96.78% for the regions of protease, reverse transcriptase and pooled protease and reverse transcriptase, respectively (Table 3.3.4). Relative positive predictive values were 98.83%, 98.04%, and 98.57% for the regions of protease, reverse transcriptase and pooled protease and reverse transcriptase, respectively.

Table 3.3.3 List of plasma/DBS paired nucleotide sequence similarities as determined by BioEdit Version 7.0.9.0

S No.	Sample ID	Viral Load RNA copies/ml	Protease sequence % similarity	Reverse transcriptase % similarity	Pol sequence % similarity
1	Baseline 102	31,000	99.27	99.34	99.47
2	Baseline 211	1,600,000	98.19	99.08	99.04
3	Baseline 236	310,000	98.13	97.64	98.19
4	Baseline 237	200,000	100.00	99.08	99.36
5	Baseline 240	15,000	99.27	98.82	98.94
6	Baseline 241	150,000	98.90	97.51	99.47
7	Baseline 259	480,000	99.63	98.96	99.47
8	Baseline 268	200,000	99.63	99.10	99.79
9	Baseline 269	400,000	100.00	97.30	99.47
10	Baseline 270	200,000	99.27	98.95	99.36
11	Baseline 274	42,000	99.27	97.77	99.47
12	Baseline 275	270,000	96.35	96.49	96.49
13	Baseline 281	1,100,000	96.34	96.07	96.07
14	Baseline 286	270,000	97.45	97.76	98.51
15	Baseline 287	680,000	96.49	97.30	99.26
16	Baseline 289	490,000	100.00	96.92	99.04
17	Baseline 290	120,000	100.00	99.48	100.00
18	Baseline 295	270,000	97.80	98.69	98.51
19	Baseline 311	1,000,000	97.80	97.52	97.87
20	Baseline 312	54,000	99.27	96.92	98.94
21	Baseline 314	89,000	98.17	98.31	99.04
22	Baseline 315	230,000	100.00	99.35	99.89
23	Baseline 319	1,200,000	98.90	99.22	99.47
24	Baseline 320	280,000	99.63	99.22	99.47
25	Baseline 321	140,000	98.16	98.57	99.36
26	Baseline 326	73,000	99.63	98.82	99.15
27	Baseline 332	18,000	98.90	96.67	98.62
28	Baseline 333	880,000	99.63	97.51	99.57
29	Baseline 334	1,000,000	99.63	98.43	99.15
30	Baseline 339	91,000	98.53	99.48	99.68
31	Baseline 342	450,000	99.63	97.94	99.79
32	Baseline 347	490,000	100.00	97.95	98.89
33	Baseline 356	23,000	96.70	97.25	97.24
34	Baseline 360	310,000	100.00	100.00	100.00
35	Baseline 364	1,300,000	98.91	99.86	99.47
		Count	35.00	35.00	35.00
		Average	**98.84**	**98.27**	**98.82**
		Maximum	100.00	100.00	100.00
		Minimum	96.34	96.07	96.07

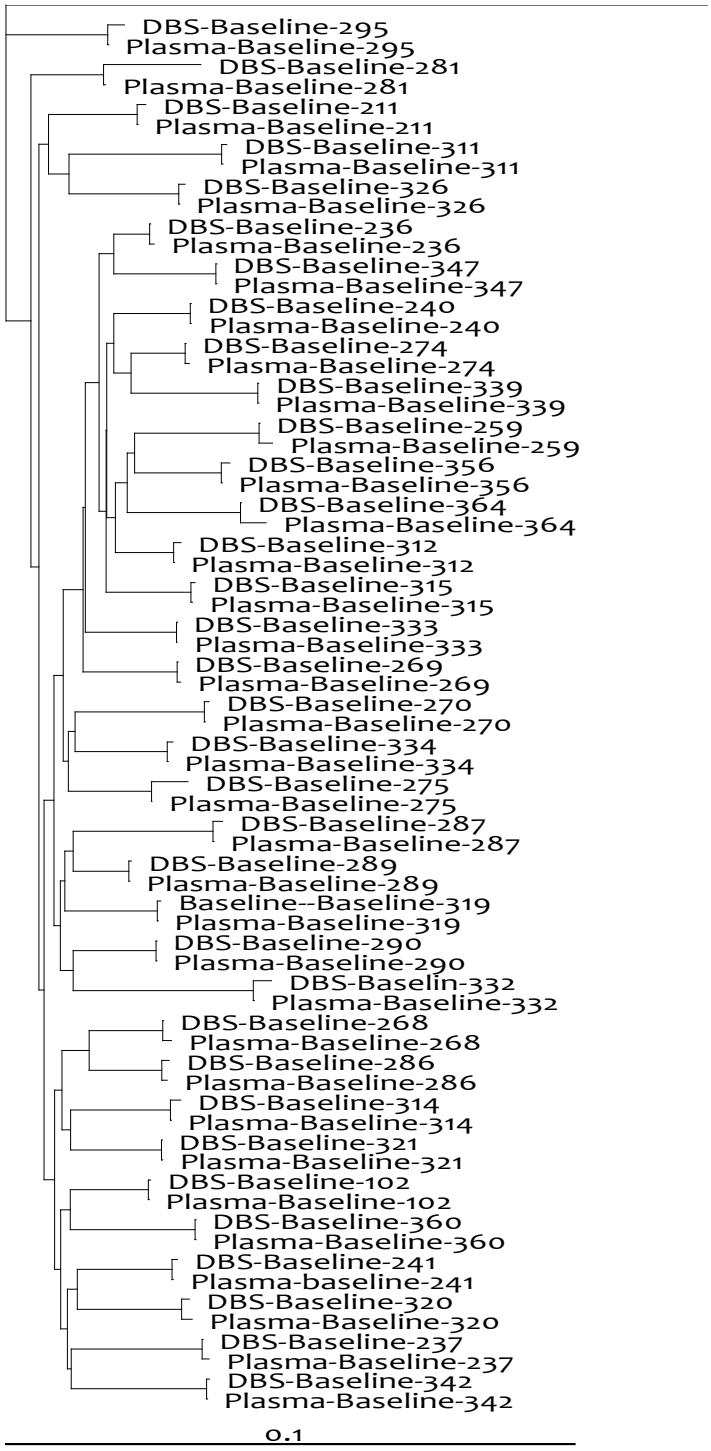


Figure 3.3.6 Phylogenetic tree of paired plasma/DBS samples from chronically infected drug-naïve participants genotyped by home-brewed genotyping system.
CPZ.CM.1998.CAM5 was used as an out-group but cropped out from the figure to spare enough space to enlarge the figure.

Table 3.3.4 Performance of DBS samples in detecting HIV-1 drug resistance mutations relative to the performance of plasma samples at amino acid level

Performance of plasma samples		Performance of DBS samples					
		Mutation at PR		Mutations at RT sites		Pooled mutations from both PR and RT regions	
		No. of mutations detected from DBS samples	No. of mutations not detected from DBS samples	No. of mutations detected from DBS samples	No. of mutations not detected from DBS samples	No. of mutations detected DBS samples	No. of mutations not detected from DBS samples
No. of mutations detected from plasma		332	10	545	19	877	19
No. of mutations not detected from plasma		6	NA*	10	NA*	16	NA*
Total		338	NA*	555	NA*	893	NA*
Performance of DBS relative to plasma	Relative sensitivity	97.08%		96.63%		96.78%	
	Positive predictive values	98.83%		98.04%		98.57	

*NA= Not available

3.4. Prevalence of ARV Drug Resistance Mutations

3.4.1 Drug resistance mutations from drug naïve recently infected participants

Forty-one serum samples were successfully amplified and genotyped from this group of participants. In all of the samples, all of the 99 codons of protease and at least the first 252 codons of the reverse transcriptase enzymes were successfully amplified and sequenced. A total of 152 drug resistance-associated mutations were detected at both the protease (151 from 14 positions) and reverse transcriptase (one NNRTI mutation from a single position) (Table 3.4.1). Of these protease mutations, however, only two major mutations were detected (G73A and N88D each in a single event) belonging to those resistance mutations designated in the 2009 WHO list. Thus, prevalence of transmitted drug resistance in Addis Ababa during 2005, as defined by the WHO's guideline, was 4.88% (100 X 2/41) for protease and 0% for NRTI and NNRTI, which were all below the 5% threshold level.

The rest of the mutations were identified as drug resistance based only on the list in the 2009 update of International AIDS Society-USA (IAS-USA). Except for the WHO-designated protease inhibitor resistance mutations G73A and N88D, all the rest were minor mutations. One mutation V90I of the reverse transcriptase belonged to NNRTI resistance mutation, according to IAS-USA but not according to WHO. The most frequent single mutation was H69K, which occurred in 39/41 (95.12%) of the samples, followed by M36I and I93L occurring in 38/41 (92.68%) and 35/41 (85.37%), respectively. The positions that showed the most diverse substitutions were M36 (with substitutions to I, L, and V) occurring 40/41 (97.56%), and K20 (with substitutions to R and M, occurring 7/41 (17.03%) of the times.

Table 3.4.1 Prevalence of drug resistance-associated mutations among recently infected drug-naïve study participants

Genotyped Region	Mutation Type	Substitution		No./41 samples (% prevalence)	Total No. (%) substitutions per position	% prevalence within the class
		From	To [Algorithm]*			
Protease	**PI resistance mutations**	I13	V [δ]	4 (9.76%)	4 (9.76%)	2.65
		G16	E [δ]	8 (19.51%)	8 (19.51%)	5.30
		K20	R [δ]	6 (14.63)	7 (17.03%)	3.97
			M [δ]	1 (2.44%)		0.66
		L33	V [δ]	1 (2.44%)	1 (2.44%)	0.66
		M36	I [δ]	38 (92.68%)	40 (97.56%)	25.17
			L [δ]	1 (2.44%)		0.66
			V [δ]	1 (2.44%)		0.66
		I62	V [δ]	1 (2.44%)	1 (2.44%)	0.66
		L63	P [δ]	10 (24.39%)	10 (24.39%)	6.62
		I64	V [δ]	1 (2.44%)	1 (2.44%)	0.66
		H69	K [δ]	39 (95.12%	39 (95.12%	25.83
		G73	**A** [ε,δ]	1 (2.44%)	1 (2.44%)	0.66
		T74	S [Iδ]	1 (2.44%)	1 (2.44%)	0.66
		V77	I [δ]	2 (4.88%)	2 (4.88%)	1.32
		N88	**D** [ε,ψ,δ]	1 (2.44%)	1 (2.44%)	0.66
		I93	L [δ]	35 (85.37%)	35 (85.37%)	23.18
	Total	**14 positions**	**17 substitution types**	**151 substitutions**	**151 (100%)**	**~ 100.00%**
Reverse transcriptase	NNRTI resistance mutations	V90	I [δ]	1 (2.44%)	1 (2.44%)	100.00
Grand total		15 positions	18 substitution types	152 at 15 positions	152 (100%)	100

* Algorithms considered: ε = WHO; δ = IAS-USA; ψ = Stanford HIV database

3.4.2 Drug Resistance Mutations from Drug-Naïve Chronically Infected Participants

The sources of sample for sequencing and genotypic analyses for this part of the work were obtained from the conventional plasma and the previously tested DBS samples, where amplification or proper sequencing was impossible from the plasma samples. The genotyping system used were also mixed, in which genotyping from the ViroSeq^{TM} V2.0 was given priority when satisfactorily done but otherwise the already validated In-house genotyping system was used to produce the sequence result. Thus, most of those sequences produced by the In-house system cover codons 8 to 99 of the protease and 1 to 252 of the reverse transcriptase. Nonetheless, these ranges are enough to make drug resistance analyses for resistance mutations to protease inhibitors, NRTI inhibitors and NNRTI inhibitors.

3.4.2.1 Resistance mutations for protease inhibitors in chronically infected drug-naïve participants

A total of 32 PI resistance-associated mutations occurring at the frequency of 293 were detected from 25 positions when all the three algorithms were considered independently (Table 3.4.2). Twenty-four of these were identified by IAS-USA algorithm alone, whereas the rest 8 were identified concordantly at least by two of the three algorithms utilized. The most prominent mutations were M36I, H69K, L63P, which occurred respectively 81/85(95.29%), 77/85(90.59%), and 29/85(34.11%) of the times. The contributions of these mutations from the total 293 were 27.65%, 26.28%, and 9.9%, respectively. Of the 25 positions harboring these resistance mutations, five harbored two or more mutations: K20 substituted with R (13/85 [15.29%]) or I (1/85 [1.38%]); M36 substituted with I (81/85 [95.29%]), L (1/85 [1.38%]), or V (1/85 [1.38%]); I64 substituted with L (1/85 [1.38%]), M (1/85 [1.38%]), or V (1/85 [1.38%]); and V82 substituted with A (4/85 [4.71%]), or I (8/85 [9.41%]). Since there is no complete agreement

between all available algorithms in defining resistance mutations, the mutations identified by at least two of the three algorithms as conferring resistance were accepted in this study as drug resistance mutations, and hence, were analyzed separately (Figure 3.4.1). There were 8 mutations at 9 different positions that were identified by two or more of the algorithms used. These mutations occurred 18/85 (21.18%) of the time. Among the major PI resistance mutations, only G48V, G73A, and I84C each occurred once (1/85 [1.38%] each); each of the rest five mutations was observed in three of the 85 samples (3/85 [~4%] each) (Figure 3.4.1).

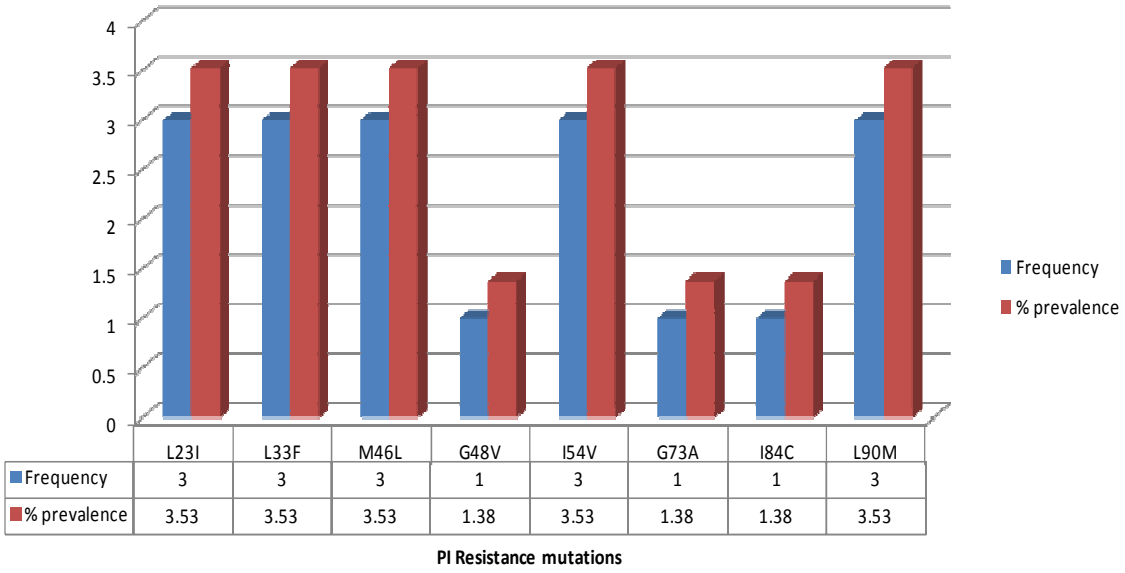


Figure 3.4.1 Frequency of detected protease inhibitors resistance mutations among chronically infected drug-naïve participants in the protease region as identified by at least two of the three genotyping algorithms
(WHO's 2009 updated list, IAS-USA's 2009 updated list, and Stanford's 2009 updated list).

Table 3.4.2 Prevalence of PI drugs resistance-associated mutations among chronically infected drug-naïve study participants

Genotyped Region and Mutation Type	Substitution		Frequency N (%)	N(%) substitutions per position	% prevalence within the class
	From	To [Algorithm]*			
	C10	I [δ]	3 (3.53)	3 (3.53)	1.02
	V11	I [δ]	2 (2.35)	2 (2.35)	0.68
	I13	V [δ]	18 (21.18)	18 (21.18)	6.14
	G16	E [δ]	8 (9.41)	8 (9.41)	2.73
	K20	R [δ]	13 (15.29)	4 (4.71)	4.44
		I [δ]	1 (1.38)		0.34
	L23	I [ε, ψ]	3 (3.53)	3 (3.53)	1.02
	L33	F [δ,ψ]	3(3.53)	3 (3.53)	1.02
	M36	I [δ]	81 (95.29)	84 (98.82)	27.65
		L [δ]	1 (1.38)		0.34
		V [δ]	1 (1.38)		0.34
	M46	L [ε, δ, ψ]	3 (3.53)	3 (3.53)	1.02
	G48	V [ε, δ, ψ]	1 (1.38)	1 (1.38)	0.34
	I54	V [ε, δ, ψ]	3 (3.53)	3 (3.53)	1.02
	Q58	E [δ]	1 (1.38)	1 (1.38)	0.34
	D60	E [δ]	4 (4.71)	4 (4.71)	1.37
	I62	V [δ]	5 (5.88)	5 (5.88)	1.71
	L63	P [δ]	29 (34.11)	29 (34.11)	9.9
	I64	L [δ]	1 (1.38)	3 (3.53)	0.34
		M [δ]	1 (1.38)		0.34
		V [δ]	1 (1.38)		0.34
	H69	K [δ]	77 (90.59)	77 (90.59)	26.28
	A71	I [δ]	2 (2.35)	5 (5.88)	0.68
		T [δ]	3 (3.53)		1.02
	G73	A [ε, δ]	1 (1.38)	1 (1.38)	0.34
	T74	S [δ]	5 (5.88)	5 (5.88)	1.71
	V77	I [δ]	3 (3.53)	3 (3.53)	1.02
	V82	A [δ]	4 (4.71)	12 (14.11)	1.37
		I [δ]	8 (9.41)		2.73
	I84	C [ε, ψ]	1 (1.38)	1 (1.38)	0.34
	L90	M [ε, δ, ψ]	3(3.53)	3 (3.53)	1.02
	I93	L [δ]	3 (3.53)	3 (3.53)	1.02
Total	25 positions	32 substitution types	293 total substitutions		~ 100%

PI resistance

* Algorithms considered: ε = WHO; δ = IAS-USA; ψ = Stanford HIV database

3.4.2.1 Resistance mutations for NRTIs and NNRTIs in chronically infected drug-naïve participants

Although this group of participants was drug-naïve, a total of 24 and 15 mutations distributed in eight and seven positions of the reverse transcriptase region were identified against NRTI and NNRTI, respectively (Table 3.4.3). However, not all of these mutations were identified unanimously by all the three algorithms used as resistance. Of the nine substitutions identified for NRTI resistance mutations, only two were identified singly by WHO (V75A) and IAS-USA (A62V) algorithms; the rest were identified either by two or three of them. Similarly, six of the eight detected NNTI resistance mutations were identified by two or all of three the algorisms (Figure 3.4.2), the only exceptions being E138A, and V179T, which were identified by IAS-USA algorithm alone. Even under this strict scrutiny, the frequency of resistance mutations among these drug-naïve treatment eligible patients were 19 for the 8 NRTI and 13 for 6 NNRTI mutations. The most dominant NRTI resistance mutation was D67N, which occurred in 5.88% of the study participants. Similarly, the most dominant NNRT resistance mutation was K103N, occurring in 4.71% of the study participants.

Table 3.4.3 Prevalence of NRTI and NNRTI drug resistance-associated mutations among chronically infected drug-naïve study participants

Genotyped Region and Mutation Type	Substitution		Frequency N	N substitutions per position	% prevalence within the class
	From	To [Algorithm]*			
NRTI Resistance mutations	M41	L [ε,δ,ψ]	3	3.53	12.50
	A62	V [δ]	3	3.53	12.50
	D67	N [ε,δ,ψ]	5	5.88	20.53
	L74	V [ε, δ]	3	3.53	12.50
	V75	A [ε]	2	2.35	8.33
	M184	V [ε,δ, ψ]	1	2	4.17
		I [ε, δ, ψ]	1		4.17
	L210	W [ε,δ,ψ]	3	3.53	12.50
	T215	Y [ε,δ,ψ]	3	3.53	12.50
Total	8 positions	9 substitution types	**24**		~ 100
NNRTI Resistance mutations	A98	G [δ,ψ]	1	1.38	6.67
	L100	I [ε,δ,ψ]	3	3.53	20.00
	K103	N [ε,δ,ψ]	4	4.71	26.67
	E138	A [δ]	1	1.38	6.67
	V179	T [δ]	1	1.38	6.67
	Y181	C [ε,δ,ψ]	3	3.53	20.00
	Y188	C [ε,δ,ψ]	1	2	6.67
		H [ε,δ,ψ]	1		6.67
Total	7 positions	8 substitution types	**15**		~ 100

* Algorithms considered: ε = WHO; δ = IAS-USA; ψ = Stanford HIV database

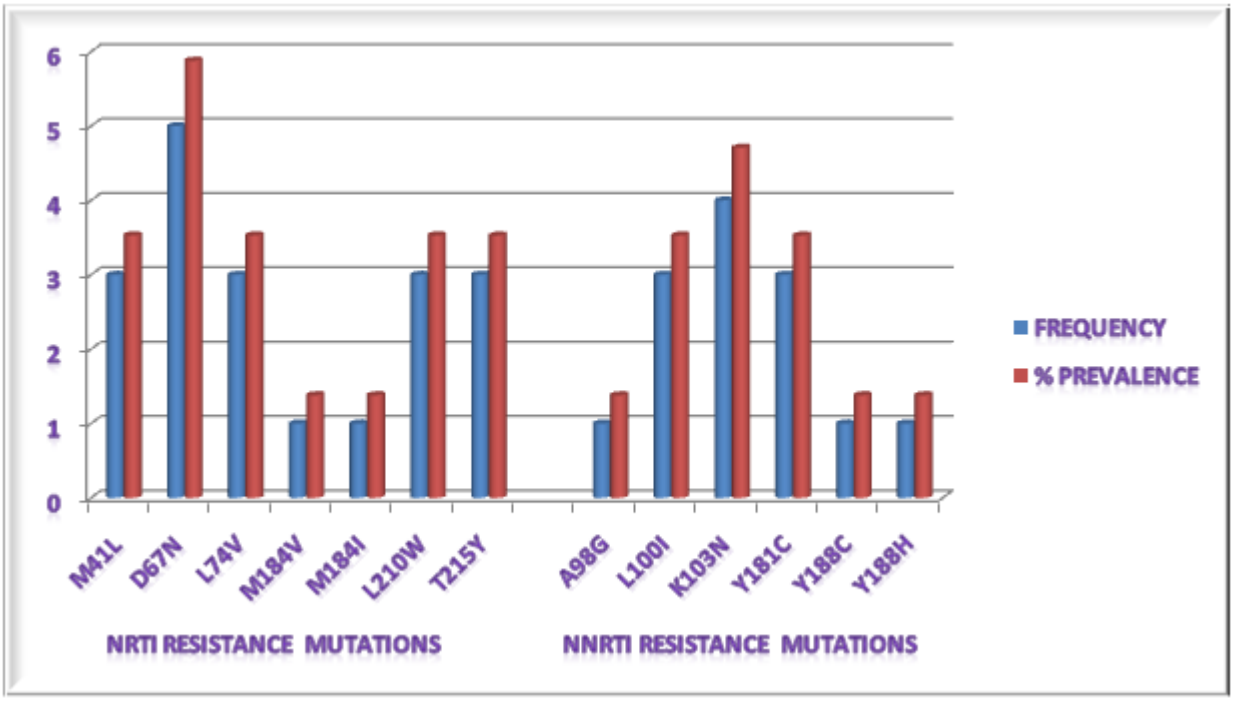


Figure 3.4.2 Frequency of detected major NRTIs and NNRTIs resistance mutations among chronically infected drug-naïve participants in the reverse transcriptase region as identified by at least two of the three genotyping algorithms
(WHO's 2009 updated list, IAS-USA's 2009 updated list, and Stanford's 2009 updated list).

3.4.3 Resistance mutations in drug-experienced chronically infected participants

3.4.3.1 Resistance mutations for protease inhibitors in chronically infected drug-experienced participants

A total of 20 protease resistance mutations from 19 positions could be identified from the 66 heavily treated participants at a total frequency of 266, as determined by the three genotyping algorithms separately. Of these, however, only four mutations from three positions and with frequency of one from each [I54A, V82A, N88D, and N88S] were accepted as relevant PI drug resistance mutations unanimously by all the three algorithms. These four mutations occurred within two individuals; so there were double PI resistance mutations in these individuals. The rest 16 substitutions, occurring from 63 of the 66 (95.45%) participants, were recognized as drug resistance-associated mutations in the IAS-USA list only (Table 3.4.4). While three (4.5%) of the

participants showed no PI resistance-associated mutation at all, each of the others have harbored from two to seven mutations. The three most prevalent of the latter mutations were H69K (63/66 [95.45%]), M36I ([93.93%]), and L89V ([87.87%]).

Table 3.4.4 Prevalence of PI drugs resistance-associated mutations among chronically infected drug-experienced study participants as identified by the three genotyping algorithms considered separately.

Substitution		Frequency N (%)	N(%) substitutions per position
From	To [Algorithm]*		
I13	V [δ]	6 (9.09%)	D (9.09%)
G16	E [δ]	6 (9.09%)	6 (9.09%)
K20	R [δ]	11 (16.67)	12 (18.18%)
	I [δ]	1 (1.52%)	
M36	I [δ]	62 (93.94%)	63 (95.45%)
	V [δ]	1 (1.52%)	
I54	A [ε,δ,ψ]	1 (1.52%)	1 (1.51%)
D60	E [δ]	4 (6.06%)	4 (6.06%)
I62	V [δ]	1 (1.52%)	1 (1.51%)
L63	P [δ]	15 (22.72%)	15 (22.72%)
I64	L [δ]	3 (4.55%)	4 (6.06%)
	V [δ]	1(1.52%)	
H69	K [δ]	63 (95.45%)	63 (95.45%)
A71	V [δ]	1(1.52%)	1 (1.51%)
T74	S [δ]	10 (15.15%)	10 (15.15%)
V77	I [δ]	1(1.52%)	1 (1.51%)
V82	A [ε,δ,ψ]	1(1.52%)	1 1.51%)
N88	D [ε,δ,ψ]	1 (1.52%)	2 (3.03)
	S [ε,δ,ψ]	1 (1.52%)	
L89	V [δ]	58 (87.88)	58 (87.88%)
L90	M [δ]	1 (1.52%)	1 (1.52%)
I93	M [δ]	6 (9.09%)	17 (25.76%)
	L [δ]	11(16.67)	
19 positions	**20 substitutions**	**Total occurrence 266**	

* Algorithms considered: ε = WHO; δ = IAS-USA; ψ = Stanford HIV database

3.4.3.2 Drug resistance mutations for NRTIs and NNRTIs in chronically infected drug-experienced participants

A very diverse and numerous resistance mutations occurred in the NRTI and NNRTI classes of ART drugs currently in use in Ethiopia. Forty-three of the 66 (65.15%) participants had at least one NRTI mutations. Twenty-four different NRTI mutations occurred with the frequency of 133 at 14 positions (Table 3.3.3.2). Eight of these 24 mutation, having frequency of 22, were recognized by single algorithm, whereas the rest 16 with aggregate frequency of 111 were recognized by two or more of the three algorithms (Table 3.4.5 and Figure 3.4.3. panel A). The eight most frequent NRTI mutations identified were M184V (39/66 [59.09%]), T215Y (11/66 [16.67%]), K70R (10/66 [15.15%]), M41L (9/66 [13.63%]), K219Q and D67N (each 8/66 [12.12%]), T215F (7/66 [10.6%]), and L210W (6/66 [9.09%), in that order. Four positions had two or more resistance mutations: V75, M184, T215, and K219 with combined frequencies of 4, 40, 18, and 9, respectively. In addition, multiple NRTI resistance mutations occurred with mean four and aggregate frequency of 28, ranging from double mutations in seven persons to 13 mutations in 1 person (data not shown).

Regarding prevalence of NNRI resistance mutations, 16 RT positions harbored 23 mutations with frequency of 109 (Table 3.4.6). These mutations were carried by 48 of the 66 (72.72%) participants. Multiple NNRTI resistance mutations occurred in 33 of the 66 (50%) persons or 33/48 (68.75%) of persons harboring NNRTI resistance mutations. Five of the 23 mutations were recognized singly by IAS-USA (V90I and E138A) or Stanford HIV drug resistance database (F227L, F227C, and K238T). The rest 18 mutations were recognized by two or more of the genotyping algorithms, and occurred with frequency of 92. K103N (23/66 [34.84%]), Y181C

(11/66 [16.67%]), A98G (10/66 [15.15%]), and G190A (7/66 [10.6%]) were the most prevalent NNRTI resistance mutations detected (Figure 3.4.3. panel B).

Overall, 40 of the 66 (60.60%) participants harbored double class NRTI and NNRTI resistance mutations. Moreover, two of the four PI resistance mutations occurred in two participants who already harbored NRTI and NNRTI resistance mutations, making them resistant to all of the three classes of drugs provided in the country. One of the participants who developed triple class resistance had already switched to 2nd line regimen (3TC,TDF, kaletra) from the initial 1st line ([D4T, 3TC, EFV]), whereas the second person stayed still on the first line regimen ([D4T, 3TC, NVP]), which lacked any protease inhibitor.

Table 3.4.5 Prevalence of NRTI drug resistance-associated mutations among chronically infected drug-experienced study participants

Substitution		Frequency N (%)	N(%) substitutions per position
From	**To [Algorithm]***		
M41	L [ε,δ,ψ]	9 (13.64%)	13.64%
A62	V [δ]	2 (3.03%)	3.03%
K65	R [ε,δ,ψ]	3 (4.55%)	4.55%
D67	E [ε]	1 (1.52%)	1.52%
D67	G [ε]	2 (3.03%)	15.15%
	N [ε,ψ]	8 (12.12%)	
T69	D [ε]	3 (4.55%)	4.55%
K70	G [ψ]	9 (13.645)	28.79%
	R [ε,δ, ψ]	10 (15.15%)	
L74	V [ε, δ]	1 (1.52%)	1.52%
V75	I [δ, ψ]	2(3.03%)	6.06%
	M [ε,ψ]	1 (1.52%)	
	T [ε,ψ]	1 (1.52%)	
F116	Y [ε,δ, ψ]	3 (4.55%)	4.55%
M184	V [ε,δ, ψ]	39 (59.09%)	60.61%
	I [ε, δ, ψ]	1 (1.52%)	
L210	W [ε,δ,ψ]	6 (9.09%)	9.09%
T215	I [ε]	1 (1.52%)	33.33%
	S [ε]	3 (4.55%)	
	Y [ε,δ,ψ]	11 (16.67)	
	F [ε,δ,ψ]	7 (10.61%)	
K219	E [ε,δ,ψ]	1 (1.52%)	15.15%
	Q [ε,δ,ψ]	8 (12.12%)	
	R [ε]	1 (1.52%)	

* Algorithms considered: ε = WHO; δ = IAS-USA; ψ = Stanford HIV databas

Table 3.4.6 Prevalence of NNRTI drug resistance-associated mutations among chronically infected drug-experienced study participants

Substitution	Frequency N (%)	N(%) substitutions per position	% prevalence within the class
V90	I [δ]	4	6.06
A98	G [δ,ψ]	10	15.15
L100	I [ε,δ,ψ]	2	3.03
K101	E [ε,δ,ψ]	6	9.09
	P [ε,δ,ψ]	1	1.51
K103	N [ε,δ,ψ]	23	34.84
	S [ε,ψ]	1	1.51
V106	A [ε,δ,ψ]	3	4.54
	M [ε,δ,ψ]	7	10.60
V108	I [δ,ψ]	6	9.09
E138	A [δ]	3	4.54
V179	D [δ,ψ]	3	4.54
	E [δ,ψ]	3	4.54
Y181	I [ε,δ,ψ]	1	1.51
	C [ε,δ,ψ]	11	16.67
Y188	L [ε,δ,ψ]	3	4.54
G190	A [ε,δ,ψ]	7	10.60
	S [ε,δ,ψ]	1	1.51
P225	H [ε,δ,ψ]	6	9.09
F227	L [ψ]	5	7.57
	C [ψ]	1	1.51
M230	L [ε,δ,ψ]	1	1.51
K238	T [ψ]	1	1.51

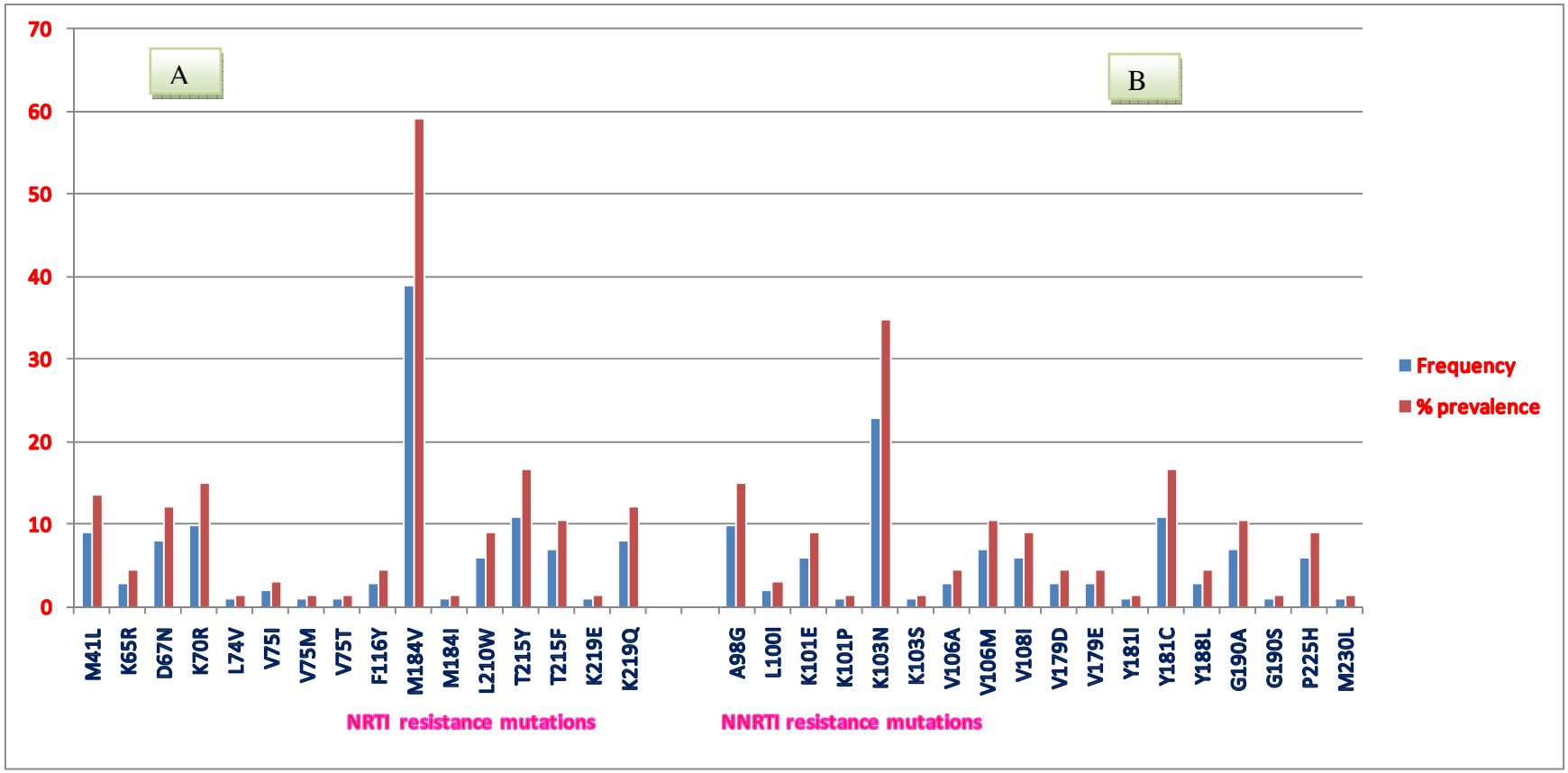


Figure 3.4.3 Frequency of detected NRTIs (panel A) and NNRTIs (Panel B)resistance mutations among chronically infeted dug-experienced participants in the reverse transcripatse region as identified by at least two of the three genotyping algorithms
(WHO's 2009 updated list, IAS-USA's 2009 updated list, and Stanford's 2009 updated list)

Patterns of Co-variation among major drug resistance mutations within and between drug classes

Many of the major resistance mutations did not occur singly, occurring simultaneously (co-varying) with another resistance mutation within the same class or outside its class. Table 3.4.7 A-E summarizes statistically significant co-varying pairs of resistance mutations observed in this study. Two major PI resistance mutations (I54A and N88S) co-varied with two other major PI resistance mutations (V82A) and L90M respectively (Table 3.4.7A). These four PI resistance mutations have also co-emerged with 10 major NRTI or NNRTI resistance mutations (Table 3.4.7D). Similarly, 15 major NNRTI resistance mutations have co-emerged within the same class in 14 combinations. A98G, K103N, K103S, and V179D have co-emerged with three and above NNRTI resistance mutations (Table 3.4.7B). The pairs of co-variations between V179D/Y181I and V179D/M230L were the most frequent of all co-variation pairs observed within the NNRTI resistance mutations, each pair co-occurring 32% of their appearances.

The largest number of co-variation within same class took place among NRTI resistance mutations (Table 3.4.7C). Twelve NRTI resistance mutations have covaried in 29 combinations. Six of these mutations covaried with five other NRTI mutations; and 2, 3, and 1 mutations covaried respectively with 4, 3 and 2 other NRTI resistance mutations. Three positions (M184I/V, T215F/Y, and K219E/Q) each showed covariation with nine other NRTI resistance mutations. While the two alternative mutations in the positions T215 and M184 correlated some times with the same mutations (e.g. T215 with M41L; T215 with V184V; and M184 with K70R; M184 with K219Q), the two alternative mutations at position K219 almost always correlated with the same mutations (e.g. with D67N; K70R; and T215F), except in one instance in which K219Q not K219E correlated with M184I and M184V. Eight pairs of covariations had

frequencies of covariation over 20%, the most frequent of all being members of thymidine analog resistance mutations M41L/L210W (41%), K70R/K219Q (38%), D67N/219Q (33%), and L210/T215Y (32%).

In addition, 13 NRTI resistance mutations were found to co-vary with 11 NNRTI resistance mutations in 30 combinations. Four of the 13 NRTI resistance mutations were able to co-vary with three or more NNRTI resistance mutations: M41L (with A98G, K103N, and Y181I); D67N (with V179D, Y181I, G190S, and M230L); K70R (with V106M, G190S, and M230L); M184V (with A98G, K101E, Y181C, and G190A); and T215Y (with A98G, L101P, V108I, and Y181I). The frequency of co-variation ranges from 7% (in 3 of the 30 pairs) to 100% (in the pair F116Y/V106M.

Table 3.4.7 Patterns of co-variation of major drug resistance mutations within and between drug classes.

A.Within PI

PI-Mutation	*Co-mutation*	*Pearson Correlation*	*% covariance*	*Significance p value (2-tailed)*	*Significance P value at <=*	*Resistance Class of co-Mutaion*
54A	V82A	1.000	100%	0.000	0.01	PI
N88S	L90M	1.000	100%	0.000	0.01	PI

B. Within NNRTI resistance mutations

NNRTI-Mutation	*Co-mutation*	*Pearson Correlation*	*% covariance*	*Significance p value (2-tailed)*	*Significance P value at <=*	*Resistance Class of Co-Mutaion*
A98G	K103N	0.312	9.73%	0.011	0.05	NNRTI
	V108I	0.307	9.42%	0.012	0.05	NNRTI
	G190S	0.294	8.64%	0.017	0.05	NNRTI
K100I	Y188L	0.386	14.90%	0.001	0.01	NNRTI
K101E	G190S	0.392	15.37%	0.001	0.01	NNRTI
V101P	G190A	0.360	12.96%	0.003	0.01	NNRTI
K103N	V108I	0.322	10.37%	0.008	0.01	NNRTI
	P225H	0.322	10.37%	0.008	0.01	NNRTI
K103S	V106M	0.360	12.96%	0.003	0.01	NNRTI
	M230L	0.360	12.96%	0.003	0.01	NNRTI
V179D	Y181I	0.568	32.26%	0.000	0.01	NNRTI
	G190A	0.397	15.76%	0.001	0.01	NNRTI
	M230L	0.568	32.26%	0.000	0.01	NNRTI
Y181I	G190A	0.360	12.96%	0.003	0.01	NNRTI

C. Within NRTI resistance mutations

NRTI-Mutation	Co-mutation	Pearson Correlation	% covariance	Significance p value (2-tailed)	Significance P value at <=	Resistance Class of co-Mutaion
M41L	D67N	0.259	6.71%	0.003	0.01	NRTI
	L74V	0.312	9.73%	0.011	0.05	NRTI
	V75I	0.445	19.80%	0.000	0.01	NRTI
	M184V	0.331	10.96%	0.007	0.01	NRTI
	L210W	0.642	41.22%	0.000	0.01	NRTI
	T215Y	0.533	28.41%	0.000	0.01	NRTI
	T215F	0.293	8.58%	0.000	0.01	NRTI
D67N	K70R	0.490	24.01%	0.000	0.01	NRTI
	M184V	0.309	9.55%	0.012	0.05	NRTI
	L210W	0.367	13.47%	0.002	0.01	NRTI
	T215F	0.475	22.56%	0.000	0.01	NRTI
	K219E	0.334	11.16%	0.006	0.01	NRTI
	K219Q	0.573	32.83%	0.006	0.01	NRTI
K70R	M184I	0.294	8.64%	0.017	0.05	NRTI
	M184V	0.266	7.08%	0.031	0.05	NRTI
	T215F	0.403	16.24%	0.001	0.01	RNTI
	K219E	0.294	8.64%	0.017	0.01	NRTI
	K219Q	0.620	38.44%	0.000	0.01	NRTI
L74V	T215Y	0.277	7.67%	0.024	0.05	NRTI
V75I	L210W	0.252	6.35%	0.042	0.05	NRTI
	T215F	0.513	26.32%	0.000	0.01	NRTI
M184I	**K219Q**	**0.334**	11.16%	0.006	0.01	NRTI
M184V	L210W	0.263	6.92%	0.033	0.05	NRTI
	T215Y	0.372	13.84%	0.002	0.01	NRTI
	T215F	0.287	8.24%	0.020	0.05	NRTI
	K219Q	0.309	9.55%	0.012	0.05	NRTI
L210W	T215Y	0.566	32.04%	0.000	0.01	NRTI
T215F	K219E	0.360	12.96%	0.003	0.01	NRTI
	K219Q	0.324	10.50%	0.008	0.01	NRTI

D. Between PI resistance and NRTI/NNRTI resistance mutations

PI-Mutation	Co-mutation	Spearman's correlation coefficient	% covariance	Significance p value (2-tailed)	Significance P value at <=	Resistance Class of co-Mutaion
I54A	M41L	0.312	9.73%	0.110	0.05	NRTI
	D67N	0.334	11.16%	0.006	0.01	NRTI
	L210W	0.392	15.37%	0.001	0.01	NRTI
	T215Y	0.277	7.67%	0.024	0.05	NRTI
	V179D	0.568	32.26%	0.000	0.01	NNRTI
	Y181I	1.000	100%	0.000	0.01	NNRTI
	G190A	0.360	12.96%	0.003	0.01	NNRTI
V82A	M41L	0.312	9.73%	0.110	0.05	NRTI
	D67N	0.334	11.16%	0.006	0.01	NRTI
	L210W	0.392	15.37%	0.001	0.01	NRTI
	T215Y	0.277	7.67%	0.024	0.05	NRTI
	V179D	0.568	32.26%	0.000	0.01	NNRTI
	Y181I	1.000	100%	0.000	0.01	NNRTI
	G190A	0.360	12.96%	0.003	0.01	NNRTI
N88S	T215F	0.360	12.96%	0.003	0.01	NRTI
	Y181C	0.277	7.67%	0.024	0.05	NNRTI
	G190A	0.360	12.96%	0.003	0.01	NNRTI
L90M	T215F	0.360	12.96%	0.003	0.01	NRTI
	Y181C	0.277	7.67%	0.024	0.05	NNRTI
	G190A	0.360	12.96%	0.003	0.01	NNRTI

E. Between NRTI and NNRTI resistance mutations

NRTI-Mutation	Co-mutation	Pearson Correlation	% covariance	Significance p value (2-tailed)	Significance P value at <=	Resistance Class of co-Mutaion
M41L	A98G	0.325	10.56%	0.008	0.01	NNRTI
	K103N	0.265	7.02%	0.031	0.05	NNRTI
	Y181I	0.312	9.73%	0.011	0.05	NNRTI
K65R	Y181C	0.293	8.58%	0.017	0.05	NNRTI
D67N	V179D	0.365	13.32%	0.003	0.01	NNRTI
	Y181I	0.334	11.16%	0.006	0.01	NNRTI
	G190S	0.334	11.16%	0.006	0.01	NNRTI
	M230L	0.334	11.16%	0.006	0.01	NNRTI
K70R	V106M	0.266	7.08%	0.031	0.05	NNRTI
	G190S	0.294	8.64%	0.017	0.05	NNRTI
	M230L	0.294	8.64%	0.017	0.05	NNRTI
L74V	A98G	0.294	8.64%	0.017	0.05	NNRTI
	V108I	0.392	15.37%	0.001	0.01	NNRTI
V75I	V108I	0.252	6.35%	0.042	0.05	NNRTI
F116Y	V106A	1.000	100%	0.000	0.01	NNRTI
M184V	A98G	0.352	12.39%	0.004	0.01	NNRTI
	K101E	0.263	6.92%	0.033	0.05	NNRTI
	K103N	0.350	12.25%	0.004	0.01	NNRTI
	Y181C	0.277	7.67%	0.024	0.05	NNRTI
	G190A	0.287	8.27%	0.020	0.05	NNRTI
L210W	V101P	0.392	15.37%	0.001	0.01	NNRTI
	Y181I	0.392	15.37%	0.001	0.01	NNRTI
T215Y	A98G	0.378	14.29%	0.002	0.01	NNRTI
	L101P	0.277	7.67%	0.024	0.05	NNRTI
	V108I	0.283	8.01%	0.003	0.01	NNRTI
	Y181I	0.277	7.67%	0.024	0.05	NNRTI
T215F	G190S	0.360	12.96%	0.003	0.01	NNRTI
K219E	V101P	0.392	15.37%	0.001	0.01	NNRTI
K219Q	V101P	0.324	10.50%	0.008	0.01	NNRTI
	G190S	0.334	11.16%	0.006	0.01	NNRTI

3.5 Genetic Diversity of HIV Isolates from the Three Groups of Participants

3.5.1 Subtype Diversity

3.5.1.1 Subtype diversity of HIV-1 isolates from recently infected drug-naïve participants

Five genotyping systems were employed in this genetic diversity study. Four of the five systems used (namely, Stanford, REGA, jpHHM, and Phylogenic analysis) were employed for all sequences. The fifth one (Simplot) was used for those sequences which gave ambiguous results or indication of recombination from results of the other genotyping programs, particularly from jpHHM program.

Regarding sub-typing of sequences from the recently infected individuals, the four programs unanimously assigned subtype C to 36 of the 41 sequences; this resulted in 87.80% concordance between them. Three separate analyses were done to determine the phylogenic relationships of the test sequences with pure clade and CRF reference sequences: at PR and RT regions separately, and the two combined as Pol. In all the three analyses, the 36 sequences were unanimously identified as subtype C (Table 3.5.1 and Figures 3.5.5-3.5.7).

Because the remaining five sequences had ambiguities and produced discrepant results (Table 3.5.1), SimPlot program was run to clarify the ambiguity. One of the sequences ("Et_97-58-098" shown at sequence No. 8 in Table 3.5.1) was sub-typed as CRF02_AG by Stanford program and phylogenetic analysis at the Pol region (Table 3.5.1 and Figure 3.5.5). While jpHHM program identified this sample to have a combined sequence from subtype G and B at the PR and RT regions in the 5' to 3' direction, respectively (Figure 3.5.1A), REGA could not assign it to any

particular subtype; rather, the genetic map of the sequence produced by this program has indicated the presence of ambiguous combinations of 5'-GKFBA-3' ends (Figure 3.5.1C). Analysis done on this sequence using SimPlot program, on the other hand, identified it to have contained combined sequences from three subtypes 5'-GAB-3' as indicated in Figure 3.4.1.1B. Surprisingly, although phylogenetic analysis at the Pol region assigned this sequence to belong to CRF01_AG, separate phylogenetic analyses for PR and RT regions provided different results from that of Pol: while the PR sequence clustered with references belonging to subtype G (in agreement with the other results), the RT sequence clustered with subtype B (Figures 3.5.6 and 3.5.7).

Sequence #3: >Et_97-58-098_Viroseq-C

This sequence is related to subtype(s): **B G**

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	-	176 - 408	**357**	G
358	958 - 1224	-	**1302**	B
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2253	-	2428 - 2660	**2609**	G
2610	3210 - 3476	-	**3554**	B

Genome map (based on HXB2 numbering)

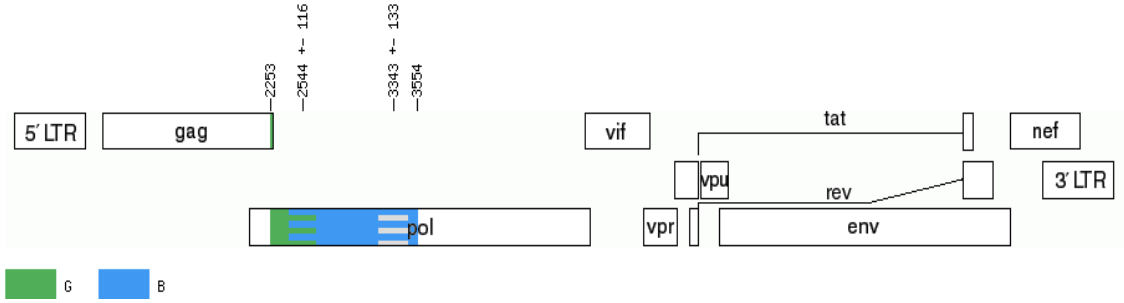


Figure 3.5.1A: Subtype identity of sequence Et_97-58-098 as determined by jpHHM base on reference sequences (A1, A2, B, C, D, F1, F2, G, H, J, 01_AE, O, and CPZ).
Positions indicated here are based on HXB2 numbering.

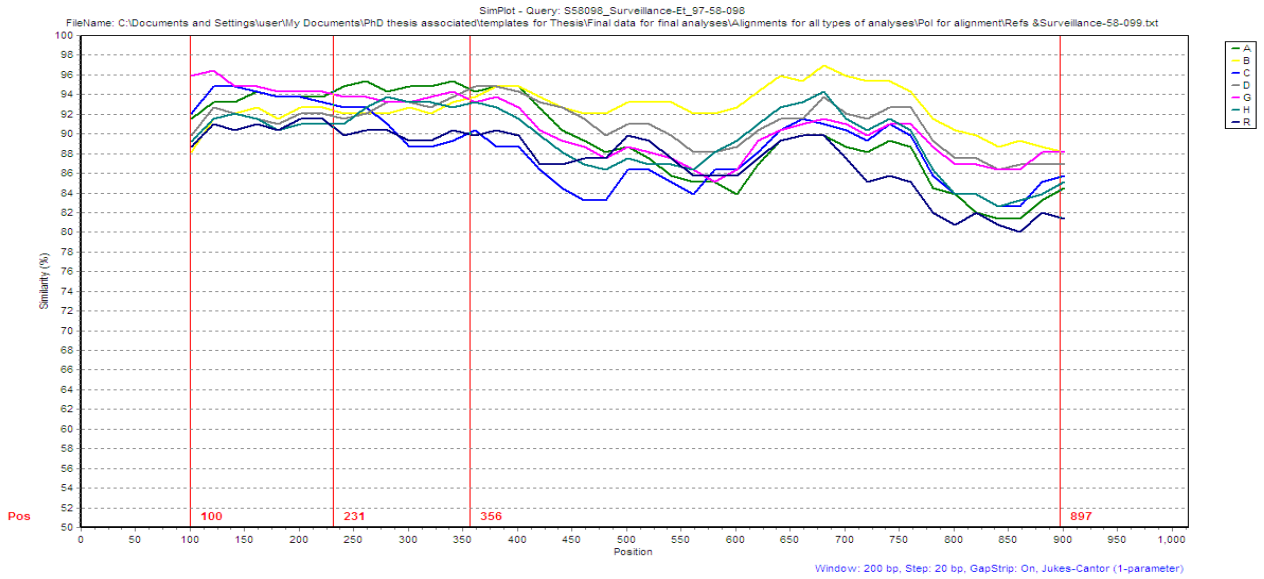


Figure 3.5.1B Subtype identity of sequence Et_97-58-098 as determined by SimPlot base on reference sequences (A, B, C, D, G, H, J).
Positions indicated here are based on true positions in the sequence.

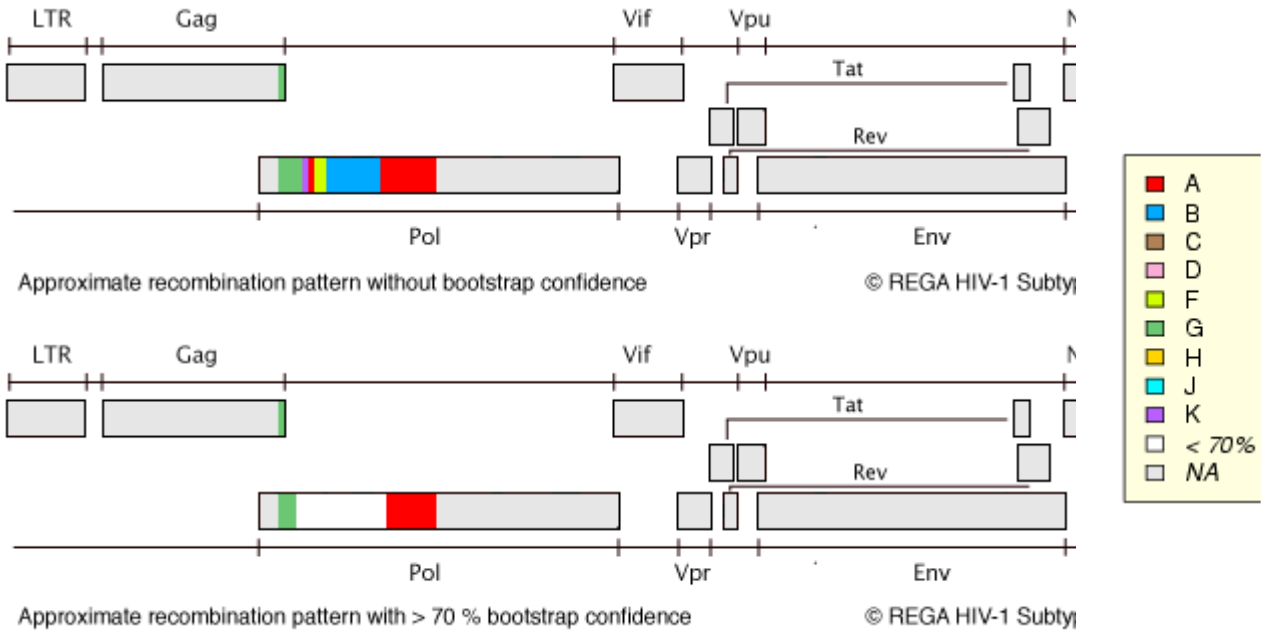


Figure 3.5.1C Subtype identity of sequence Et_97-58-098 as determined byREGS Version 2.0
Clustered with a pure subtype and CRF or sub-subtype with bootstrap >70 %, with detection of recombination in the pure subtype bootscan, and failure to classify as a CRF or sub-subtype by bootscan analysis.

Table 3.5.1. Subtype assignment by five genotyping programs for sequences from recently infected drug-naïve participants

S No.	*Sample ID*	***Sub-typing by Stanford***	***Sub-typing by jpHMM***	***Sub-typing by SimPlot***	***Sub-typing REGA***	*Subt-yping by Phylogenyetic analyses*		
						Pol	**PR**	**RT**
1	Surveillance-ET_97-57-137	C	C	ND	C	C	C	C
2	Surveillance-Et_97-57-195	C	C	ND	C	C	C	C
3	Surveilance-Et_97-57-224	C	C	ND	C	C	C	C
4	Surveilance-Et_97-57-242	C	C	ND	C	C	C	C
5	Surveilance-Cleaning-97-58-004	C	C	ND	C	C	C	C
6	Surveilance-Et_97-58-048	C	C	ND	C	C	C	C
7	Surveilance-Et_97-58-063	C	BC	C	C	C	C	C
8	Surveilance-Et_97-58-098	CRF_02AG	BG	ABG	GKFBA	AG	G	B
9	Surveilance-Et-97-58-116	C	C	ND	C	C	C	C
10	Surveilance-Cleaning-97-58-139	C	C	ND	C	C	C	C
11	Surveilance-Et_97-58-163	C	BC	BC	**BC**	C	C	C
12	Surveilance-Cleaning-97-58-356	C	C	ND	C	C	C	C
13	Surveilance-Et_97-58-477	C	C	ND	C	C	C	C
14	Surveilance-Et_97-59-035	C	C	ND	C	C	C	C
15	Surveilance-Cleaning-97-59-064	C	C	ND	C	C	C	C
16	Surveilance-Cleaning-97-59-147	C	C	ND	C	C	C	C
17	Surveilance-Et_97-59-417	C	C	ND	C	C	C	C
18	Surveilance-Et_97-60-029	C	C	ND	C	C	C	C
19	Surveilance-Cleaning-97-60-108	C	C	ND	C	C	C	C
20	Surveilance-Et_97-60-113	C	C	ND	C	C	C	C
21	Surveilance-Et_97-60-195	C	C	ND	C	C	C	C
22	Surveilance-Et_97-60-208	C	C	ND	C	C	C	C
23	Surveilance-Et_97-60-230	C	C	ND	C	C	C	C
24	Surveilance-Et_97-60-233	C	C	ND	C	C	C	C
25	Surveilance-Et-97-60-237	C	C	ND	C	C	C	C
26	Surveilance-Et_97_60_239	C	C	ND	C	C	C	C
27	Surveilance-Et_97-60-285	C	C	ND	C	C	C	C
28	Surveilance-Et_97-60-317	C	C	ND	C	C	C	C
29	Surveilance-Et_97-60-355	C	C	ND	C	C	C	C
30	Surveilance-Et_97-60-375	C	C	ND	C	C	C	C
31	Surveilance-Et_97-61-056	C	C	ND	C	C	C	C
32	Surveilance-Et_97-61-057	ND	C	ND	C	C	C	C
33	Surveilance-Et_97-61-065	C	C	ND	C	C	C	C
34	Surveilance-Cleaning-97-61-073	C	C	ND	C	C	C	C
35	Surveilance-Cleaning-97-61-099	C	C	ND	C	C	C	C
36	Surveilance-Et_97-61-201	C	C	ND	C	C	C	C
37	Surveilance-Et_97-61-304	C	C	ND	C	C	C	C
38	Surveilance-Et_97-61-315	C	C	ND	C	C	C	C
39	Surveilance-Cleaning-97-61-326	C	C	ND	C	C	C	C
40	Surveilance-Et_97-61-421	C	BC	Inditerminate	C	C	C	C
41	Surveilance-Cleaning-97-61-432	C	C	ND	C	C	C	C

Similarly, the sequence "Surveilance-Et_97-58-063" was identified by jpHHM program to have contained sequences of B and C at the RT and PR regions respectively (Figure 3.5.2A). However, Stanford program and phylogenetic analyses have identified the sequence to belong to subtype C at both regions. In agreement with the latter analyses, SimPlot also identified the sequence as subtype C in both the PR and RT regions (Figure 3.5.2B). The third source of discrepancy was "Surveilance-Cleaning-97-58-163", which was identified as subtype C by Stanford program and the phylogenetic analysis but BC by jpHHM (Figures 3.5.3A). REGA program also identified it as subtype C, but showed indication of recombination between largely C at the 5'- end and B at the -3' end. When SimPLot was run, the result was in agreement with that of jpHHM, namely BC (Figure 3.5.3B).

Sequence #2: >Et_97-58-063_Viroseq-C

This sequence is related to subtype(s): **B C**

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	-	-	413	C
414	-	-	1302	B
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2253	-	-	2665	C
2666	-	-	3554	B

Genome map (based on HXB2 numbering)

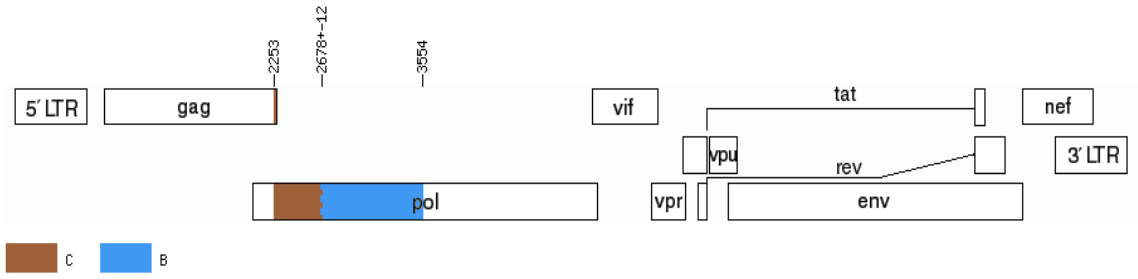


Figure 3.5.2A: Subtype identity of sequence Et_97-58-063 as determined by jpHHM base on reference sequences (A1, A2, B, C, D, F1, F2, G, H, J, 01_AE, O, and CPZ).
Positions indicated here are based on HXB2 numbering.

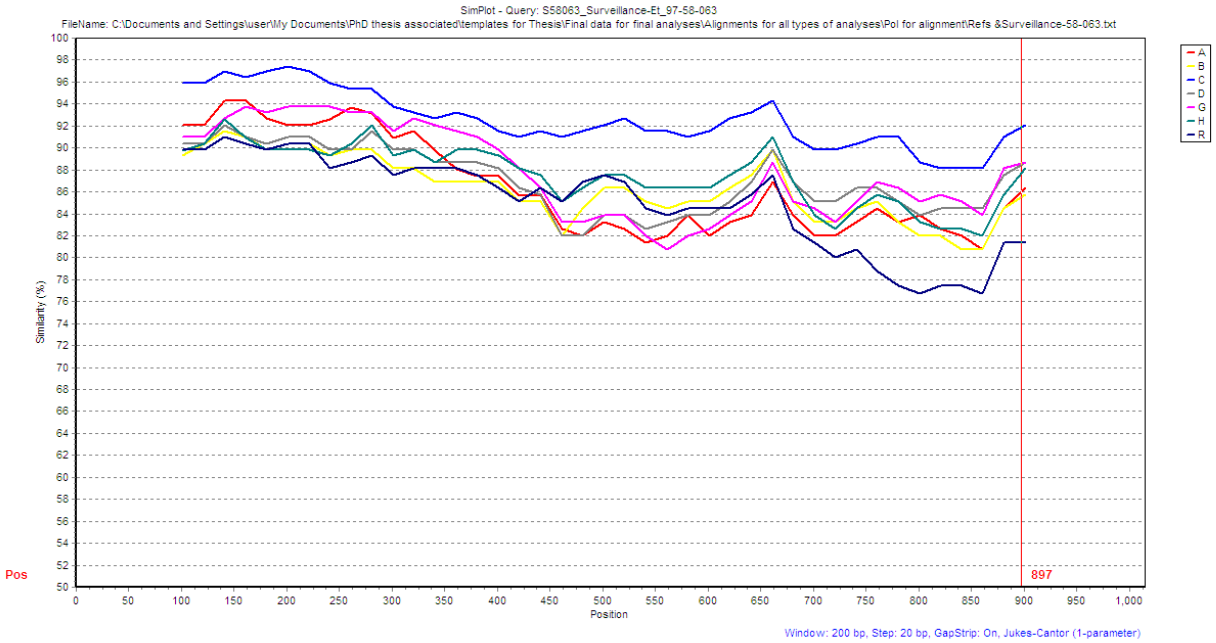


Figure 3.5.2B: Subtype identity of sequence Et_97-58-063 as determined by SimPlot as subtype C base on reference sequences (A, B, C, D, G, H, J).
Positions indicated here are based true positions in the sequence.

Sequence #7: >Et_97-58-163_Viroseq-C

This sequence is related to subtype(s): **B C**

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 - 257	732 - 821	**780**	**C**
781	-	-	**1196**	**B**
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2253	2253 - 2509	2984 - 3073	**3032**	**C**
3033	-	-	**3448**	**B**

Genome map (based on HXB2 numbering)

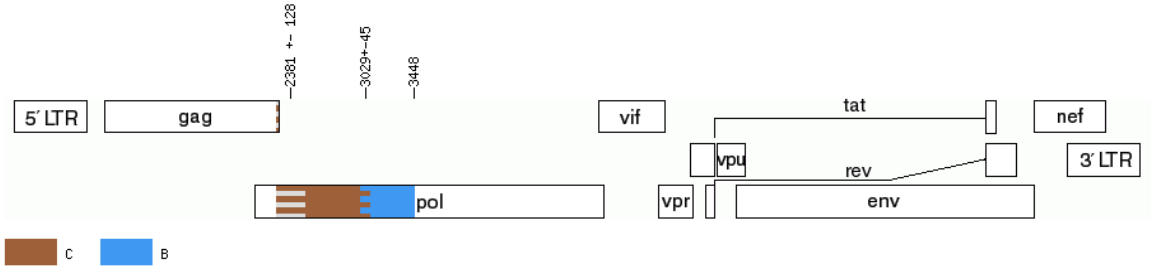


Figure 3.5.3A: Subtype identity of sequence Et_97-58-163 as determined by jpHHM base on reference sequences (A1, A2, B, C, D, F1, F2, G, H, J, 01_AE, O, and CPZ).
Positions indicated here are based on HXB2 numbering.

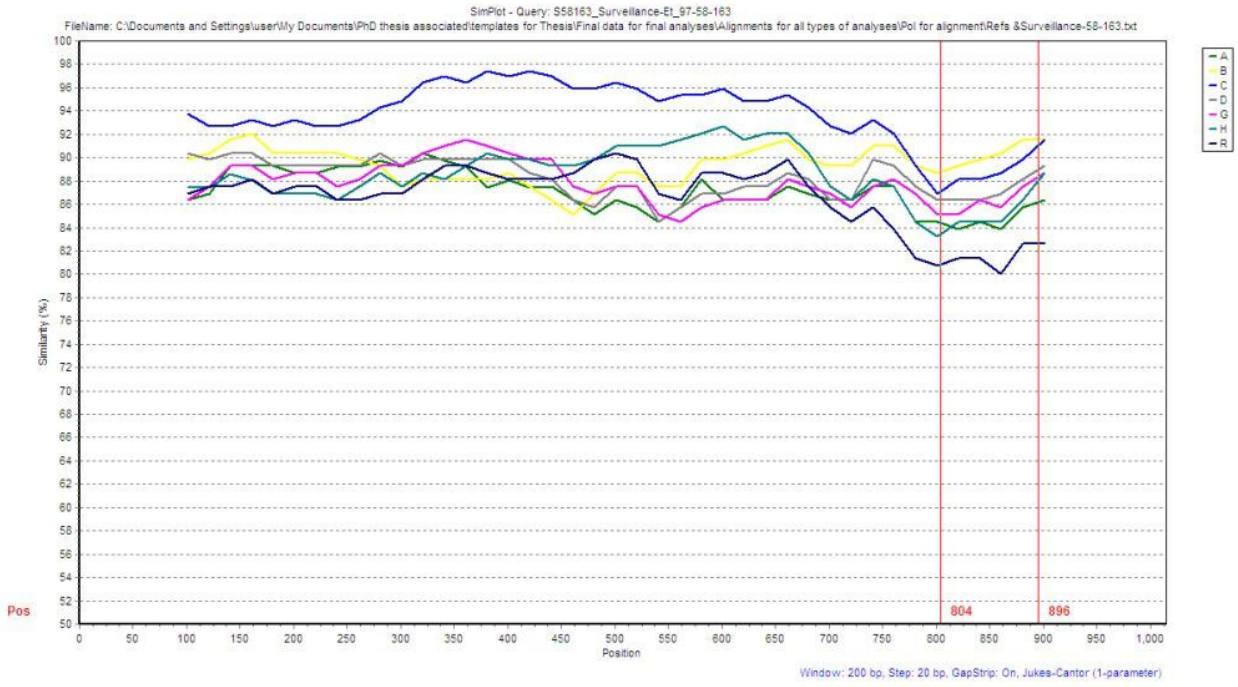


Figure 3.5.3B: Subtype identity of sequence Et_97-58-163 as determined by SimPlot as subtype C base on reference sequences (A, B, C, D, G, H, J).
Positions indicated here are based on true positions in the sequence.

The last two problematic sequences were “Et_97-61-057” and “Et_97-61-421”. The former could not be genotyped by the Stanford program, but was sub-typed as C by phylogenetic analysis, jpHHM, and REGA (not shown). SimPlot analysis could not clearly define the subtype of “Et_97-61-057”. On the other hand, sequence “Et_97-61-421” was identified as BC by jpHMM (Figure 3.5.4A) and subtype C by the other programs except SimPlot, which could not clearly determine the subtype (Figure 3.5.4B).

Sequence #4: >Et_97-61-421_Viroseq-C

This sequence is related to subtype(s): **B C**

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	-	64 - 95	**64**	**B**
65	-	414 - 500	**423**	**C**
424	-	-	**1302**	**B**
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2253	-	2316 - 2347	**2316**	**B**
2317	-	2666 - 2752	**2675**	**C**
2676	-	-	**3554**	**B**

Genome map (based on HXB2 numbering)

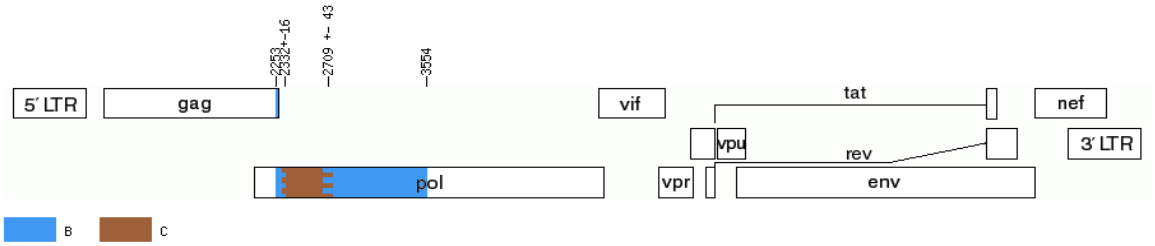


Figure 3.5.4A: Subtype identity of sequence Et_97-61-421 as determined by jpHHM base on reference sequences (A1, A2, B, C, D, F1, F2, G, H, J, 01_AE, O, and CPZ).
Positions indicated here are based on HXB2 numbering.

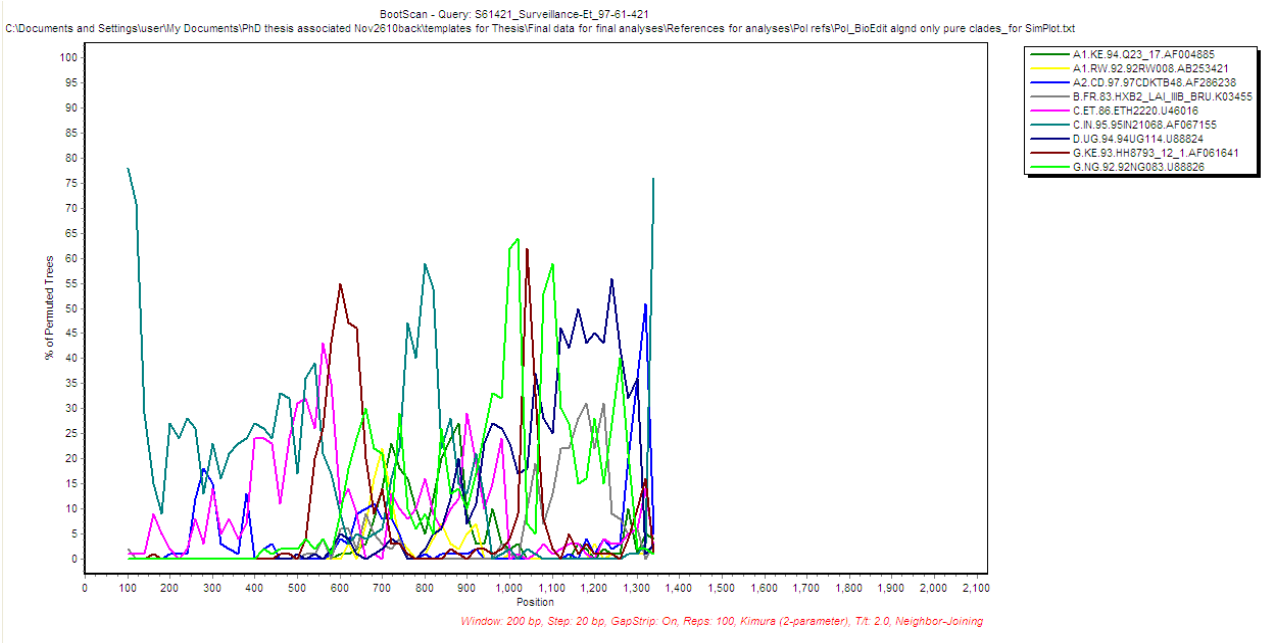


Figure 3.5.4B: Subtype identity of sequence Et_97-58-421 as determined by SimPlot as subtype C base on reference sequences (A, B, C, D, G, H, J).
Positions indicated here are based on true positions in the sequence.

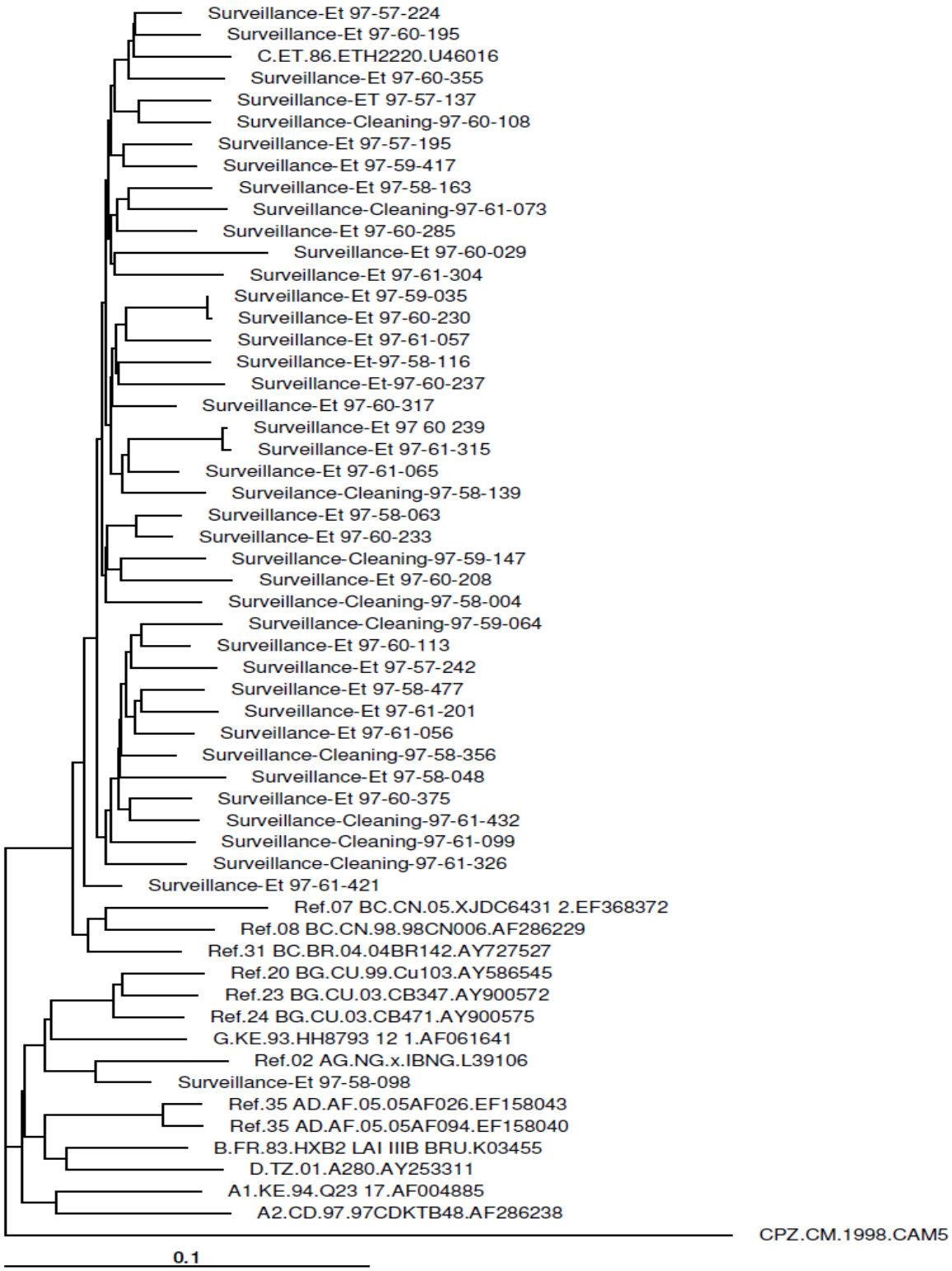


Figure 3.5.5 Phylogenetic relationship of research sequence with pure clade and CRF references at the Pol region

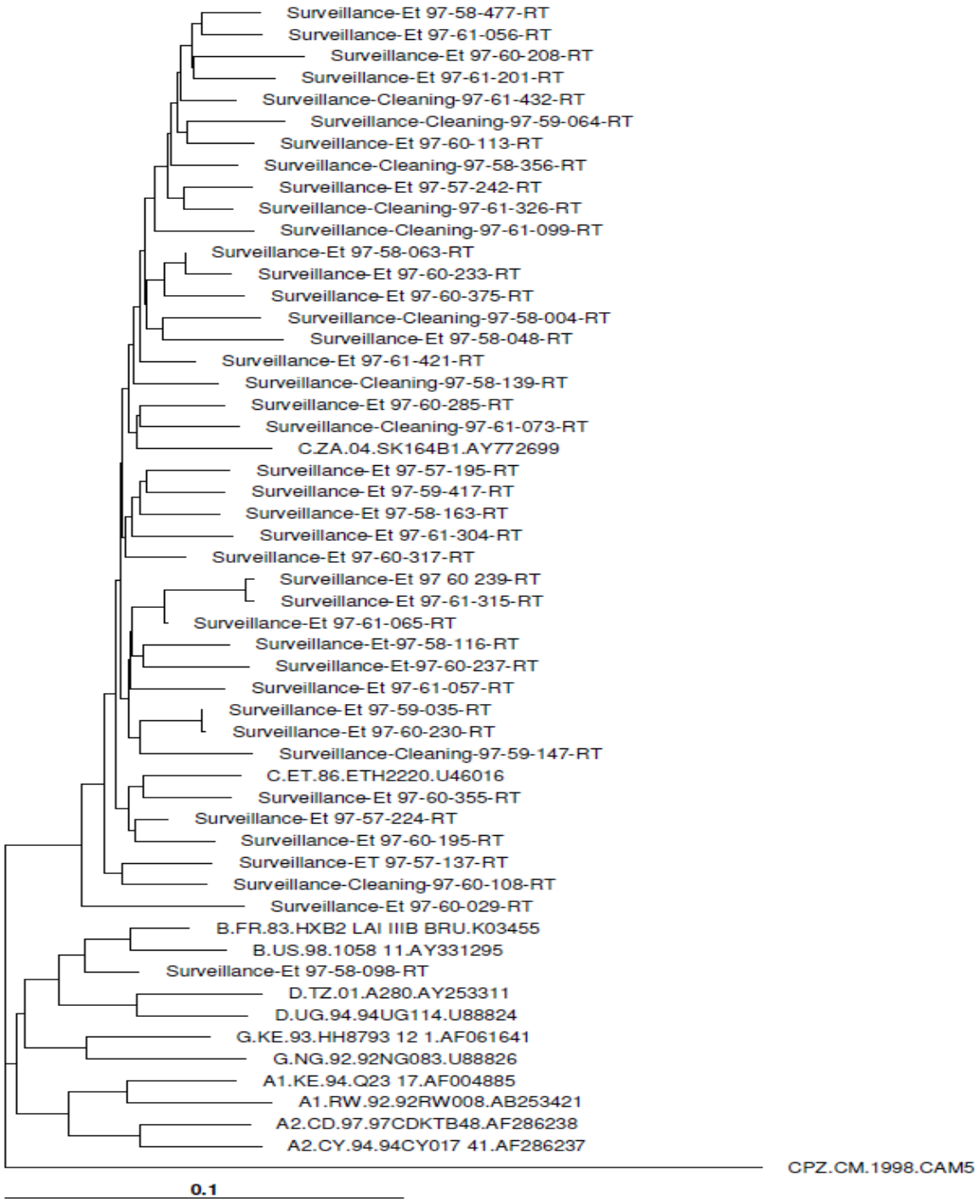


Figure 3.5.6 Phylogenetic relationships of sequences from recently infected drug-naïve patients at the RT region.

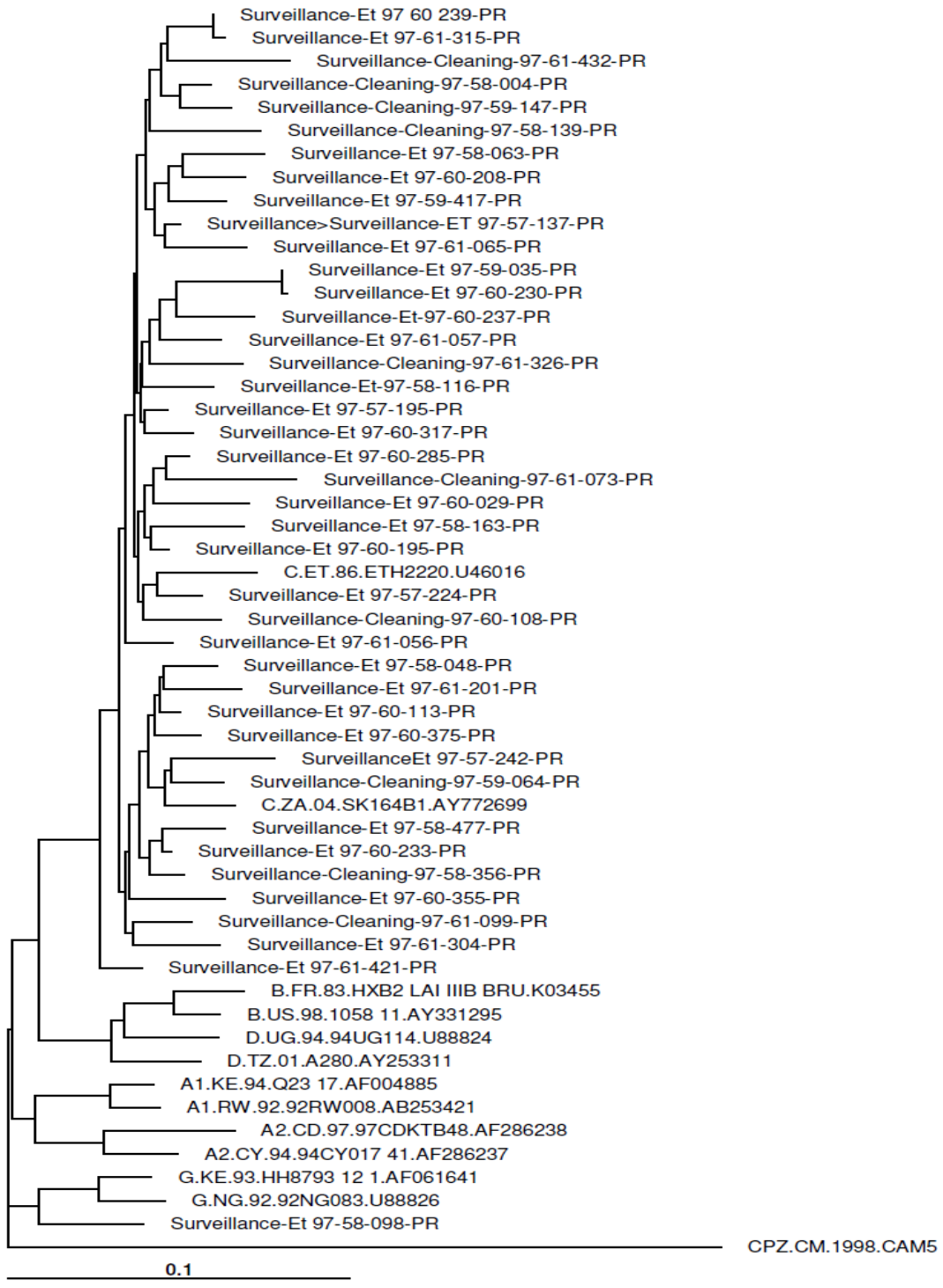


Figure 3.5.7: Phylogenetic relationships of sequences from recently infected drug-naïve patients at the PR region

3.5.1.2 Subtype diversity of HIV-1 isolates from chronically infected drug-naïve participants

A total of 84 samples from chronically infected drug naïve patients were successfully sequenced and subtyped. While 38 of these (45%) were genotyped by the ViroSeq^{TM} assay, the rest 46 (55%) were genotyped by the In-house assay. Regarding source of samples for genotyping, 22 (26%) of the sequences used in this analysis were taken from DBS while 62 (74%) were from plasma. All algorithms used in subtyping the sequences have agreed 100% on the identity of the sequences in this group (Table 3.5.2 and Figures 3.5.8). All of them belonged subtype C at PR, RT, and combined Pol region. Therefore there was no need to run the SimPlot program. This unanimous assignment to subtype C was achieved in spite of differences in genotyping assays (In-house vs ViroSeq) and sample sources (plasma vs DBS) used.

Figure 3.5.8 Phylogenetic relationship of research sequences with pure clade and CRF reference sequences

Table 3.5.2 Subtype assignment by four genotyping programs for sequences from chronically infected drug-naïve participants

Sample ID	*Subtype by Stanford*	*Subtype by jpHMM*	*REGA*	*Subtype by Phylogeny*		
				Pol	PR	RT
Baseline-102ViSq-plasma	C	C	C	C	C	C
Baseline-115ViSq-plasma	C	C	C	C	C	C
Baseline-154ViSq-plasma	C	C	C	C	C	C
Baseline-155ViSq-plasma	C	C	C	C	C	C
Baseline-156ViSq-plasma	C	C	C	C	C	C
Baseline-157ViSq-plasma	C	C	C	C	C	C
Baseline-165ViSq-plasma	C	C	C	C	C	C
Baseline175ViSq-plasma	C	C	C	C	C	C
Baseline-183ViSq-plasma	C	C	C	C	C	C
Baseline-190ViSq-plasma	C	C	C	C	C	C
Baseline-192ViSq-plasma	C	C	C	C	C	C
Baseline-197ViSq-plasma	C	C	C	C	C	C
Baseline-200ViSq-plasma	C	C	C	C	C	C
Baseline203ViSq-plasma	C	C	C	C	C	C
Baseline-204ViSq-plasma	C	C	C	C	C	C
Baseline-205ViSq-plasma	C	C	C	C	C	C
Baseline-211BViSq-plasma	C	C	C	C	C	C
Baseline-212ViSq-plasma	C	C	C	C	C	C
Baseline214ViSq-plasma	C	C	C	C	C	C
Baseline220ViSq-plasma	C	C	C	C	C	C
Baseline-226ViSq-plasma	C	C	C	C	C	C
Baseline-227ViSq-plasma	C	C	C	C	C	C
Baseline-230ViSq-plasma	C	C	C	C	C	C
Baseline 234InHouse-DBS	C	C	C	C	C	C
Baseline-235InHouse-DBS	C	C	C	C	C	C
Baseline-236InHouse-plasma	C	C	C	C	C	C
Baseline-237Inhouse-plasma	C	C	C	C	C	C
Baseline-238Inhouse-DBS	C	C	C	C	C	C
Baseline-239Inhouse-DBS	C	C	C	C	C	C
Baseline-240Inhouse-plasma	C	C	C	C	C	C
Baseline-241ViSq-plasma	C	C	C	C	C	C
Baseline-253Inhouse-DBS	C	C	C	C	C	C
Baseline-258Inhouse-DBS	C	C	C	C	C	C
Baseline-259Inhouse-plasma	C	C	C	C	C	C
Baseline-264Inhouse-DBS	C	C	C	C	C	C
Baseline-265Inhouse-DBS	C	C	C	C	C	C
Baseline-266Inhouse-plasma	C	C	C	C	C	C
Baseline-269Inhouse-pasma	C	C	C	C	C	C
Baseline270ViSq-plasma	C	C	C	C	C	C
Baseline-271Inhouse-DBS	C	C	C	C	C	C
Baseline-273Inhouse-DBS	C	C	C	C	C	C
Baseline-274Inhouse-plasma	C	C	C	C	C	C

Table 3.5.2 …Continued

Sample ID	*Subtype by Stanford*	*Subtype by jpHMM*	*REGA*	*Subtype by Phylogeny*		
				Pol	PR	RT
Baseline-275ViSq-plasma	C	C	C	C	C	C
Baseline-281Inhouse-plasma	C	C	C	C	C	C
Baseline-Cleaning-286Inhouse-p	C	C	C	C	C	C
Baseline-287Inhouse-plasma	C	C	C	C	C	C
Baseline-288Inhouse-DBS	C	C	C	C	C	C
Baseline-289ViSq-plasma	C	C	C	C	C	C
Baseline-290Inhouse-plasma	C	C	C	C	C	C
Baseline-291ViSq-plasma	C	C	C	C	C	C
Baseline-293Inhouse-DBS	C	C	C	C	C	C
Baseline-295Inhouse-DBS	C	C	C	C	C	C
Baseline-296Inhouse-DBS	C	C	C	C	C	C
Baseline-297Inhouse-DBS	C	C	C	C	C	C
Baseline-306Inhouse-DBS	C	C	C	C	C	C
Baseline-311Inhouse-plasma	C	C	C	C	C	C
Baseline-312ViSq-plasma	C	C	C	C	C	C
Baseline-314Inhouse-plasma	C	C	C	C	C	C
Baseline-315Inhouse-plasma	C	C	C	C	C	C
Baseline-318Inhouse-DBS	C	C	C	C	C	C
Baseline-319Inhouse-plasma	C	C	C	C	C	C
Baseline-320Inhouse-plasma	C	C	C	C	C	C
Baseline-321Inhouse-plasma	C	C	C	C	C	C
Baselien-Cleaning-322Inhouse-p	C	C	C	C	C	C
Baseline-325Inhouse-DBS	C	C	C	C	C	C
Baseline-326Inhouse-plasma	C	C	C	C	C	C
Baseline-329Inhouse-DBS	C	C	C	C	C	C
Baseline-330Inhouse-DBS	C	C	C	C	C	C
Baseline-332ViSq-plasma	C	C	C	C	C	C
Plasma-baseline-333DR-inhouse	C	C	C	C	C	C
Baseline-334ViSq-plasma	C	C	C	C	C	C
Baseline-337Inhouse-DBS	C	C	C	C	C	C
Baseline-339ViSq-plasma	C	C	C	C	C	C
Baseline-342ViSq-plasma	C	C	C	C	C	C
Baseline-343Inhouse-DBS	C	C	C	C	C	C
Baseline-345ViSq-plasma	C	C	C	C	C	C
Baseline-346Inhouse-DBS	C	C	C	C	C	C
Baseline-347ViSq-plasma	C	C	C	C	C	C
Baseline-349Inhouse-plasma	C	C	C	C	C	C
Baseline-356ViSq-plasma	C	C	C	C	C	C
Baseline-360Inhouse-plasma	C	C	C	C	C	C
Baseline-361Inhouse-DBS	C	C	C	C	C	C
Baseline-363ViSq-plasma	C	C	C	C	C	C
Baseline-364ViSq-plasma	C	C	C	C	C	C

3.5.1.3 Subtype diversity of HIV-1 isolates from chronically infected drug-experienced participants

Sixty-six samples from this group of patients were successfully sequenced and genotyped using the In-house genotyping assay. Sequence analyses by the four genotyping programs produced concordant subtyping results in the majority of the sequences (64/66 [96.96%]). All but two of the concordant genotyping results identified the sequences as subtypes C (62/66 [93.93%]), the two exceptions (2/66 [4.59%]) being identified as subtype B by all the four programs. Summaries of the genotyping results are found in Table 3.5.3 and Figure 3.5.9-3.5.11. For the remaining two of the 66 sequences (Table 3.4.1.3) for which discordant genotyping results was observed, SimPlot analysis was run to resolve the discrepancies. In one of these cases (sequence "Experienced-V1311DRZ2"), it was shown by jpHHM program that it contained a nucleotide sequence belonging to subtype C in the first 45 nucleotides of the 5'-end, while the remaining long fragment of the 3', including the PR and RT regions, contained sequences belonging to subtype B (Figure 3.5.12A). However, Stanford, REGA and phylogenetic analyses at PR and RT separately and in combination at Pol have identified this sequence to belong to subtype C. SimPlot analysis could not resolve the discordance observed here; rather, it produced a result difficult to interpret (Figure 3.5. 12B).

Table 3.5.3 Subtype assignment by four genotyping programs for sequences from chronically infected drug-experienced participants.
ND= Not Determined

Sequence ID	*Subtyping by Stanford*	*Subtyping by jpHMM*	*Subtyping by REGA*	*Subtyping by SimPlot*	*Subtyping by Phylogeny*		
					Pol	PR	RT
Experienced-V1309DRZ2	C	C	C	ND	C	C	C
Experienced-V1311DRZ2	C	**B**C	C	Indeterminate	C	C	C
Experienced-V1312DRZ	C	C	C	ND	C	C	C
Experienced-V1315DRZ	C	C	C	ND	C	C	C
Experienced-V1316DRZ	C	C	C	ND	C	C	C
Experienced-V1317DRZ	C	C	C	ND	C	C	C
Experienced-V1320DRZ	C	C	C	ND	C	C	C
Experienced-V1323DRZ	C	C	C	ND	C	C	C
Experienced-Cleaning-V1324	C	C	C	ND	C	C	C
Experienced-V1329DRZ	C	C	C	ND	C	C	C
Experienced-V1330DRZ	C	C	C	ND	C	C	C
Experienced-V1334DRZ	C	C	C	ND	C	C	C
Experienced-Cleaning-V1342	C	C	C	ND	C	C	C
Experienced-V1343DRZ	B	B	B	ND	B	B	B
Experienced-V1344DRZ	C	C	C	ND	C	C	C
Experienced-V1349DRZ	D	A1D	NA	A1D	D	A1	D
Experienced-Cleaning-V1351	C	C	C	ND	C	C	C
Experienced-Cleaning-V1355	C	C	C	ND	C	C	C
Experienced-V1356DRZ	C	C	C	ND	C	C	C
Experienced-Cleaning-V1358	C	C	C	ND	C	C	C
Experienced-V1362DRZ	C	C	C	ND	C	C	C
Experienced-V1363DRZ	C	C	C	ND	C	C	C
Experienced-V1365DRZ	C	C	C	ND	C	C	C
Experienced-V1370DRZ	C	C	C	ND	C	C	C
Experienced-V1374DRZ	C	C	C	ND	C	C	C
Experienced-V1376DRZ	C	C	C	ND	C	C	C
Experienced-Cleaning-V1378	C	C	C	ND	C	C	C
Experienced-V1382DRZ	C	C	C	ND	C	C	C
Experienced-V1383DRZ	C	C	C	ND	C	C	C
Experienced-V1389DRZ	C	C	C	ND	C	C	C
Experienced-V1392DRZ	C	C	C	ND	C	C	C
Experienced-V1393DRZ	C	C	C	ND	C	C	C
Experienced-Cleaning-V1394	C	C	C	ND	C	C	C

Table 3.5.3 …. continued

Sequence ID	*Subtyping by Stanford*	*Subtyping by jpHMM*	*Subtyping by REGA*	*Subtyping by SimPlot*	*Subtyping by Phylogeny*		
					Pol	PR	RT
Experienced-V1404DRZ	C	C	C	ND	C	C	C
Experienced-V1408DRZ	C	C	C	ND	C	C	C
Experienced-V1415DRZ	C	C	C	ND	C	C	C
Experienced-Cleaning-V1421	C	C	C	ND	C	C	C
Experienced-V1430DRZ	C	C	C	ND	C	C	C
Experienced-V1431DRZ	B	B	B	ND	B	B	B
Experienced-V1435DRZ	C	C	C	ND	C	C	C
Experienced-V1437DRZ2	C	C	C	ND	C	C	C
Experienced-V1442DRZ	C	C	C	ND	C	C	C
Experienced-Cleaning-V1448	C	C	C	ND	C	C	C
Experienced-Cleaning-V1604	C	C	C	ND	C	C	C
Experienced-V1607DRBL	C	C	C	ND	C	C	C
Experienced-V1626DRBL	C	C	C	ND	C	C	C
Experienced-V1641DRBL	C	C	C	ND	C	C	C
Experienced-V1660DRBL	C	C	C	ND	C	C	C
Experienced-V1666DRBL	C	C	C	ND	C	C	C
Experienced-V1667DRBL	C	C	C	ND	C	C	C
Experienced-V1676DRBL	C	C	C	ND	C	C	C
Experienced-V1679DRBL	C	C	C	ND	C	C	C
Experienced-V1682DRBL	C	C	C	ND	C	C	C
Experienced-Cleaning-V1692	C	C	C	ND	C	C	C
Experienced-V1694DRBL	C	C	C	ND	C	C	C
Experienced-V1696DRBL	C	C	**C**	ND	C	C	C
Experienced-V1699DRBL	C	C	C	ND	C	C	C
Experienced-V1505	C	C	C	ND	C	C	C
Experienced-V1512	C	C	C	ND	C	C	C
Experienced-V1514	C	C	C	ND	C	C	C
Experienced-V1955	C	C	C	ND	C	C	C
Experienced-V1971	C	C	C	ND	C	C	C
Experienced-V1905	C	C	C	ND	C	C	C
Experienced-V1578	C	C	C	ND	C	C	C
Experienced-V1580	C	C	C	ND	C	C	C
Experienced-V1587	C	C	C	ND	C	C	C

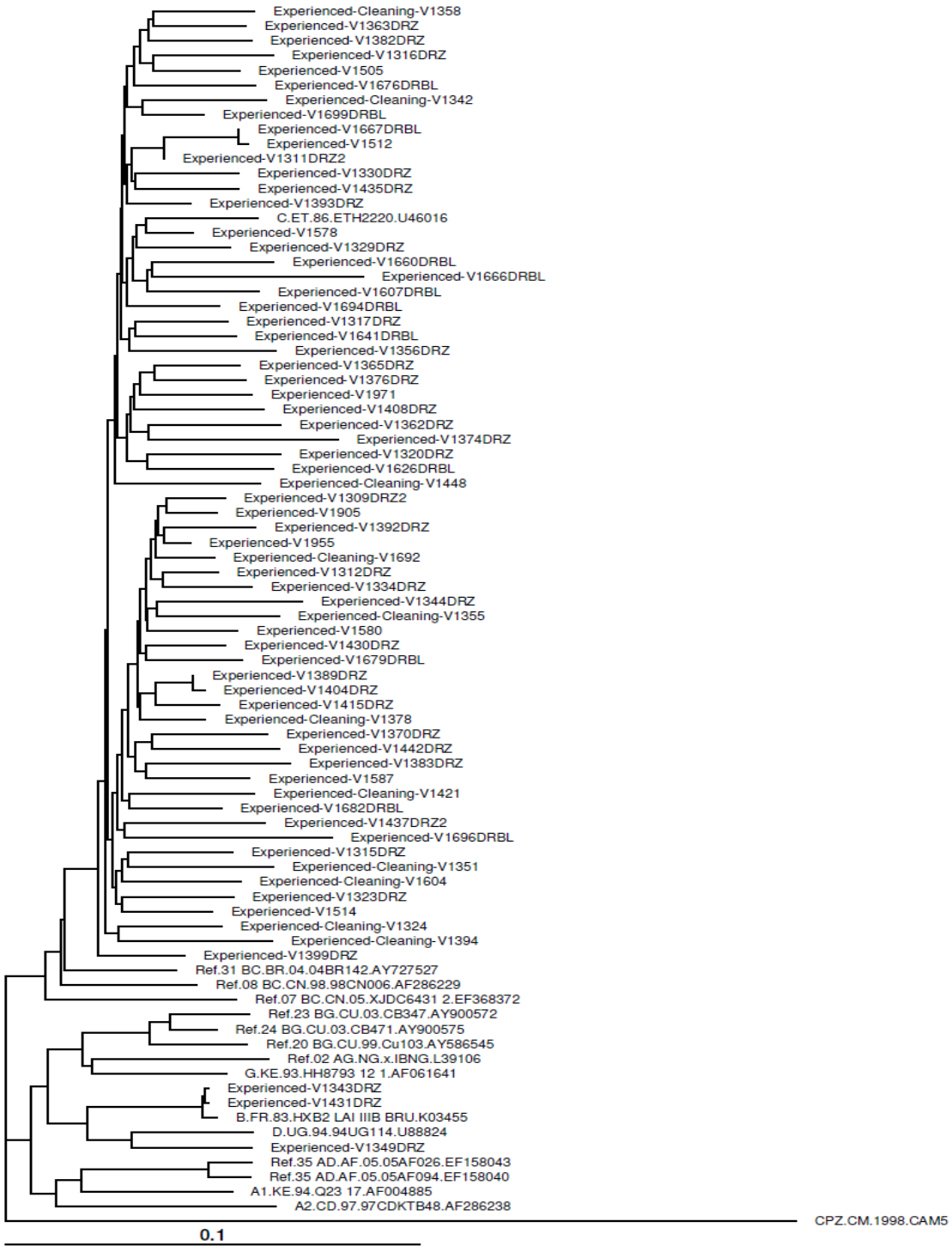


Figure 3.5.9 Phylogenetic relationship of sequences in the Pol region from chronically infected drug-experienced participants with pure clade and CRF reference sequences at Pol region

Figure 3.5.10 Phylogenetic relationship of sequences from chronically infected drug-experienced participants with pure clade and CRF reference sequences at PR region

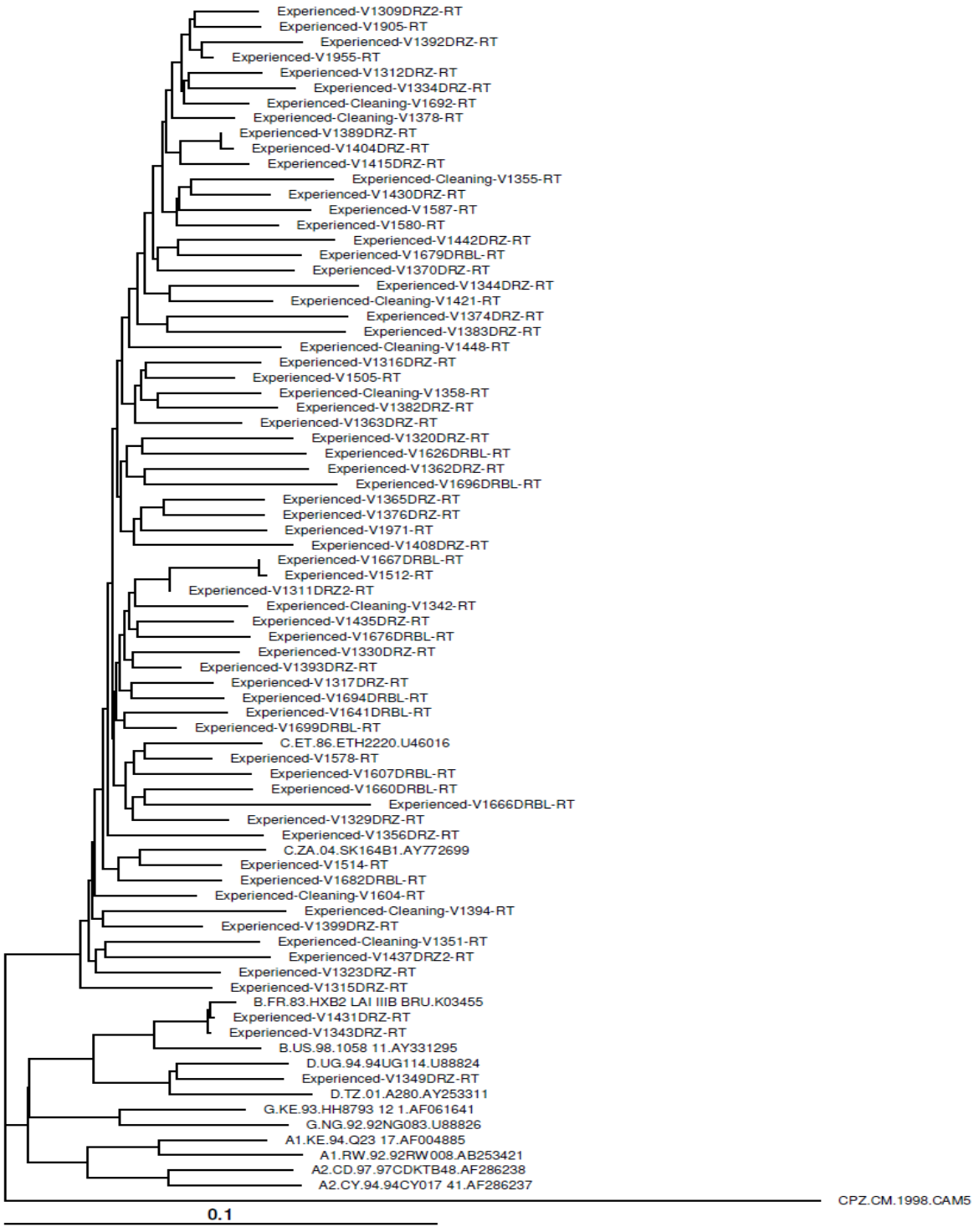


Figure 3.5.11 Phylogenetic relationship of sequences from chronically infected drug-experienced participants with pure clade and CRF reference sequences at RT region

Sequence #2: >V1311DRZ2

This sequence is related to subtype(s): **B C**

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 - 45	-	**45**	**C**
46	46 - 67	-	**1037**	**B**
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2277	2277 - 2321	-	**2321**	**C**
2322	2322 - 2343	-	**3311**	**B**

Genome map (based on HXB2 numbering)

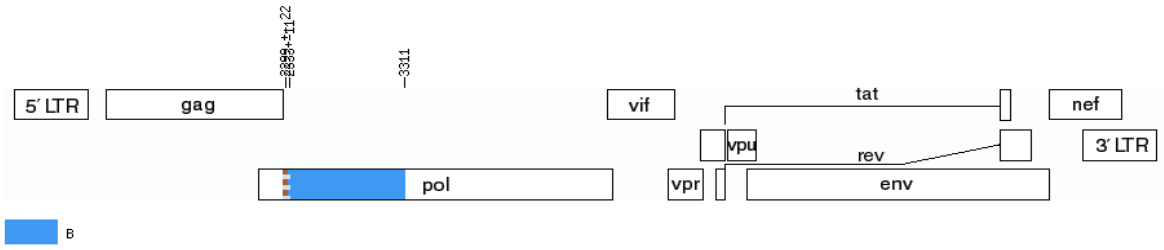


Figure 3.5.12A Subtype identity of sequence Experienced-V1311DRZ2 as determined by jpHHM base on reference sequences (A1, A2, B, C, D, F1, F2, G, H, J, 01_AE, O, and CPZ).
Positions indicated here are based on HXB2 numbering.

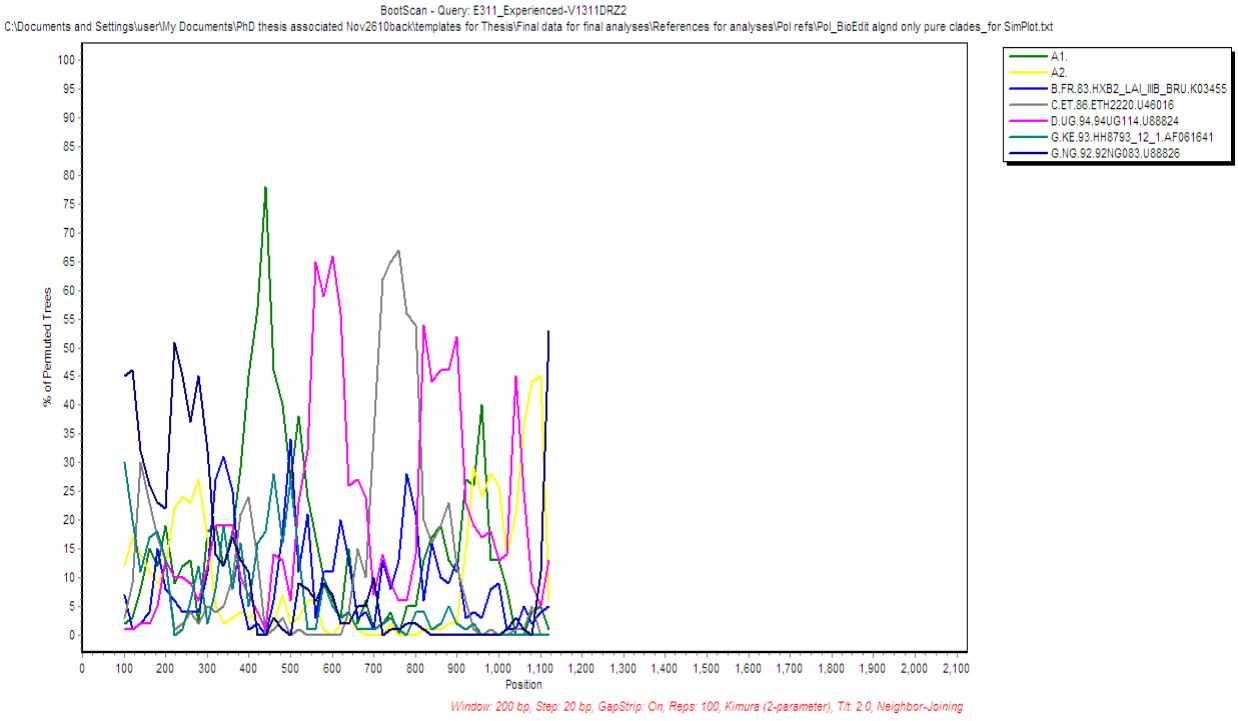


Figure 3.5.12B: Subtype identity of sequence Et_97-58-421 as determined by SimPlot as subtype C base on reference sequences (A, B, C, D, G, H, J).
Positions indicated here are based on true positions in the sequence.

The other discrepant genotyping result was obtained from sequence "Experienced-V1349DRZ", whose subtype could not be determined by the REGA program, but was identified as subtype D by Stanford and phylogenetic analysis at the Pol region (Table 3.5.3 and Figure 3.5.9). Phylogenetic analyses done on PR and RT regions separately have identified these sequences to belong to subtypes A1 and D, respectively. jpHHM and SimPlot, on the other hand identified, this sequence as a recombinant A1D (Figures 3.5.13A and 3.5.13B), a complete agreement with the separately done phylogenetic analyses for the PR and RT regions.

Sequence #2: >V1349DRZ2

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

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	-	292 - 330	317	A1
318	-	-	1039	D
Position based on HXB2 numbering [pred_recombination] [recombination_incl_UR_and_BPI] [UR_and_BPI]				
2277	-	2568 - 2606	2593	A1
2594	-	-	3315	D

Genome map (based on HXB2 numbering)

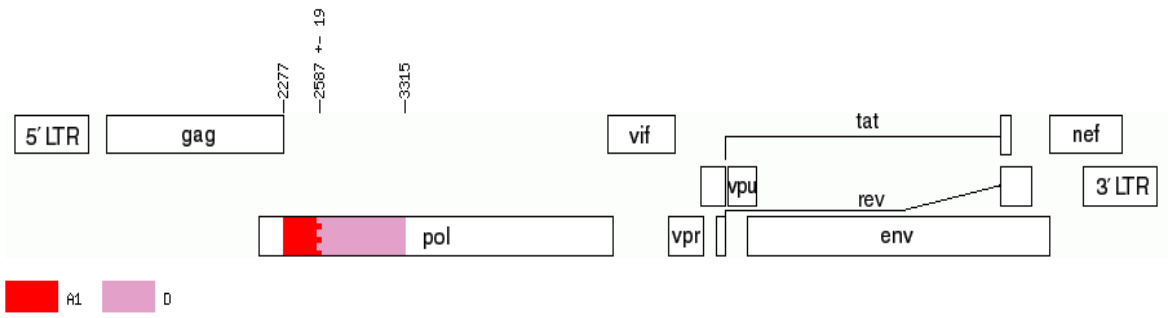


Figure 3.5.13A: Subtype identity of sequence Experienced-V1349DRZ2 as determined by jpHHM base on reference sequences (A1, A2, B, C, D, F1, F2, G, H, J, 01_AE, O, and CPZ).
Positions indicated here are based on HXB2 numbering.

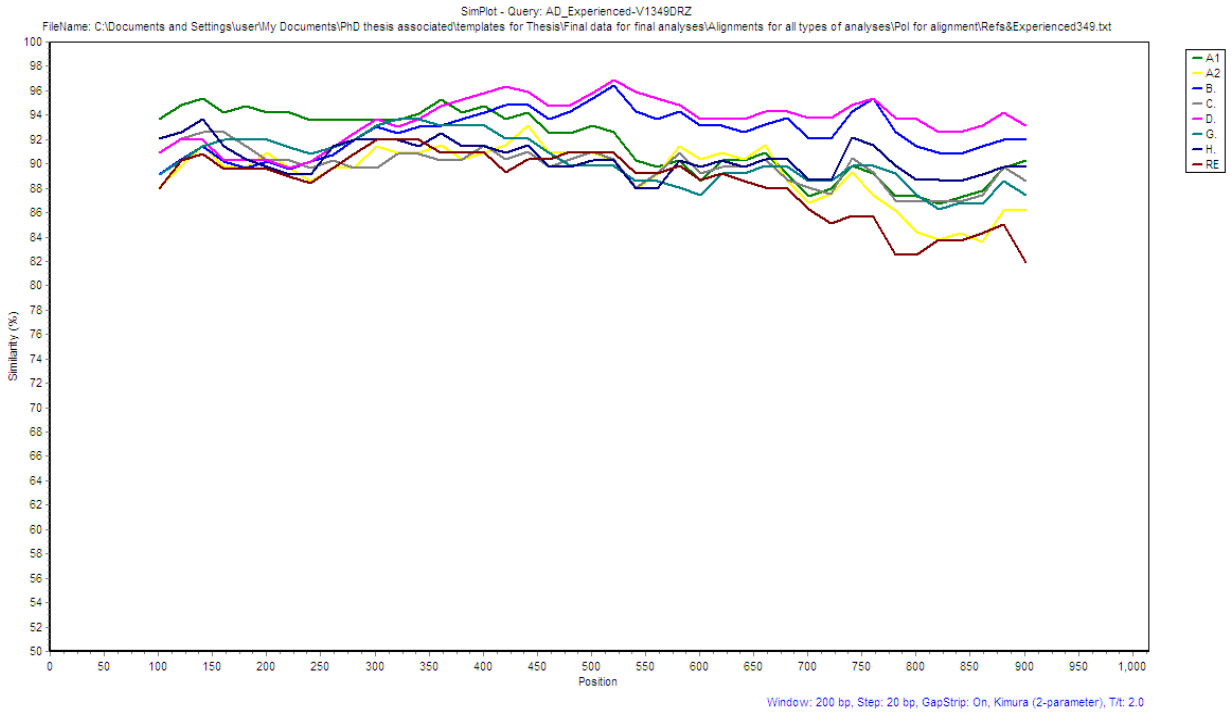


Figure 3.5.13B Subtype identity of sequence Experienced-V1349DRZ2 as determined by SimPlot as subtype A1D base on reference sequences (A, B, C, D, G, H, J).
Positions indicated here are based on true positions in the sequence.

3.5.2 Overall Comparisons between Sequences of the three Groups.

3.5.2.1 HIV-1 Non-drug resistance mutations at drug resistance and non-drug resistance sites

Genotypic analysis has identified a total of 24 non-drug resistance mutations among the recently infected study participants at the drug resistance sites of the Pol region. Of these, 18 cases involved four NNRTI non-resistance mutations at three resistance sites (A98S (15/41[37%]); K101R (1/41[2%]); K101Q (1/41[2%]); and V179I (1/41[2%])). One and five cases, respectively, were non-resistance at NRTI (K65E (1/41[2%])) and PI resistance positions (V82I (5/41[12%])). In addition, over 47 non-resistance mutations occurred at non-resistance sites of the protease region. The most common of these non-resistance sites were T12S (16/41[39%]); K14R (10/41[24%]); L19I (11/41[27%]); I15V (32/41[78%]); R41K (39/41[95%]); and L89M (38/41[93%]). Similarly, over 55 non-resistance mutations occurred at non-resistance sites of the

reverse transcriptase region, the most common ones being V35T (35/41[85%]); E36A (22/41[54%]); T39E (30/41[73%]); K122E (35/41[85%]); D123S (21/41[51%]); D123G (11/41[27%]); A158S (22/41[54%] K173A (35/41[85%]); K173T (14/41[34%]); and D177E (35/41[85%]).

Among the chronically infected drug-naïve participants, 24 PR mutations considered by one of the three algorithms as resistance but not by the other two algorithms were recorded. The most common of these were M36I (81/85[95%]); L63P (29/85[34%]); H69K (77/85[91%]); and I93L (81/85[95%]). Five more non-resistance mutations were observed at drug resistance sites of the PR region: V32E (1/85[1%]); N83T (1/85[1%]); N88G (3/85); N88K (1/85[4%]); and L90W (1/85[1%]). In addition, over 34 non-resistance mutations at PR non-resistance sites occurred among all of participants in this study group. The four most prevalent mutations from this group were I15V (66/85[78%]); L19I (36/85[42%]); R41K (77/85[91%] and L89M (69/85[81%]). With regards to occurrence of mutations at the NRTI resistance sites, four have been detected, including M41W (1/85[1%]); D67E (1/85[1%]); V179I (1/85[1%]); M184L (3/85[4%]). One mutation, A62V, which is considered as resistance only by IAS-USA but not by others was also detected in the 3/85 (4%) of the 85 sequences. Similarly, five non-resistance mutation types were observed at NNRTI resistance sites: V90G (1/85[1%]); A98S (26/85[31%]); L100V (1/85[1%]); K101R (2/85[2%]); and V179I (1/85[1%]). Fifty-two mutations at non-resistance sites to NRTIs and NNRTIs were observed in the RT region of the Pol. Among them, those mutations with prevalence rate of 20% or more include: V35T (74/85[87%]); E36A 43/85[51%); T39E (67/85[69%]); S48T (29/85[34%]); K122E (79/85[93%]); D123S (44/85[52%]); D123G (19/85[22%]); I135T (18/85[21%]); I158S (34/85[40%]); T166R (17/85[20%]); K173A (68/85[80%]); Q174K (35/85[41%]); D177E (69/85[81%]); G196K (17/85[20%]); T200A

(76/85[89%]); Q207E (62/85[73%]); R211K (61/85[72%]); V245Q (75/85[88%]); E248D (18/85[21%]); and D250E (23/85[27%]).

In chronically infected drug-experienced participants, six cases of a single non-resistance mutation was observed at resistance site of the protease region, V82I (6/66[9%]). The non-resistance sites of the protease region also harbored numerous non-resistance mutations, among which the most common ones were T12S (60/66[91%]); I15V (48/66[73%]); L19I (19/66[29%]); R41K (63/66[95%]); and L89M (51/66[77]). Likewise, four non-resistance mutations were observed at NRTI resistance sites, including T69K (1/66[2%]); T69W (1/66[2%]); K70G (9/66[14%]); and V75L (1/66[2%]). V179E, which is a non-resistance mutation located at NNRTI resistance site, occurred in three (5%) of the 66 participants. The RT region in which no mutation confers any resistance to NRTIs and NNRTIs has been also shown to have undergone amino acid substitutions in 63 of the 66 participants. The most common (with prevalence rate of 20% or above) of these non-resistance mutations at non-resistance sites included: T20R (17/66[26%]); V35T (55/66[83%]); E36A (41/66[62%]); T39E (46/66[70%]); T39D (11/66[17%]); E40D (14/66[21%]); S48T (29/66[44%]); V60I (15/66[23%]); K173R (26/66[39%]); N175Y (51/66[77%]); G196N (56/66[85%]); E204Q (55/66[83%]); Q207V (47/66[71%]); and P243S (56/66)[85%]).

3.5.2.2 Diversity and evolution of genomic composition

The overall sequence diversity of all the three groups at the Pol region was determined based on pair-wise analysis of the sequences using default parameters in the Jukes-Cantor method using MEGA4 (Jukes and Cantor, 1969; Tamura *et al*., 2007). The analyses have shown that mean diversity calculations as the function of the number of base substitutions per site within the entire

subpopulations was 0.005. Analysis of pair-wise average sequence evolutionary divergence within each of the three groups of studies (within group comparison) has also shown that the number of base substitutions per site were 0.005, 0.004, and 0.006 from sequences of recently infected drug-naive, chronically infected drug-naïve, and chronically infected drug-experienced participants, respectively. Combining two groups together (recently infected drug-naïve and chronically infected drug-naïve into one as drug-naïve group) gave a within-group divergence rate of 0.004 base substitutions per site. Similarly, estimates of evolutionary divergence over sequence pairs between the three groups (between groups comparison) was determined as the number of base substitutions per site using the same software. The result of this analysis showed that average divergence between recently infected drug-naïve and chronically infected drug-naive was 0.004; while that between recently infected drug-naïve and chronically infected drug-experienced participants was 0.006. The same measurement for chronically infected drug-naive and chronically infected drug-experienced groups was 0.005. Splitting the whole sequence only into two groups (drug-naïve and drug-experienced) has produced a mean between-groups divergence rate of 0.005 base substitutions per site.

Sequence variability was also determined using the SNAP (Synonymous Non-Synonymous Analysis Program) software developed for calculating synonymous and non-synonymous substitution rates (Korber, 2000) available at the Los Alamos National Laboratory website (www.hiv.lanl.gov). Those sequences which were identified as non-subtype C in this study were omitted from the SNAP analyses as they had the potential to exaggerate the non-synonymous substitutions rate. Five groupings were made for use in the SNAP analyses: the three homogenous groups separately (recently infected drug-naïve, chronically infected drug-naïve, and chronically infected drug-experienced) and two more groups made by combining chronically

infected drug naïve and chronically infected drug-experienced (Chronic infection as opposed to recent infection) and recently infected drug-naïve and chronically infected drug-naïve (Drug-naïve group as opposed to drug-experienced).

Table 3.5.4 summarizes the overall mean synonymous and non-synonymous rate in the five groupings. The highest rates for both synonymous and non-synonymous substitutions was recorded from chronically infected drug-experienced study groups (0.134566, and 0.173064, respectively) followed by the combined group chronically infected (0.111358 and 0.119422, respectively) and recently infected drug-naïve (0.106792 and 0.076561, respectively). Regarding the mean non-synonymous to synonymous ratio, two groups had values >= 1: chronically infected drug-experienced (1.2861) and the combined chronically infected (1.0724).

The cumulative mean codon-by-codon ratio of synonymous to non-synonymous substitutions have shown striking patterns (Figure 3.5.14 Panel A-Panel E). In recently infected drug-naïve participants, non-synonymous substitutions (areas covered by heavy dashed blue lines in Figure 3.5.14 Panel A) were greater than synonymous substitutions at the first 100 codons, where the protease region is located, and at the region in the RT located after codons 220 in the alignment (around codon 112 in the RT protein). In the chronically infected drug-naive sequences, synonymous substitutions were greater right from the beginning in the protease region to around codon 260 in the alignment (roughly at codon 158 of the RT protein), after which the non-synonymous substitution took over (Figure 3.5.14 Panel B). However, in chronically infected drug-experienced sequences non-synonymous substitutions rate remained higher than those of synonymous substitutions starting from the beginning at the PR region to the end of RT region

(Figure 3.5.14 Panel C), although the extent was marginal until codon 120 in the alignment (around codon 18 of the RT protein).

Table 3.5.4 Overall mean synonymous (ds) and non-synonymous (dn) codon substitution rates and the non-synonymous to synonymous ratio (dn/ds) among various groups of the study.

Study groups	*Mean ds substitutions*	*Mean dn substitutions*	*dn/ds*
Recently infected drug naive	0.106792	0.076561	0.7169
Chronically infected drug-naive	0.093497	0.076387	0.08170
Chronically infected drug-experienced	0.134566	0.173064	1.2861
Total Drug-naive	0.098613	0.07711	0.7819
Total chronically infected	0.111358	0.119422	1.0724

In contrast, the groups composed of two combinations from the three homogenous ones have shown either nearly parallel substitution rate between synonymous and non-synonymous in the chronic infection group (Figure 3.5.14 Panel D) or gross domination of synonymous substitutions over non-synonymous ones in the drug-naïve group (Figure 3.5.14 Panel E). Two exceptions to these observations should be noted at codons following 260 in the alignment (around codon 170 of the RT protein) in the chronic group and the first 80 codons in both the alignment and the PR protein in the drug-naïve groups, where the non-synonymous substitution rates marginally surpassed the synonymous substitutions.

Shannon entropy test was also used to measure site-specific sequence diversity of one group (query) relative to another group (background). The test was performed on sequences of nucleotide and amino acid residues from the three homogenous groups and the two combined groups described. A total of 1037 nucleotide and 346 amino acid sites were compared from each group of sequences (Data not shown). The highest overall mean nucleotide entropy value from

the three homogenous sequence groups was recorded among the chronically infected drug-experienced (0.154695) followed by recently infected drug-naïve (0.134446) and chronically infected drug-naive (0.134018). Similarly, the highest overall mean amino acid entropy value was recorded in the chronically infected drug-experienced group (0.34868497) followed by chronically infected drug-naïve (0.25553468) and recently infected drug-naïve (0.230630058). Computation of overall mean entropy difference between the background (ancestral) and query group has shown that the query sequence groups had greater values than the background ancestral sequences for both nucleotide and amino acids, except for nucleotide sequence comparison between recently infected drug-naïve and chronically infected drug-naïve groups, where the background ancestral group (recently infected drug-naive) had slightly greater overall mean values (greater by 0.000426) than its counterpart in chronically infected drug-naïve group. However, none of these overall mean entropy differences were statistically significant ($P>0.05$ in all comparisons). Table 3.5.5 summarizes the overall mean entropy values among the different sequence groups.

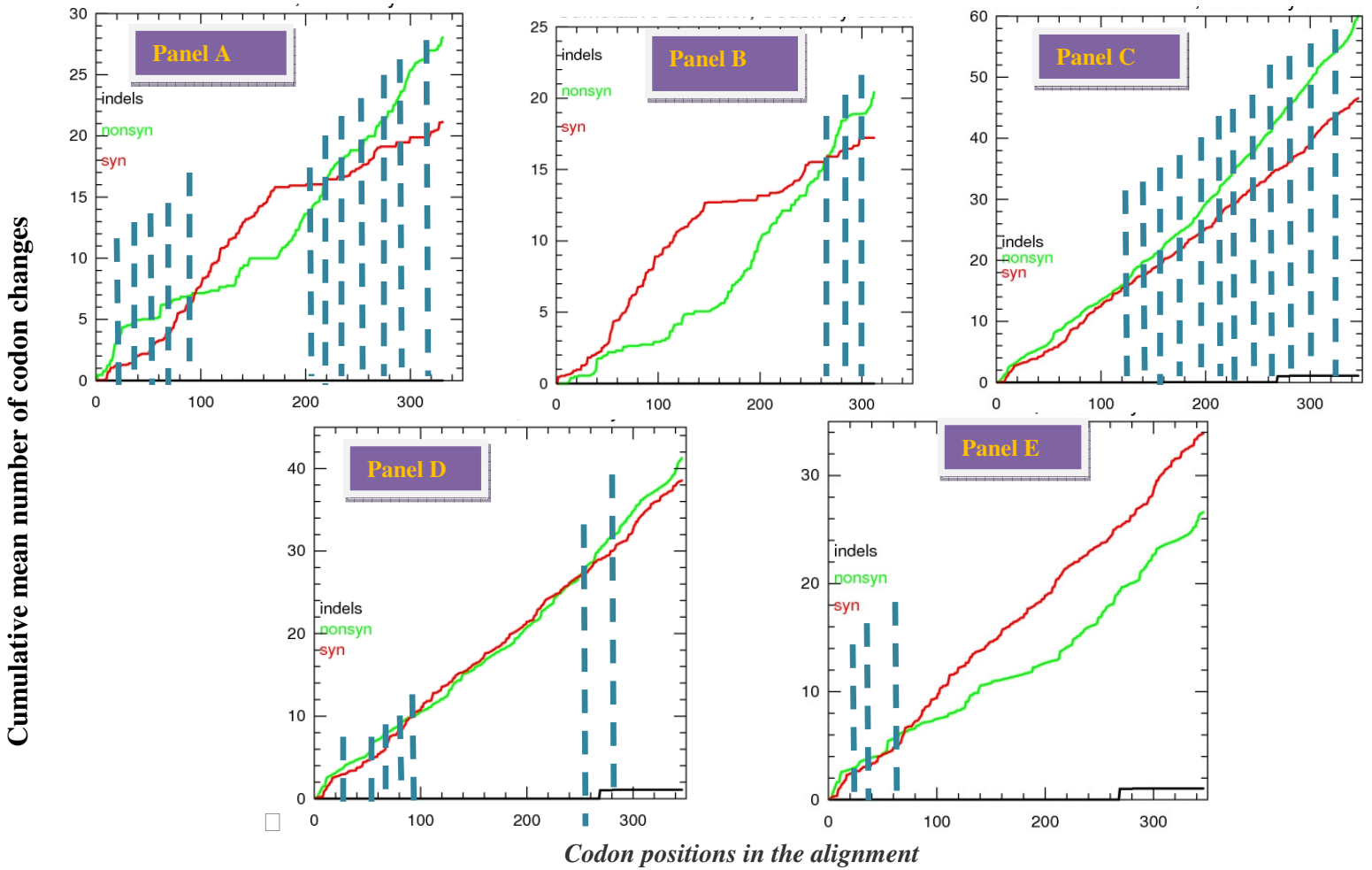


Figure 3.5.14 Cumulative dS/dN codon-by-codon behavior
Panel A, Recently infected drug-naïve; Panel B, Chronically infected drug-naïve; Panel C, Chronically infected drug-experienced; Panel D, Chronic infection (Chronic drug-naive and chronic drug-experienced infection); Panel E, Drug-naïve (recently and Chronically infected). Areas covered with heavy blue dashed lines indicate values in which the rate of non-synonymous substitutions exceed the rate of synonymous substitutions

Of the 1037 sites in the nucleotide sequences in the pairs of sequence groups under comparison, statistically significant greater entropy values than differences were shown in 44 sites between recent infection drug-naïve Vs chronic infection drug-naive (group A); in 69 sites between recent infection drug-naïve Vs chronic infection drug-experienced (group B); in 70 sites between chronic infection drug-naïve Vs chronic infection drug-experienced (group C); in 85 sites between all drug naïve Vs drug-experienced (group D); and in 64 sites between recent infection Vs all chronic infection (group E). In four of these five compared pairs of groups, the query groups (but not the background groups) had more number of sites with statistically significant greater entropy values than the differences: among group B, 44/69 for query vs 25/69 for background; among group C, 53/70 for query vs 17/70 for background; among group D, 69/85 query vs 16/85 background; and among group E, 38/64 for query vs 26/68 for background. The only exception was seen among group A, in which the background sequence group (not the query group) had marginally more number of sites with statistically significant greater entropy values than the differences (23/44 vs 21/44, respectively).

Likewise, comparisons of pairs of aligned sequences with 346 amino acid sites in ancestral (background) and query groups of sequences have identified sites with statistically significant greater entropy values than the difference: 26/346[7.51%] among group A; 43/346[12.43%] among group B, 42/346[12.19%] among group C; 50/346[14.45%] among group D; and 31/346[9.0%] among group E (Table 3.5.6). In all the five scenarios of comparisons, the query groups have produced overwhelmingly greater number of statistically significant entropy values, accounting between the lowest 63.38% among group A to the highest 92.0% among group D. In all the cases in which chronically infected drug-experienced group was incorporated as part of the query sequence group, the number of statistically significant higher entropy values exceeded that

of the background groups by at least four-fold. Splitting the amino acid sites under comparison into PR and RT regions has shown that more statistically significant higher entropy values were contributed by the RT region than the PR region with percentage proportions of 50%, 86.05%, 83.33%, 90.0%, and 74.19% among groups A, B, C, D, and E, respectively (computations not shown in the table).

Results of site-by-site measurement of entropy differences between that of the background and the query sequence sets are depicted in Figure 3.5.15 (Panels A1-E2). It is clearly visible from the figures that more diversity, as measured by more entropy values, was observed both at nucleotide and amino acid levels among the query sequence sets (represented by red lines below the scale 0) rather than among the background ones (represented by red lines above the scale 0). Moreover, except in the case of comparison A, statistically significant higher entropy values in the RT region of the query sequence sets were localized to certain adjacent nucleotide/amino acid sequence groups. These localized sites coincided with resistance mutation sites for NRTI and NNRTI. If the amino acid sequence sets alone is considered, one can see the first localized sites at about position 130 in the alignment, where the NRTI resistance mutation site M41 is situated; the second, around position 155 extending to about 170 in the alignment, where NRTI resistance mutation sites K65, K70, D67, L74, and V75 are situated; the third patch is around positions 200 in the alignment, where NNRTI resistance mutation sites A98, L100, K100, K103, V106, V108I, as well as NRTI resistance mutation site F116Y are situated. The fourth and the fifth groups are those immediately before positions 300 in the alignment, where NRTI resistance mutation site M184 and NNRTI resistance mutation sites V179, Y181I, Y188, and G190A; and those after positions 300 in the alignment, where NRTI resistance mutation sites L210, T215, K219, and NNRTI resistance mutation sites like P225 and M230 are concentrated.

Table 3.5.5 Overall mean entropy values of the three homogenous and two combined groups

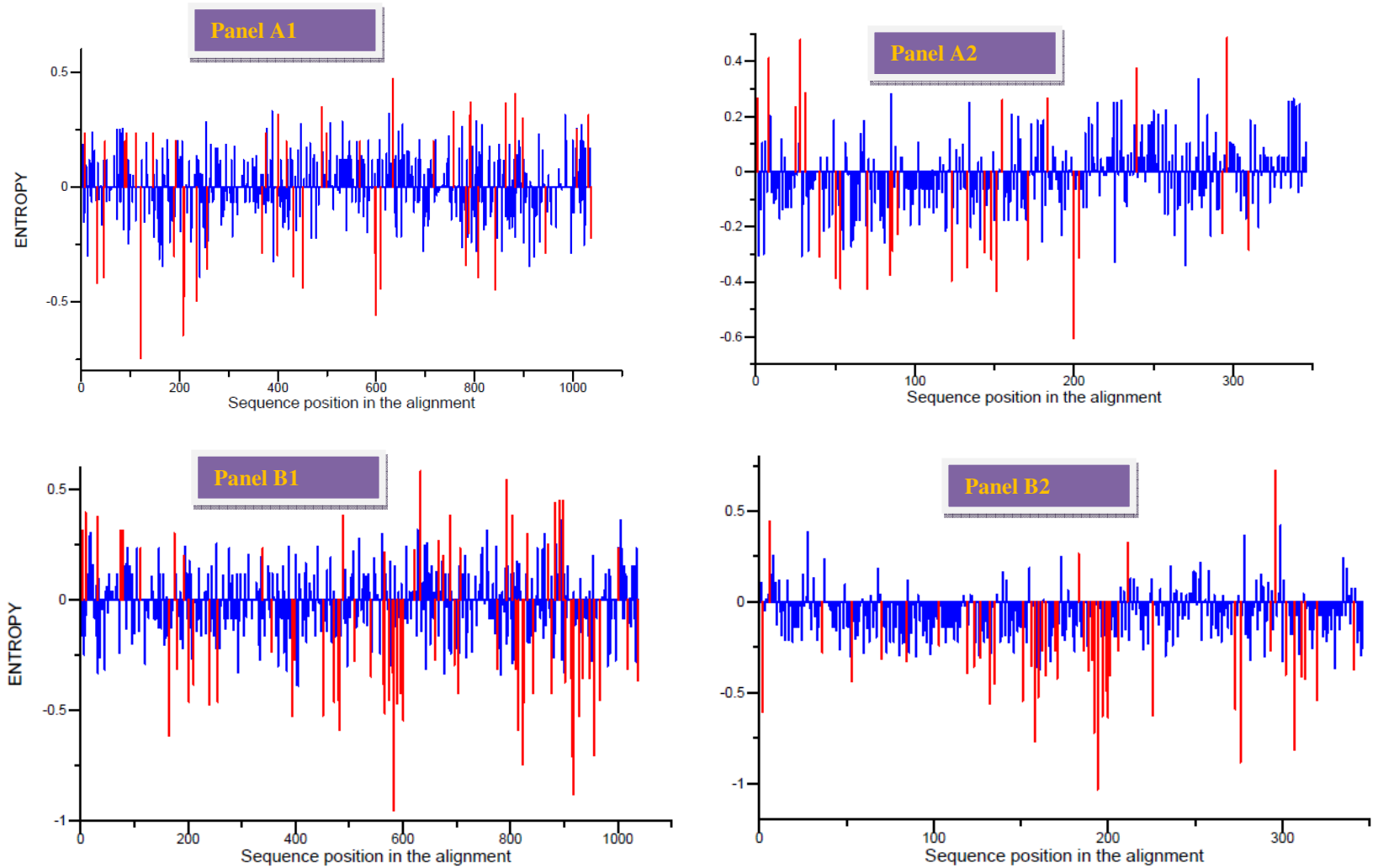
Background Vs Query	***Nucleotide***				***Amino acid***			
	Background entropy (Hb)	Query entropy (Hq)	(Hb-Hq)	P-value	Background entropy (Hb)	Query entropy (Hq)	(Hb-Hq)	P-value
A	0.134446	0.134018	0.000426	0.71703	0.230630058	0.25553468	-0.02493064	0.510723
B	0.13444605	0.154695	-0.020265896	0.686137	0.230630058	0.34868497	-0.11805202	0.430694
C	0.134018304	0.154694605	-0.0206927	0.678882	0.255534682	0.34868497	-0.09312717	0.409971
D	0.140634875	0.154694605	-0.014077071	0.66841851	0.258384393	0.34868497	-0.09026301	0.42315
E	0.13444605	0.15034104	-0.01589	0.66865	0.230630058	0.31109249	-0.08052601	0.452168

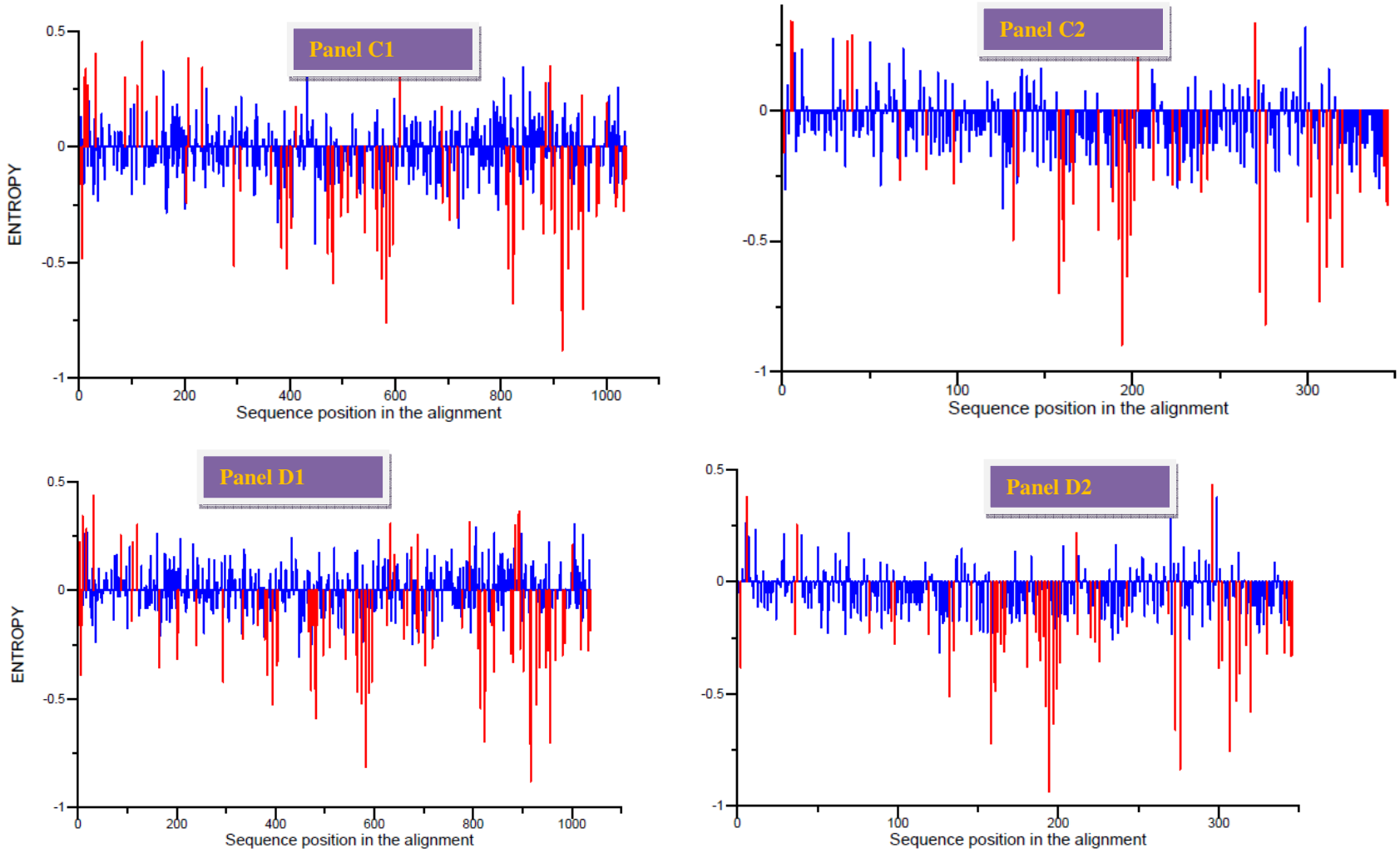
*****Key for groupings: **A,** Recently infected drug-naïve Vs Chronically infected drug-naïve; **B,** Recently infected drug-naïve Vs Chronically infected drug-Experienced; **C,** Chronically infected drug-naïve Vs Chronically infected drug-experienced; **D,** All Drug-naïve Vs experienced; **E**, Recently infected Vs chronically infected drug-experienced

Table 3.5.6 Summary of number of amino acid sites in homogenous and combined sequences with statistically significant greater entropy values than the difference between the query and background entropy values

Compared groups (Background Vs Query)*	*Region of Test site*	*Number (%) sites with statistically significant greater entropy values than the difference*		
		In the Query	In the Background	Total
A	Pol	17 (63.38%)	9 (34.62%)	26 (100%)
B	Pol	39 (90.70%)	4 (9.30%)	43 (100%)
C	Pol	36 (85.71%)	6 (14.29%)	42 (100%)
D	Pol	46 (92%)	4 (8%)	50 (100)
E	Pol	25 (80.65%)	6 (19.35%)	31 (100%)

*****Key for groupings: **A,** recently infected drug-naïve Vs chronically infected drug-naïve; **B,** recently infected drug-naïve Vs chronically infected drug-experienced; **C,** chronically infected drug-naïve Vs chronically infected drug-experienced; **D,** all drug-naïve Vs experienced; E, recently infected Vs chronically infected drug-experienced.


Panel A1
ENTROPY
0.5
0
-0.5
0
200
400
600
800
1000
Sequence position in the alignment
Panel A2
0.4
0.2
0
-0.2
-0.4
-0.6
0
100
200
300
Sequence position in the alignment
Panel B1
ENTROPY
0.5
0
-0.5
-1
0
200
400
600
800
1000
Sequence position in the alignment
Panel B2
0.5
0
-0.5
-1
0
100
200
300
Sequence position in the alignment


Panel C1
ENTROPY
0.5
0
-0.5
-1
0
200
400
600
800
1000
Sequence position in the alignment
Panel C2
0
-0.5
-1
0
100
200
300
Sequence position in the alignment
Panel D1
ENTROPY
0.5
0
-0.5
-1
0
200
400
600
800
1000
Sequence position in the alignment
Panel D2
0.5
0
-0.5
-1
0
100
200
300
Sequence position in the alignment

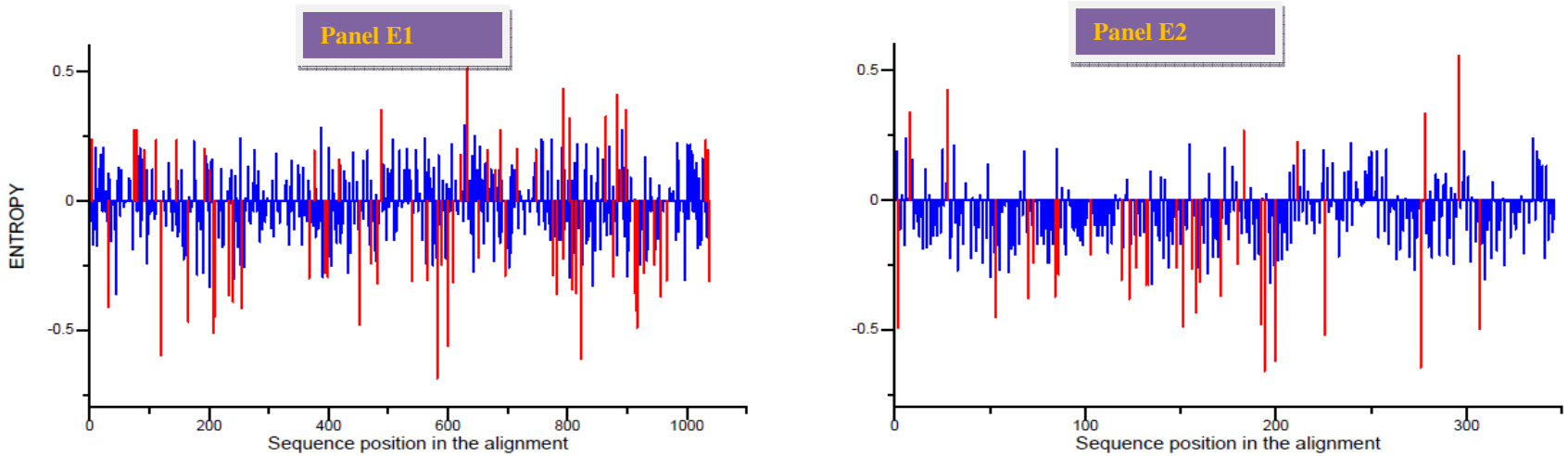


Figure 3.5.15 Site-by-site nucleotide and amino acid entropy differences at Pol region between background (ancestral) and query sequence groups. The blue lines indicate the sites at which entropy differences between the background (ancestral) and query sequences were not statistically significant, while the red lines indicate sites at which the differences were statistically significant. The values above the scale 0 were all contributed by the background entropy values, whereas the values below the scale 0 were those of the query sequences.

Key: Panels A1 and A2, entropy differences in nucleotide and amino acid sequences, respectively, between recently infected drug-naïve background and chronically infected drug-naïve query sets; Panels B1 and B2, entropy differences in nucleotide and amino acid sequences, respectively, between recently infected drug-naïve background and chronically infected drug experienced query sets of sequences; Panels C1 and C2, entropy differences in nucleotide and amino acid sequences, respectively, between chronically infected drug-naïve background and chronically infected drug experienced query sets of sequences; Panels D1 and D2, entropy differences in nucleotide and amino acid sequences, respectively, between all drug-naïve (recent and chronic infection) background and drug-experienced query sequence sets; Panels E1 and E2, entropy differences in nucleotide and amino acid sequences, respectively, between recent infection Background and all chronic infection query sequence sets.

CAHPTER 4: DISCUSSIONS

4.1 Discussions on demographic features and clinical laboratory measurements

4.1.1 Recently infected participants

The overall prevalence of HIV infection among the source population considered in this study, i.e., pregnant women enrolled in the 2005 round of ANC-based sentinel, was 13.7%. This figure was greater by 2% than the 11.7% reported by HAPCO (MOH/HAPCO, 2006). Two reasons might attribute for this discrepancy: first, analysis in this study has included data from Federal Police Hospital, which the HAPCO report treated separately; and second, the HAPCO report included only three sites in its analysis of overall prevalence in Addis Ababa (MOH/HAPCO, 2006, page 37), whereas in this study all the six sites with acceptable specimen handling and unambiguous data entry were included. HIV prevalence at specific sampling sites in this study has also revealed that specimens collected from Federal Police Hospital had the highest estimate (over 20%) followed by Kazanchis Clinic (over 19%), Gulele Clinic (over 14%), T/Haimnot and Kefitenga 23 Clinics (each with over 10%), and Akaki Clinic (about 9%), in that order. Except for the estimates for Federal Police Hospital, estimate in this study were slightly greater than estimates in the HAPCO report (Figure 4.1.1).

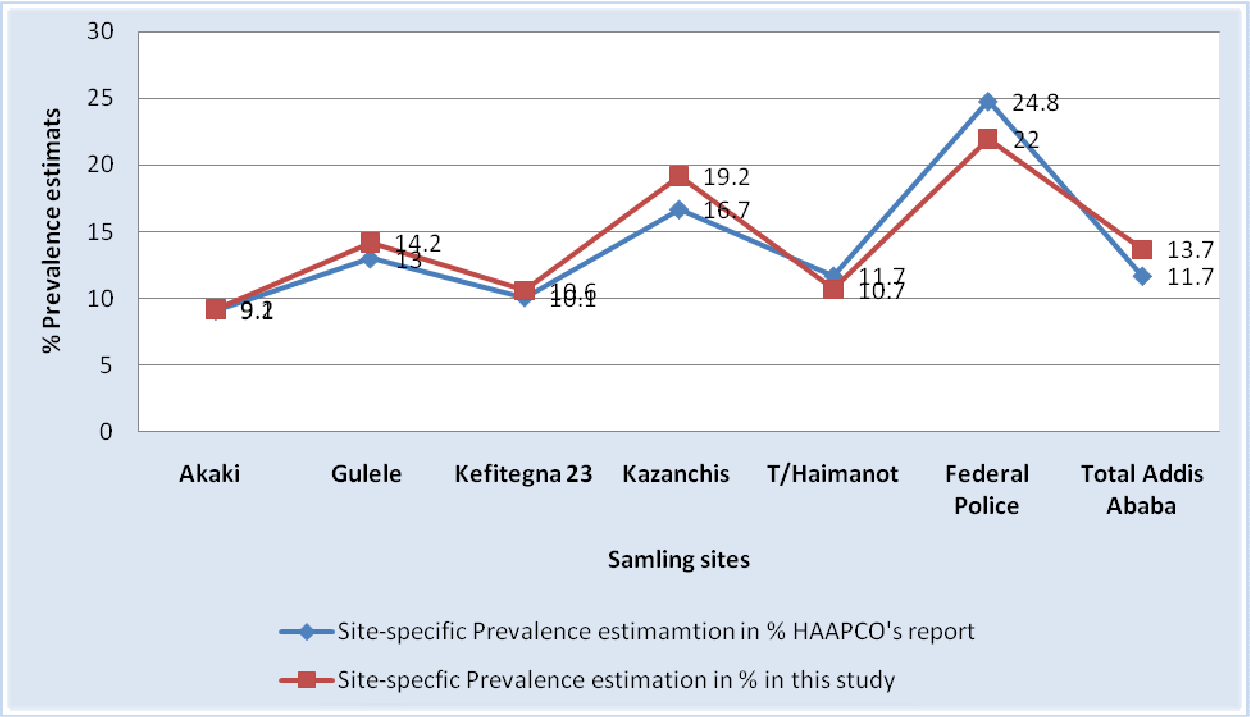


Figure 4.1.1 HIV prevalence estimates in Addis Ababa from the 2005 ANC-based sentinel sites as reported by HAPCO (MOH/HAPCO, 2006) and in this study.

Although 300 HIV-positive samples were collected from the six sampling sites, 11 of them had incomplete data, reducing the number of usable samples to 289. However, only 77 samples satisfied the criteria set by WHO (Bertagnolio and Sutherland, 2005; Myatt and Bennett, 2008) in its threshold survey schemes to investigate emergence of transmitted HIV drug resistance varieties. The fact that only 63/77 (81.82%) eligible samples had detectable viral RNA showed that either there were poor sample storage or sample processing practices at sample collection sites, or use of sera samples instead of plasma samples might have resulted in this reduced viral RNA detection. Although sera samples in general have the tendency to reduce the rate of viral RNA detection by at about 21% (Rodriguez, *et al*., 1997), this could not have been the cause of reduced detection in this study since all of the study participants were acutely infected. Under this setting, it is unlikely for RNA level to be undetectable, regardless of the sample type. It seems,

therefore, that poor sample storage/handling practices contributed for poor RNA detection and genotyping performance recoded in this study. In fact, there were differences among the samples collected from the six sites regarding success of RNA detection and genotyping: samples from Gulele and Akaki clinics had the highest RNA detection rate (over 93%) and genotyping success rate (over 72%), respectively. On the contrary, samples from Kazanchis clinic had the lowest RNA detection rate (about 67%) followed by samples from Kefitegna 23 (about 77%), whereas those from Kefitegna 23 clinic had the lowest genotyping success rate (31%) followed by those from Kazanchis (34%), showing poor specimen handling at the latter two sites. Fortunately though, the number of successfully genotyped samples from all the sites (41/63 [67.08%]) was sufficient to proceed with the survey, since it met the sequential resistance testing criteria set in the WHO Guideline.

4.1.2 Chronically infected drug-naïve participants

This group consisted of chronically HIV infected persons who were eligible and ready to start on ARV treatment. The majority of them (54.1%) were men, and over three-fourth of the study participants in this group were 40 or below years of age, which is economically productive age group. This figure was greater than that documented among new ART attendees in Southern India, where about 51% of the attendees were between ages 30-40 years (Badiger *et al*., 2010), and Malawi, where about 55% of the new ART attendees were below the age of 40 (van Oosterhout *et al*., 2005). In addition, over 74% of the participants belonged to WHO clinical stages III and IV. This high percentage is not unexpected since the health situation of all of the participants was deteriorating to the extent of needing initiation of ARV therapy. However, the proportion in this study was much greater than the proportion documented in Malawi, where only 44% of patients newly starting ART belonged to WHO stages III and IV (van Oosterhout *et al*., 2005).

The median and mean plasma viral RNA were well over 200,000 copies/ml, with peak 2.5 million copies/ml and nearly 70% of the participants having over 100,000 copies/ml. This high baseline viral load, particularly those with ≥100,000 copies/ml, has been demonstrated in several studies to be associated with poorer ARV treatment responses (Hirsch, 2005; Holmberg *et al.,* 2003; Langford *et al*., 2007). Similarly, the CD4+ T cell count was unfavorable for the majority of the participants, with median and mean values of 103.50 and 105.95 cells/µL, respectively. About 92% and 23% of the participants had values ≤200 and ≤50 cells/µL, respectively. Like the case in the plasma viral RNA level, the very low baseline CD4+ cell level documented in this study could be an indication of poor prognosis for disease progression, survival, and treatment response, as has been shown in other similar studies (Hirsch, 2005; Holmberg *et al*., 2003; Langford *et al*., 2007). Considering the facts that most of the participants' WHO stage (III and IV), very high baseline plasma viral RNA level (mean value of >420,000 copies/ml), and low baseline CD4+ T cell level (mean value of 105.95 cells/ µL), it can be concluded that ARV treatment in these patients started late, and that favorable treatment outcome would not be sustained for longer period, as high nadir viral load (Langford *et al*., 2007; Loannidis *et al*., 2000) and low CD4+ T cell count (May *et al*., 2010; Monforte *et al*., 2005) have been shown to negatively affect later virological and immunological responses.

Although the highest viral load was recorded among women, comparison of mean viral loads between those of women and men showed that plasma viral RNA in men (487,173 copies/ml) was higher than that observed in women (343,512 copies/ml), though the value was not statistically significant (p= 0.215). Same unfavorable condition was observed more on men than women with regard to baseline CD4+ T cell level: mean CD4+T cell count in men (90.4 cells/µL)

was significantly lower than that observed in women (124.97 cells/μL) (p= 0.017). Such higher mean viral load and lower CD4+ T cell count in men than women was also documented in other studies elsewhere (Kiwanuka *et al*., 2010; Langford *et al*., 2007).

4.1.3. Chronically infected drug-experienced participants

This group of participants consisted of ART treated men (over 53%) and women (over 45%), with overall mean and median treatment duration of 42.09 months (N=262; SD=17.339) and 40.02 months (Range= 112 months), respectively. The mean and median duration of treatment for men and women were 42.90 (Std. Error= 1.395) and 44.0 (Std. Error= 1.736) months for men and 40.93 and 38.0 months for women, respectively. The difference in treatment duration between the two genders was not statistically significant (p= 0.137). These data clearly demonstrated that this group of patients was heavily treated with ARV drugs. Similarly, the overall mean and median duration of sero-conversion were 55.38 months (SD= 25.175 months) and 52.0 months (R=154 months), respectively, showing the degree of chronicity of the infection in this group. There was no statistically significant differences (p= 0.330) in the mean duration of sero-conversion between men (56.79 [Std Error= 2.020 months]) and women (53.72 [Std. Error= 2.445 months). With regard to age, the overall mean age of these participants was 37.76 years (N= 255; SD=9.822). However, men's mean age (40.83 years) was significantly greater than that of women's (34.15 years) (p= 0.000).

WHO clinical stage profile of the participants has shown that over 84% of study participants in this group belonged to the more advanced baseline WHO stages III and IV. There was no correlation between gender and WHO stages (Spearman's rho= 0.095, p= 0.132). Same proportion (85%) of the two advanced WHO stages was observed in Africa from a study comprising patients who were under ARV therapy in four scale-up cohorts (May *et al*., 2010).

However, baseline WHO clinical stage profile of the chronically infected treatment starting study participants in this study showed reduced prevalence (74%) in stages III and IV compared to the prevalence in this group (84%). The high prevalence of people in stages III and IV in this group of participants could be because of people's ignorance of the availability of and/or reluctance to seek for ARV treatment services in the earlier periods of ART scaling-up.

The laboratory test data analyses have shown that mean current plasma RNA level, baseline CD4+ T cell count, and current CD4+ T cell count were, respectively: 289,128.31 copies/ml (Std. Error= 90946.16), 114.41 cells/µL (Std. Error= 6.22), 237.77 cells/µL (Std. Error= 10.17). Point of concern here is that compared to patients who start ART in other places (Frater *et al*., 2002; Trotta *et al*., 2010), participants in this study (both chronically infected drug-naïve and drug-experienced) had considerably lower mean baseline CD4+ T cell counts, showing that the patients' virological and immunological situations have already deteriorated when they sought for treatment. While baseline CD4+ T cell count was found to show significant association with subsequent immunological responses (p= 0.002) its association with current virological response (current plasma viral RNA) was not statistically significant (p= 0.370). This direct association between baseline CD4+ count and subsequent immunological response was also observed in a 7-year follow-up study by Lok *et al*. (2010), in which patients who began ART at lower CD4 cell counts remained to have lower CD4 cell counts (<200 cells/µL) compared to those who began with higher CD4 counts. Similar result was documented by Lawn *et al*. (2006), who observed that patients who began ART at baseline CD4+ T cell count of <50 cells/µL remained at <200 cells/µL for longer period, although they were able to achieve equivalent or greater immunological recovery within the first 48 weeks of ART.

Of the 266 participants in this group, 83 (31.2%) had virological failure with >150 copies/ml. This viremic rate is greater than the one reported in South Africa by over a fold, where only 14% of the patients under treatment for over 12 months were reported to have detectable viral load (El-Khatib *et al*., 2010). However, it is less than observed in the United States, which was reported to be 63% (Richman *et al*., 2004). Fifty-one of these 83 (61%) viremic participants or 19% of all participants in this study group fell on WHO's treatment failure category threshold of >=5,000 (WHO, 2010). Mean plasma viral RNA level of men (345,816 copies/ml [Std. Error= 73921 copies/ml]) was greater than observed in women (183,565 copies/ml [Std. Error= 135739 copies/ml]). However, the difference was not statistically significant (p=0.404).

The importance of viral load value in predicting clinical benefits under varied CD4+ T cell counts was emphasized by several published results. For example, it was reported that at CD4+ counts 200–350 cells/μL patients with viral load of ≥300,000 copies/ml had a four-fold risk of developing AIDS compared to those with viral load of 3000 copies/ml; and this risk was reported to increase further with viral load of >100,000 copies/ml under any CD4+ count (Langford *et al*., 2007). In a related development, it has been established that there is a gender difference in regards to viral load at a given CD4+ T cell count. At lower CD4+ counts (e.g. ≤50 cells/μL), women have greater mean viral load than men; but at higher CD4+ counts (e.g. >350 cells/μL) women show lower mean viral load than men (Langford *et al*., 2007).

Quite surprisingly, however, the observation in the current study group was contrary. At both lower CD4+ count strata (≤50 cells/μL, 51-101 cells/μL, and 102-201 cells/μL) and higher CD4+ count strata (202-301 cells/μL and 302-501 cells/μL), the number of women with higher viral load groups (i.e, ≥10,001-50,000 copies/ml) was lower than that of men (Figure 4.1.2). The

reason for this unexpected result is not clear, but it could not be because Ethiopians are different from the other population in this regard, since participants in the chronically infected drug-naive study group (part of this study in the previous section) followed the same pattern of more women than men with higher viral load at low CD4+ count and lower viral load at higher CD4+ count (see Figure 3.2.1). Rather, non-biological reasons such as poor adherence, alcohol and drug consumptions like *Chat*, might have contributed for poorer virological response in men, as men are prone to be involved in these kinds of activities more than women.

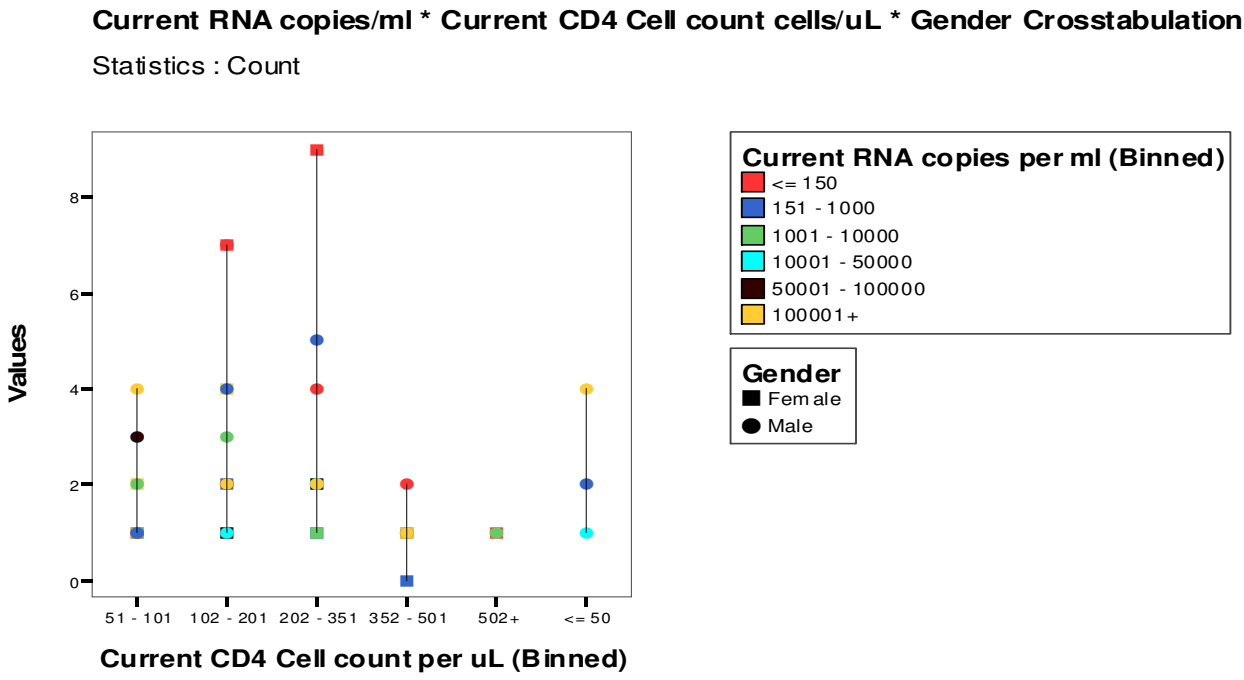


Figure 4.1.2 Comparison of plasma viral load distribution between chronically infected drug-experienced women and men study participants at various stratified CD4+ T cell counts

Another important observation in this study is that virological responses of the patients to the various ART drugs varied from regimen to regimen. Plasma viral RNA detection rate for each initial 1st line ART drug regimen has shown the following: 26/70 (37%) for 1a (D4T-3TC-NVP), 14/28 (50%) for 1b (D4T-3TC-EFV), 20/90 (22%) for 1c (AZT-3TC-EFV), and 22/70 (31%) for 1d (AZT-3TC-NVP). However, when WHO's virological failure criterion was considered, as

measured by the detection of mean plasma viral RNA ≥5,000 copies/ml, virological failure rates among the four 1st line ART regimens used during initiation of the treatment were 19/70 (27%) for 1a (D4T-3TC-NVP), 6/28 (21%) for 1b (D4T-3TC-EFV), 10/90 (11%) for 1c (AZT-3TC-EFV), and 15/70 (24%) for 1d (AZT-3TC-NVP). Therefore, it seems that D4T-containing regimens (25/98 [~26%]) were less effective than AZT-containing regimens (25/160 [~16%]) in bringing favorable virological response. Of the 55 patients for whom regimen switching was done, 16 were switched to 2nd line while the rest 39 changed to a different 1st line. Among the newly introduced 1st line regimens, plasma viral detection was still observed from 3/9 (33%) of 3TC-TDF-NVP, 5/28 (18%) of 3TC-TDF-EFV, 1/28 (4%) of 3TC-TDF-EFV. But strict virological failure according to WHO's criterion occurred only in 1/14 (11%) of 3TC-TDF-NVP. Of the two 2nd line regimens, plasma viral detection was observed among patients taking 3TC-TDF-kaletra (5/14 [36%]), with mean viral load of 21,176 copies/ml, and ABC-TDF-kaletra (1/2 [50%]), with mean viral load 639 copies/ml. However, only in 1/14 (7%) of the regimen 3TC-TDF-kaletra did virological failure occur with plasma viral load over 5,000 copies/ml.

4.2 Discussions on genotypic results

4.2.1 Performance of in-house genotyping assay

The commercial assays used to detect HIV drug resistance, ViroSeq™ HIV-1 Genotyping and TruGene®, were developed and optimized for subtype B, the variant most prevalent in resource-rich countries. Although the two assays have been claimed to genotype non-B subtypes as well as the B subtype (Eshleman *et al*., 2005b; Jagodzinski *et al*., 2003), different comparative studies have shown inconsistencies and inadequacy of the two assays to successfully genotype non-B subtypes including recombinant forms (Beddows *et al*., 2003; Maes *et al*., 2004). This problem prompted many investigators to develop and use In-house-brewed genotyping systems including primers

which could anneal to highly conserved regions in the Pol and amplify diverse subtypes (Snoek *et al*., 2005). The In-house genotyping assay validated in this study under the setting of Ethiopian HIV-1 isolates was developed in an attempt to address this problem and was optimized for the purpose of genotyping using DBS samples (McNulty *et al*., 2007).

The performance of the RT-PCR primer sets in this assay in amplifying the Ethiopian isolates was 100%, although weakly amplified products could sometimes be observed during gel detection. Its genotyping performance relative to the commercial ViroSeq™ assay was also 100%. Comparison of the sequences produced from the In-house assay with sequence products of ViroSeq™ commercial assay has shown that at the nucleotide level it had mean concordance of 98.72% and sensitivity of 94.26% and positive predictive value of 99.47% at the amino acid level in the Pol region. The small discordance observed in these comparisons could largely be attributed to operator's interpretation discrepancies during manual editing, as this process sometimes requires subjective decisions.

One of the important features of this assay was that it was capable of amplifying samples having low viral load (as low as 500 copies/ml), although originally designed to amplify samples with viral load of 1000 copies/ml. Although not many drug-resistance mutations were detected because the samples used in this comparison were drawn from drug-naïve patients, use of the assay in the other part of this study, that involved samples from drug-experienced patients, has demonstrated that drug-resistance mutations at the protease and reverse transcriptase sites were well detected. In general, the performances of the In-house assay relative to the ViroSeq™ system in amplifying and genotyping HIV-1 Ethiopian isolates were more than satisfactory and hence the in-house system should preferably be used since it is less expensive and technically

less demanding. For instance, using in-house system it is possible to work with RNA samples extracted with any RNA extraction method, unlike the ViruSeq™ system, which produces poor amplification, if any, unless the RNA extraction is done with the system's own extraction method.

Owing to its capability of amplifying and genotyping fragment lengths of only about 1023 base pair (McNulty *et al*., 2007), one drawback of this genotyping assay is its inability to produce longer range fragment from the *pol* compared even to ViroSeq™, which amplifies about 1800 base pair fragments (Eshleman *et al*., 2005b). This calls for further optimization of the assay to increase the length of the region to be amplified and sequenced. First of all, reports are emerging that mutations at positions different from the commonly recognized drug resistance sites at the protease and reverse transcriptase regions can confer resistance to the three globally used drug classes. For example, certain mutations at *gag* cleavage sites such as A431V or I437V in NC-SP2-p6 (Dam *et al.,* 2009), and NC/p1 cleavage site (K436A and I437T) (Nijhuis *et al.,* 2005) can directly confer protease resistance in addition to compensating for fitness losses. It has also been shown repeatedly that mutations at the connection and RNase H domains of reverse transcriptase, which are situated out of the range of commonly genotyped RT region (the polymerase domain), can render the mutant variant resistant to NRTI (Nikolenko *et al*., 2005; Yap *et al.,* 2007). Moreover, the enzyme integrase, the gene of which is situated in continuum with the RT region, is part of the polyprotein Pol and the target for the drug class integrase inhibitor. Detection of drug resistance mutations for this target will be integral part of treatment monitoring in the future when drugs against the enzyme are introduced. Since the in-house assay tested in this study cannot amplify and genotype this region, it would be of no help when integrase inhibitors are introduced in the country.

It is therefore of great benefit, while using the currently validated in-house assay, if other in-house drug resistance genotyping systems are developed or/and optimized, or the current one improved to cover all the *gag* and *pol* regions that are relevant for the detection of drug resistance conferring mutations using only a single amplification and sequencing protocol. Actually, an in-house genotyping assay already existed that was developed and optimized to detect mutations from the whole protease and the entire reverse transcriptase (RT) gene, including the gag cleavage site region and RNase H region (Snoeck *et al*., 2005). Unlike the genotyping assay validated in the current study, which uses outer and inner primer pairs that can produce a 1023 bp long amplicon and six sequencing primers that amplify 15-99 residues of the protease and 1-256 residues of the reverse transcriptase, the assay developed and optimized by Snoeck *et al*. (2005) comprises of outer and inner pairs of primers that can produce 2878 bp and 2853 bp amplicons respectively, and 10 sequencing primers that are capable of sequencing right from the 3'-end of p24 to the middle of integrase region. The advantage of this assay is that it has already been tested in all major subtypes and some CRFs of group M viruses (Snoeck *et al*., 2005). The optimization it needs for adopting to our setting would be to extend the region in the 3'-end to encompass the whole of integrase, and improve its sensitivity for subtype C and CRF02_AG to amplify, just like the other subtypes tested, at lower RNA copy number than the 5000 copies/ml it has demonstrated for the two subtypes.

4.2.2 Performance of DBS for detection of drug resistance mutations

After having checked the usability of the in-house assay, it was then utilized for genotyping HIV-1 isolates from DBS and plasma samples of drug-naïve patients. RNA extraction and amplification of the target Pol region was successful in every DBS samples tested for which plasma counterpart was available. The 100% amplification success could be due to the high viral

load in all the samples; the minimum RNA copy in the tested DBS samples was 15,000 copies/ml. However, given success stories of amplification in other studies with similar sample storage conditions but with viral RNA of $\leq$ 2000 copies/ml (Masciotra *et al*., 2007), $\geq$500 copies/ml (Johannessen *et al*., 2009), and even from samples with undetectable viral load (Buckton *et al*., 2008), it can be assumed that the assay would amplify if the samples in this study had lower viral loads than the one observed. Yet, further work is required at lower viral load to provide definitive conclusion, since published results from some other studies, which evaluated the suitability of DBS to quantify plasma RNA, have reported that the highest viral load value below which viral RNA could not be detected was 2084 copies/ml (Lofgren *et al*., 2009). One more study conducted under the African settings also reported that the minimum viral loads that could be detected from DBS samples in South Africa and Burkinafaso were 3100 and 1550 copies/ml, respectively (Viljoen *et al*., 2010), showing the need to evaluate performance of DBS at low viral load under the Ethiopian setting.

Similar to the success of amplification, genotyping was also successful for all DBS samples attempted, although the quality of some of the sequences was lower (as measured by the intensity of noises in the chromatogram and weak peaks, compared to others but tolerable since manual editing could fix the problems. This 100% genotyping success was remarkable when compared to success rate reported from Cameroon (92.5%) (McNulty *et al*., 2007) for DBS samples stored under comparable storage temperatures (-20^{0}C in this study but -20^{0}C and -70^{0}C in the Cameroonian study) and similar in-house genotyping assay. However, the length of storage time was different in the two studies; it was only 2.5-4 years in this study but 6-7 years in the Cameroonian study. Other studies which used in-house genotyping system also reported over 90% success (Johannessen *et al*., 2009; Youngpairoj *et al*., 2008). However, success rate in some

studies which used commercial genotyping assays was very much reduced; 57.5% using ViroSeq assay (Youngpairoj *et al*., 2008) and 78.8% using Trugene® (Hallack *et al*., 2008).

Concordance rate of genotyping between the DBS samples and their plasma counterparts in the current study was also excellent. With nucleotide mean concordance rate of 98.84%, 98.27%, and 98.82%, for PR, RT, and Pol respectively, it must be among the most successful DBS evaluation attempts done. Over 89% (31/35) of the DBS sequences had ≥98% concordance with sequences of their plasma counterparts at the nucleotide level. Only about 11% (4/35) had concordance between 96.34% - 97.9%. Similar scale of concordant genotyping result between DBS and plasma samples were reported from the study of Hallack *et al*. (2008), who obtained 99.3% using Trugene®; McNulty *et al*. (2007), 98.1% using the same genotyping assay as in the current study; Bertagnolio *et al*. (2007), with median concordance of 99.95% using a different in-house genotyping system; and Masciotra *et al*. (2007), with mean 96.8% concordance using ViroSeq assay.

Moreover, the sensitivity in this study of DBS genotyping relative to plasma genotyping at the amino acid level was 97.08%, 96.63%, and 96.78%, for PR, RT, and Pol respectively. Hence, the power of DBS genotyping to correctly detect the presence of a given mutation (positive predictive value) was high: 98.83%, for PR, 98.04%, RT, and 98.57% for Pol. Taken together, it can be concluded that HIV-1 drug resistance genotyping assay evaluated on DBS samples in this and other similar studies have given ample evidences that DBS is appropriate for use in resource –limited countries like Ethiopia, since it is of low cost, low demand for stringent storage conditions (provided it is transported to the genotypic laboratory promptly; or stored at 4^0C ordinary refrigerator temperatures for a maximum of 1 year; or as long as 5-6 years if stored at -

20^0C). Although excellent genotyping results could be produced from DBS regardless of the genotyping assay used, it is recommended that in-house assays optimized under the existing setting be used given the superior positive rate of amplification in the in-house genotyping assays compared to the commercial assays.

4.2.3 Discussions on drug resistance mutations

4.2.3.1 Prevalence rate of HIV-1 transmitted drug resistance mutations

The life expectancy and the quality of life of people infected with HIV have changed dramatically since the advent of HAART (Little *et al*., 2002). However, this success was accompanied with emergence and transmission of drug resistance virus variants among the treated and newly infected persons, respectively. This problem has been restricted to the industrialized countries, where HAART has been in use for longer time, until recently (Buckton, 2008). With recent introduction and scaling-up of HAART in resource-constraint countries, the threat of emergence and transmission of HIV drug resistance has become a global concern (Bennett, 2006). In Ethiopia, where ART scale-up program began since 2005, and where over one million people are living with HIV/AIDS (UNAIDS, 2010) the magnitude of emergence and transmitted drug resistance has not been known. This study has investigated the issue from three groups of patients: recently HIV-1 infected ARV drug-naïve pregnant women, chronically infected supposedly ARV drug-naïve but ready to start treatment, and chronically infected and heavily treated patients.

There were a total of 129 drug-naïve participants for whom genotypic drug resistance mutations were investigated. Forty-one of these were recently infected, whereas the remaining 85 were chronically infected. Only the former group of participants met WHO's criteria for consideration

in transmitted drug resistance surveillance scheme, because they are considered least likely to be previously exposed to ART and in whom therapy-driven drug resistance was unlikely (Myatt and Bennett, 2008). Therefore this section of the study reports the prevalence of HIV-1 transmitted drug resistance from genotyping data generated from the recently infected participants only. Drug resistance mutation profile from this group of participants showed that most of the mutations observed were polymorphic, according to the WHO's list of transmitted drug resistance mutations (Bennett *et al*., 2009). Only two PI resistance mutations (G73A and N88D) and no major NRTI or NNRTI mutations were detected in this group, thus making transmitted drug resistance rate in Addis Ababa about 4.9%, which falls under the category "low prevalence".

However, extrapolating this to the whole country is difficult as drug exposure histories prior to the 2005 scale-up could have varied from region to region. Thus, similar surveillance studies covering major cities and towns of the country should be conducted in the future in order to determine prevalence of transmitted drug resistance in Ethiopia. Even the rate obtained in this study might not show the true prevalence of transmitted resistance in Addis Ababa for several reasons. For one thing, the data were generated from population sequencing, which detects only sequences from viral population represented by over 20% of the total population (Metzner *et al*., 2005; Palmer *et al*., 2005). If in this study variants harboring resistance mutations were represented in <20% of the quaispecies, the 4.9% low category might have provided a wrong sense of safety, since drug resistant variants present even in samples with undetectable viral load and samples in which drug resistance mutations were not detected by conventional population sequencing were found by deep sequencing to have harbored drug resistance mutations that could reduce the potency of the treatment (Bergroth *et al*., 2009; Buckton *et al*., 2010; Halvas *et al*., 2010; Johanson *et al*., 2008b; Metzner *et al*., 2005).

Secondly, the samples for which genotyping was done were drawn from a single anatomical site, that is serum. Drawing conclusions from analysis conducted only on serum samples could lead into erroneous impression that reduced transmitted drug resistance prevalence exists in Addis Ababa, while viruses from different compartments (breast milk, genital tract, semen, CSF, PBMC, etc) could contain genetically distinct variants including those harboring drug resistance mutations that might serve as reservoir for later dissemination (Becquart *et al*., 2002; Bergroth *et al*., 2009; Tirado *et al*., 2004; Pilger *et al*., 2010).

Thirdly, absence of drug resistance mutation in this study could be due to reversion of previously transmitted drug resistance into either wild type, because of reduced fitness in the absence of drug selective pressure, or mutations that are not recognized by the current algorithms as resistance. Progressive reversion (or disappearance) of some drug resistance mutations (such as Y181C and K219Q of NRTI resistance mutations) into wild type or nonresistance mutations in the absence of drug selective pressure has been documented in several studies, although some others (such as M41L, T69N, K103N, and T215Y/F) could stay stable for longer time (Gandhi *et al*., 2003; Pao *et al*., 2004; Turner *et al*., 2004). In this regard, two more mutations (V82I of Protease in five cases, and another K65E of reverse transcriptase) occurred at their respective resistance sites in the current study. It can therefore be assumed that these two mutations could be revertants from the previous resistance mutations (V82I from one of the following V82A/T/F/S/C/M/L; and K65E from K65R). For example, reversion of the resistance mutation V82T into non-resistance mutation V82I requires only replacing the second nucleotide ‘C’ in the degenerate codon ‘ACU’ or ‘ACC’ or ‘ACG’ of threonine with a ‘U’ to produce ‘AUU’, ‘AUC’, or ‘AUG’, respectively, of Isoleucine. If these six additional mutations (five cases of V82T and

one case of K65E) were products of reversion from previously acquired drug resistance mutations, the rate of transmitted drug resistance would have been computed to 8/41 (or nearly 20%), since they occurred in different individuals.

Another important issue is that the 4.9% resistance prevalence rate in Addis Ababa may not hold true at this time. Samples for this study were taken in 2005 when very few people had access to ARV treatment, since official ART scale-up began only during that year; thus little, if any, treatment-driven drug resistance mutations could have emerged that would be transmitted to newly infected individuals. Given that ART is widely used in this country to treat more patients now than before, the six years elapsed time between now and the first sampling time is long enough to change the transmission pattern. In fact this kind of observation was documented from rural Malawi, where in just three years after commencement of ART scale-up, five out of 40 drug-naïve newly diagnosed patients showed NRTI (12.5%) and four NNRTI (10%) resistance mutations with one individual showing both (2.5%) (Bansode *et al*., 2010).

Several more investigations have already reported that when HIV-infected people gained back their health upon ARV treatment, they were engaged in risky sexual activities (Elford *et al.,* 2002; Shafer *et al*., 2011), a disturbing phenomenon that would facilitate transmission of treatment-driven drug resistance mutations to newly infected sex partners. If these risky sexual behaviors existed in Ethiopia for the last six years (there is no reason why it would not), then there is a very good chance that transmitted resistance will have already increased substantially, since virological failure among heavily treated patients has already reached 31.2%, and treatment-driven resistance for NRTI, NNRTI and double class drug resistance mutations reached over 65%, 70% and 60% respectively (reported in this study elsewhere in the section addressing

drug resistance mutations in drug-experienced patients). Thus, all these shortcomings should be taken into consideration while interpreting the currently determined 4.9% prevalence rate of transmitted drug resistance from Addis Ababa. This, therefore, calls for undertaking a new transmitted surveillance study, which should enlarge its scope both in the number of regional sites, inclusion of private health institutions and diversity of study populations.

4.2.3.2 HIV-1 drug resistance mutations among chronically infected drug-naïve participants

Although this group of participants was supposed to be drug-naïve, drug resistance mutations were detected for all the three classes of ARVs, albeit in small proportions. Twenty-two drug resistance mutations, minor and polymorphic mutations at drug resistance sites have been detected: eight major PI resistance mutations (L23I [3.53%], L33F [3.53%], M46L [3.53%], G48V [1.38%], I54V [3.53%], G73A [1.38%], I84C [1.38%], and L90M [3.53%]); eight NRTI mutations (M41L [3.53%], D67N [5.88%], L74V [3.53%], M184V/I [2.35%], L210W [3.53%], and T215Y [3.53%]); and 6 NNRTI mutations (A98G [1.38%], L100I [3.53%], K103N [4.71%], Y181C [3.53%], and Y188C/H [2.35%]). Of note here is that three (3.5%) participants already harbored type-1 Thymidine Analogue Mutations (TAMs), which confer higher level reduction in clinical response and cross resistance to ABC, ddI, and TDF (Shafer and Schapiro, 2008).

This drug resistance mutation profile is different from that obtained from drug-naïve patients in North Western Ethiopia, where no major PI resistance mutations were detected but 1.1% and 2.2% of these people harbored V75I (NRTI resistance) and G190A (NNRTI resistance) respectively (Kassu *et al*., 2007), both of which absent in the current study. However, studies from elsewhere on ART starting drug-naïve patients have reported more-or-less similar prevalence rate for the above mutations detected in the current study. For example, a study

carried out in Mexico to determine prevalence rate of drug resistance mutations from chronically infected drug-naïve patients has found that prevalence of NRTI resistance mutations T215Y/C, F77L, D67N/S, and M184V was at the rate of 3% for the first two, and 2% for the latter two (Escoto-Delgadillo *et al*., 2005). This is in addition to an overall resistance detection rate of 6% to delavirdine and nevirapine, 4% to efavirenz, and 2% nelfinavir (Escoto-Delgadillo *et al*., 2005). Others have reported a bit higher resistance mutation rate than the report in the current study. Hogg *et al*. (2006), for instance, have reported a baseline resistance rate of 7.8% to any class of ARV drugs among drug-naïve patients. They further reported that 73.6%, 18.9%, and 7.6% of these patients exhibited resistance to one class, two classes, and three classes of ARV drugs (Hogg *et al*., 2006).

Three potential sources of drug resistance mutations responsible to these drug-naïve participants could be speculated. One obvious possibility is that they might have been transmitted from persons who were on ART. Given the fact that all of these patients were chronically infected and therefore were seeking to start ART, it would be logical to think that transmitted resistance mutations would not sustain this long without reverting to wild type or non-resistant forms (Pao *et al*., 2004), an assumption which may lead to exclude the possibility of their being transmitted. However, there are also observations that some transmitted resistance mutations (such as K103N, V179E, T215Y, etc.) can persist up to 7 years without reversion (Brenner, *et al*., 2004). In fact, some of the mutations detected in this study were of among these types, although some others were among those which were observed to disappear or revert to wild type in a prolonged absence of drug-selective pressure (Pao *et al*., 2004; Turner *et al*., 2004). It is thus possible that some of the mutations detected here could be transmitted.

Another possibility for the occurrence of these resistance mutations could be due to undisclosed previous exposure to the drugs as monotherapy or combination therapy. Again this scenario is possible under the setting in which HIV-infected patients in Ethiopia were found just before ART scaling-up was launched in 2005. It was documented that antiretroviral drugs were in use since 2000 prior to the launching of ART Initiative in Ethiopia in 2003; but few people had access to these drugs from black market (MOH, 2005). Between 2003 and 2005, five ARVs were distributed to selected Urban Dwellers Association Pharmacies and Ethiopian Red Cross Pharmacies in Addis Ababa: three NRTIs (AZT, 3TC and D4T) and two NNRTIs (NVP and EFV). Thus, the persons who harbored the 22 drug resistance mutations in this study might be among those patients who had been using drugs before the scale-up.

The third possible source of the resistance mutations could be presence of already established viral lineages that contained drug resistance mutations independent of treatment, which circulate among HIV-infected persons even before the introduction of ART in Ethiopia. This kind of phenomena has been documented among Ethiopians and other people elsewhere. For example, Loemba *et al*. (2002) have found treatment-independent K70R and G190A naturally harbored by ARV drug-naïve two Ethiopian Jews. Hue *et al*. (2009) also identified five treatment-independent viral clusters containing mutations conferring cross-resistance to antiretroviral drugs prescribed today in the United Kingdom. Using reconstruction of dated Phylogeny through Bayesian Markov chain Monte Carlo inference, the authors have concluded that these reservoir lineages originated between 1997 and 2003 and have persisted in the HIV-infected population for up to 8 years.

In order to test if such viral reservoir viral lineages existed in Ethiopia in the early years of HIV infections, one has to run genotyping using early sequences at the *pol* region. In fact, 10 previously submitted (perhaps before ART scale-up began in Ethiopia and so presumably drug-naive) Ethiopian sequences were retrieved from the Los Alamos HIV database, and analyzed for the presence of drug resistance mutations. Two of them already harbored multidrug resistance mutations: PI resistance mutations L23I, M46I, F53L, I54V, V82F, I85V, L90M; NRTI resistance mutations F77L, F116Y, Q151M; and NNRTI resistance mutations K101E, Y181I, G190A. As can be seen here, three of the seven PI resistance mutations detected in this study (namely, L23I, I54V, and L90M) were already harbored in patients who were infected years before ART roll-out; and in one of the other PI resistance mutation position detected in this steady, there was M46L instead of M46I harbored in the earlier patients. Among NNRTI resistance mutations, there was also a-Y181I mutation in the previous infection instead of Y181C found in the current infection. This may be an indication that HIV variants already harboring drug resistance mutations were circulating within previously infected people in Ethiopia.

Regardless of how the drug-naïve patients may have acquired the resistance mutations, what is worrying most should be the therapeutic implications of their presence when ART administration starts. With the selective growth advantage of the mutant viruses when the infected persons are taking ARV drugs, they proliferate and reduce the potency of the treatment (Brenner *et al*., 2000). This is in addition to their involvement in flaring emergence of more drug resistance mutations (Little *et al*., 2002). In support of this, several cohort investigations have already found the link between length of treatment success and existence of resistance mutations prior to commencement of ART. For example, Hogg *et al*. (2006) observed that prevalence of drug resistance mutation to any class of ARV rose to 26.5% after the first 30 months of treatment from

7.8% observed at baseline; and mortality was eminent with each emergence of any resistance, particularly with NNRTI resistance mutations. Similarly, Little *et al*. (2002) also reported an increase in prevalence of multi drug resistance mutations from 3.8% at baseline to 10.2% after treatment. This raised prevalence was accompanied with poor virological and immunological responses: after initiation of ART, the time to viral suppression was longer and virologic failure shorter for patients with baseline drug resistance mutations than those without (Little *et al*., 2002). These and more other studies showed that prior exposure of patients to viral variants with drug resistance mutations hamper the patients' ART options severely. Thus, the participants in our study with the drug resistance mutations indicated above might have suffered from poor treatment gains.

4.2.3.3 Drug resistance mutations from drug-experienced

In this group of study, about 31% (83/266) of patients had virological treatment failure in its loose definition of having detectable viral load (≥150 copies/ml). However, only 66 of the 83 (~80%) failing samples were genotyped. About 70% of the participants from genotyped specimens or about 55% of all viremic participants harbored at least one confirmed major drug resistance mutation from any class. The 70% prevalence rate of drug resistance mutations among these heavily treated patients in this report is a little lower than the rate reported from some sub-Saharan African countries: e.g. 78% in Sweto, South Africa (El-Khatib *et al*., 2010) and 79% in Cote d'Ivoire (Adje-Toure *et al*., 2003). But an earlier study from Gabon had reported lower rate (58%) (Vergne *et al*., 2002) compared to the one reported this study. The rate in the USA and Canada, where it was reported to be over 76% (Kantor and Katzensein, 2004; Richman *et al*., 2004), is more-or-less similar to the report in the other African countries (except in Gabon). With prevalence rate of 83% of viremic patients, drug resistance in France was reported to be slightly

higher than the African and North American reports (Costagliola *et al*., 2007). It seems that prevalence of drug resistance mutation rate seen in Africa, where ART treatment has been introduced relatively recently, is comparable to that in North America and Europe, where ARV treatment has been in place for more than a decade-and-half. Several reasons could be attributed for this high prevalence: presence of resistant variants circulating among drug-naïve persons prior to treatment initiation; absence of treatment outcome monitoring tests such as viral load and drug resistance tests; suboptimal treatment due to poor adherence or poor drug absorption or some other factor related to the genetic background. This is alarming considering the limited therapeutic options Africans have.

In this study, the mutation rates according to drug class were: 2/66 (3%) for any major resistance mutation against PI, 41/66 (62%) against NRTI and 45/66 (68%) against NNRTI. While any major RT resistance for double class NRTI and NNRTI was found 40/66 (~61%), any triple class mutation was 2/66 (3%). The low mutation rate for PI resistance must be because the majority patients were not exposed to PI drugs; only one of the two PI resistance harboring patients was exposed to PI drugs. While any PI resistance mutation prevalence in this study (3%) was comparable to other reports from Africa (South Africa, 2% [El-Khatib *et al*., 2010]; Cote d' Ivore, 1.5% [Adje-Toure *et al*., 2003]), it was much lower than those reported in resource-rich countries (USA, 40.5% [Richman *et al*., 2004]; Italy 19.9% [Tozzi *et al*., 2006]; and France 7% (Costagliola *et al*., 2007). This shows the limited availability and hence limited use of PI drugs in African countries.

Regarding prevalence of any NRTI resistance mutations, the finding in this study (62%) was more-or-less closer to the one reported in other African countries (South Africa, 64% [El-Khatib

et al., 2010)]; and Cote d' Ivore, 56.3% [Adje-Toure *et al*., 2003]). Whereas reports from India (90% [Kumarasamy *et al*., 2009]), France (74%, [Costagliola *et al*., 2007]), and USA (71.4%, [Richman *et al*., 2004]) showed worse NRTI resistance prevalence rate than the reports from previously mentioned African countries including from the current study. However, the report from Italy was much lower (28.5%, [Tozzi *et al*., 2006]). By comparison, resource-constraint countries, including Ethiopia in this study, showed higher NNRTI resistance mutation rates than resource-rich countries: 68% Ethiopia (this study); 81% South Africa (El-Khatib et al., 2010); 87% Botswana (Doualla-Bell *et al*., 2009), and 65% India, (Kumarasamy *et al*., 2009). An exception to this generalization is the report from Cote d' Ivore, where only 3.5 % of the patients had resistance mutations for NNRTI, which may not be surprising since only one patient was treated with NNRTI-containing regimen (Adje-Toure *et al*., 2003). The prevalence in the USA, France and Italy was 25.2% (Richman *et al*., 2004), 44% (Costagliola *et al*., 2007), and 57.7% (Tozzi *et al*., 2006), respectively. It seems that where PI resistance is less, NNRTI resistance is more (e.g. in the reports above in the USA, 40.5% for PI resistance Vs 25.2% NNRTI resistance). In general, the data show that people in resource-limited countries who were ART treated and who already developed resistance are left with narrower future options particularly for NNRTI use.

Specific resistance mutations against PIs, NRTIs and NNRTIs

In regards to specific resistance mutations in each class of ART drugs detected from this study, five major PI resistance mutations at four sites occurred among two individuals: I54A (1.51%), V82A (2.51%), N88D/S (3.03), and L90M (1.51%). While one of these two patients was on PI drugs, the other was not. The latter patients might have acquired the resistant variants during

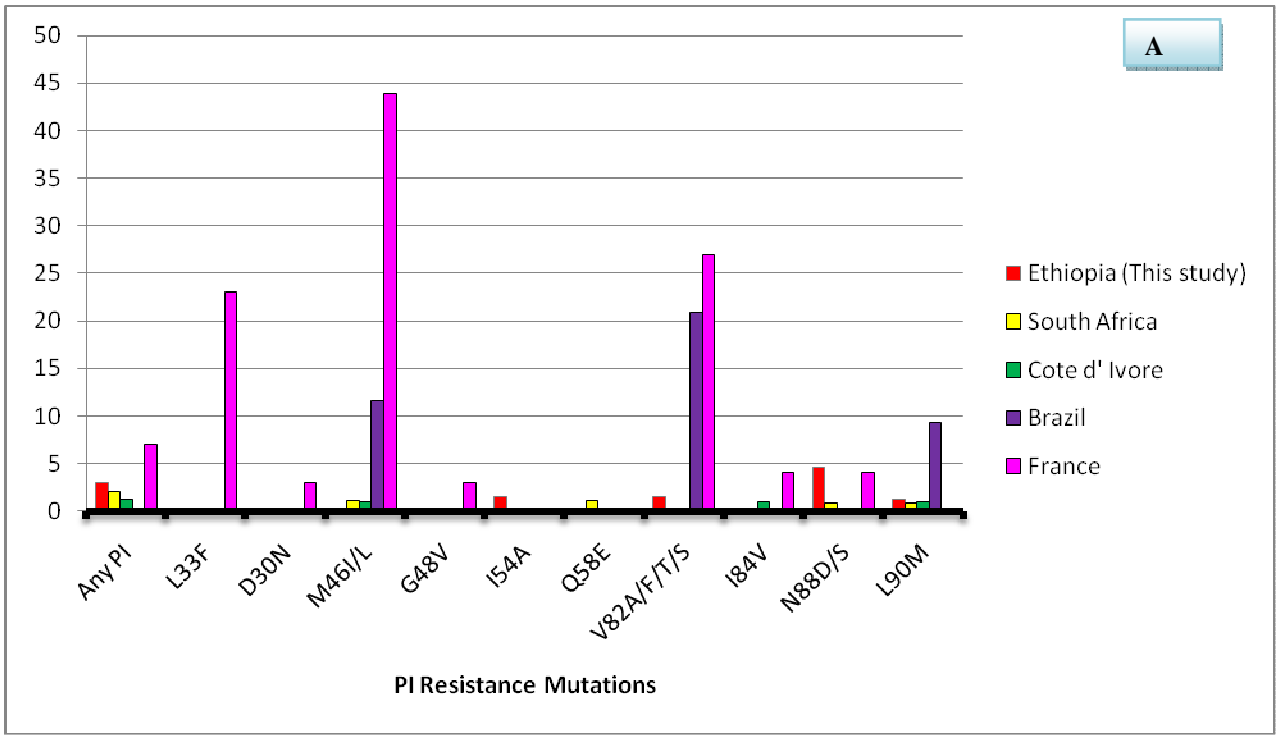
primary infection (transmitted) or could be that the variants mutated without drug exposure from wild type or from intermediate genetic background that is prone for change.

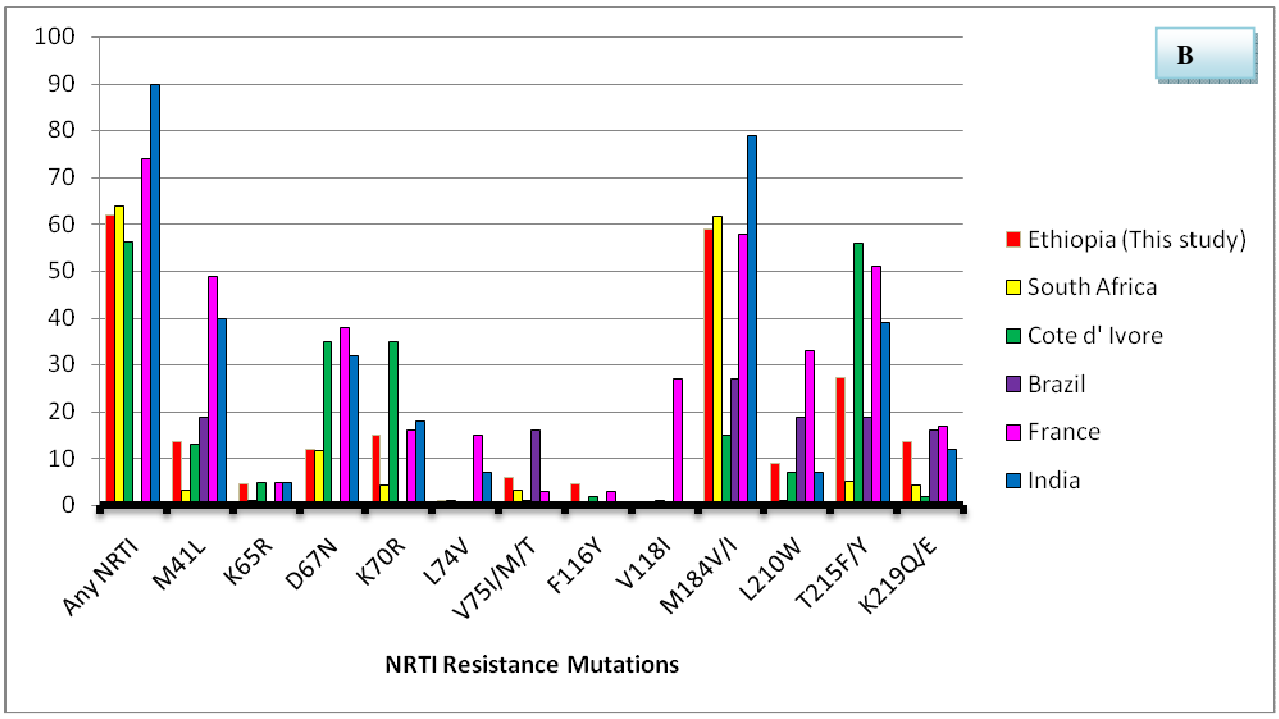
The most common major NRTI resistance mutations observed from successfully genotyped viremic patients were M184V/I (59.09%), T215F/Y (27.27%), K70R (15.15%), K219Q/E (13.64%), M41L (13.63%), D67N (12.12%), and L210W (9.09%). Although all the three type I Thymidine analog mutations occurred individually at various rates, they were found to have occurred simultaneously in 4 (6.1%) patients. However, individual mutations also occurred in pairs at various rates. For example, M41L and L210W co-occurred 41.22% of the time (p=0.000); M41L and T215Y, 28.41% of the time (p=0.000); and L210W and T215Y, 32.04% of the time (p=0.000). Moreover, two members of type II analog D67N (2/66) and K219Q (1/66) co-occurred with the complete battery of type I analogs. This is in addition to the occurrence of individual members of type II analogs with members of type I analogs singly or in pairs. For example, D67N and K70R co-occurred 24% of the time (p=0.000); D67N and T215F, 22.56% of the time (p=0.000); D67N and K219Q, 32.83% of the time (p<0.001); D67N and K219E, 11.16% of the time (p<0.001); K70R and T215F, 16.24% of the time (p=0.001); K70R and K219Q, 38.44% of the time (p=0.000), etc.

Figures 4.2.1.A-C summarize prevalence rates of individual resistance mutations to PI (Panel A), NRTI (Panel B), and NNRTI (Panel C) detected in this study, and compares them with those reported from patients of various countries distributed in four continents (Africa, Latin America, Europe and Asia). In this comparison, it can be seen that, except for Cote d' Ivore, the most prevalent NRTI mutation in all others was M184V/I; in Cote d' Ivore, T215F/Y was the most prevalent one (Adje-Toure *et al*., 2003). In all cases including in this study, M41L, D67N, K70R

and K219Q/E were most abundant next to M184V/I and T215F/Y. K65R and F116Y were observed at the lowest rate in all the six countries. In fact, F116Y was reported only from this study (4.55%), France (3%) (Costagliola *et al*., 2007), and Cote d' Ivore (2%) (Adje-Toure *et al*., 2003). There seems country-dependent variation in the prevalence of some NRTI mutations. For example, D67N was more prevalent in Cote d' Ivore (Adje-Toure *et al*., 2003), France (Costagliola *et al*., 2007) and India (Kumarasamy *et al*., 2009) compared to the other three countries; K70R was more abundant in Cote d' Ivore than the other countries; M184V/I was more prevalent in India (Kumarasamy *et al*., 2009) than the rest countries, although Ethiopia, South Africa and France are following having more-or-less similar rate.

The leading primary NNRTI resistance mutation in this study was K103N (34.84%) followed by Y181C/I (18.18%), and V106A/M (each 15.15%), L190S/A (12.12%), Y188L/C/H (4.54%). Major secondary mutations detected included L100I (3.03%), L101E/P/H (15.15%), P225H (9.09%). These major secondary mutations co-occurred with the primary mutations in several occasions: e.g., K100I and Y188L, in 14.29% of the time; and K101E and G190S, in 15.37% of the time. Minor NNRTI resistance mutations detected include A98G (15.15%), V108I (9.09%), and V179D (4.54%). A different study from Italy showed reduced prevalence for most NNRTI mutations (K103N, 14.5%; V106A, 1.4%; and comparable rate for others (Y181C/I, 17.2%) (Rosina *et al*., 2000). However, reports from Southern Africa region, where subtype C is dominating like Ethiopia, have shown an elevated rate for some mutations compared to observed in this study (e.g. K103N, 47.9%, V106A/M, 26.6%, and Y188C/L/H 13.8% in South Africa (El-Khatib *et al*., 2010); and V106A/M 30%, Y181C 23%, and G190A ~23% in Botswana (Doualla-Bell *et al*., 2009).


A
50
45
40
35
30
25
20
15
10
5
0
Any PI
L33F
D30N
M46I/L
G48V
I54A
Q58E
V82A/F/T/S
I84V
N88D/S
L90M
Ethiopia (This study)
South Africa
Cote d' Ivore
Brazil
France
PI Resistance Mutations


B
100
90
80
70
60
50
40
30
20
10
0
Any NRTI
M41L
K65R
D67N
K70R
L74V
V75I/M/T
F116Y
V118I
M184V/I
L210W
T215F/Y
K219Q/E
Ethiopia (This study)
South Africa
Cote d' Ivore
Brazil
France
India
NRTI Resistance Mutations

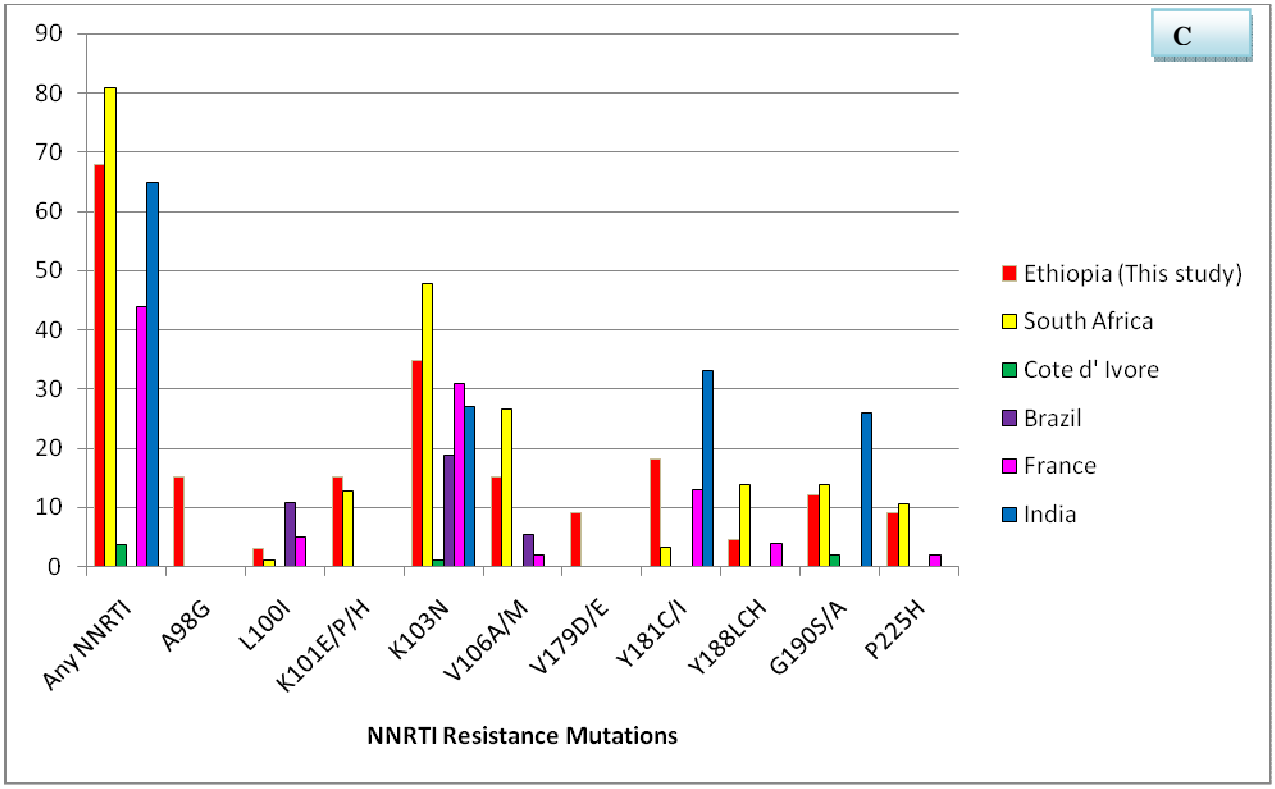


Figure 4.2.1. Prevalence of selected PI (Panel A), NRTI (Panel B), and NNRTI (Panel C) resistance mutations (in %) among Ethiopia (this study), and published reports from four other countries.

Associations of NRTI major resistance mutations observed in the context of NRTI drugs used in the treatment

The potential therapeutic consequences of NRTI resistance mutations observed in this study are disturbing. First of all, mutations have been detected against all the regimens used both in the starting and the current therapy. It was found out that 33/98 (~37%) of stavudine (D4T)-containing and 32/163 (~20%) of zidovudine (AZT)-containing regimens have developed at least one major NRTI resistance mutations in the initial regimens. The difference between these two regimens in selecting resistance mutations was statistically significant, more D4T than AZT selection was observed (Mean difference of 2.6875 %; SD= 3.0048; 95 CI= 1.08632-4.28868; p= 0.003). Perhaps the toxicities associated with D4T (DHHS, 2011; WHO, 2010) could have prompted patients to escape their dose, which would leave them under suboptimal treatment. In

the current therapy, resistance mutations were developed 18/51 (35%), 40/155 (26%), and 7/51 (14%) against D4T-, AZT-, and Tenofovire (TDF)-containing regimens, respectively. In this case, there was no statistically significant difference in the selection rate between D4T and AZT (Mean difference= 4.0; SD= 9.42338; 95% CI= -9.02136-1.0213; p= 0.110). In both initial and current therapies, all M184V, thymidine analog resistance and non-thymidine alnalog resistance mutations have occurred against these NRTI containing regimens. For example, M184V occurred in 19/33 (58%) of D4T- and 19/32 (59%) of AZT-containing failing regimens in the initial therapy; and 7/18 (39%) of D4T-, 26/40 (65%) of AZT-, and 5/7 (71%) TDF-containing currently failing regimens. This may not be surprising given that Lamivudine (3TC), the drug that selects the mutation M184V/I (Maserati *et al*., 2010; Whitcomb *et al*., 2003), was incorporated in all of the regimens used in treating the patients. However, the phenotypic effect of this mutation might not be as bad as its genotypic prevalence rate since published observations have shown that this mutation also enhances susceptibility of the viral isolates to the three background NRTI drugs used; D4T, AZT and TDF (Diallo *et al*., 2003; Miller *et al*., 2002).

Secondly, due to the rampant prevalence of thymidine analog resistance mutations, there is a remarkably high potential to reduce the efficiency of alternative even non-thymidine analog NRTIs such as ABC and ddI, should next-line switching become necessary. This can be seen clearly when one considers the high prevalence rate of failing thymidine analog NRTI regimens against individual thymidine analog resistance mutations such as M41L (11% in D4T-containing and 10% AZT-containing), L210W (13% in AZT-containing), T215Y and T215F (17% and 6% respectively in D4T-containing and 15% and 10% in AZT-containing), D67N and K70R (6% each in D4T-containing, and 15% and 23% respectively in AZT-containing regimens), along with some compensatory mutations such as T215S/I.

This potentially restricting therapy option will further be exacerbated by wide spread abundance of resistance mutations usually selected by non-thymidine analogs such as abacavir (ABC) and 3TC. In this regard, K65R prevalent in 9% and 6% of failing D4T- and AZT-based regimens in the initial and current therapy respectively, is known to render the viral isolates resistant to TDF, ABC and 3TC (Moyle, 2001; Theys and Vandamme, 2008). It has been reported in several published results that K65R occurs in viruses treated with D4T/ddI and ddI/3TC (Boucher *et al*., 2006; Hawkins *et al*., 2007), or TDF/3TC pairs (Maserati *et al*., 2010). Moreover, occurrence in this study of other non-thymidine analog resistance mutations such as L74V and V75T (each 6% in D4T-containing), V75I (3% in AZT- and 14% in TDF-containing), V75M (3% in AZT-containing), and F116Y (11% in D4T- and 3% in AZT-containing) regimens constricts further the already narrow alternative next-line treatment options, because these resistance mutations are selected by 3TC, ddI and ABC, the latter two being components of alternative first-line and part of the second-line regimens in the Ethiopian ARV treatment options (MOH/HAPCO, 2007b). Figures 4.2.2 summarize prevalence of individual NRTI resistance mutations in regimens containing AZT and D4T within initial therapy (Panel A) and regimens containing AZT, D4T, and TDF in the current therapy regimens (Panel B).

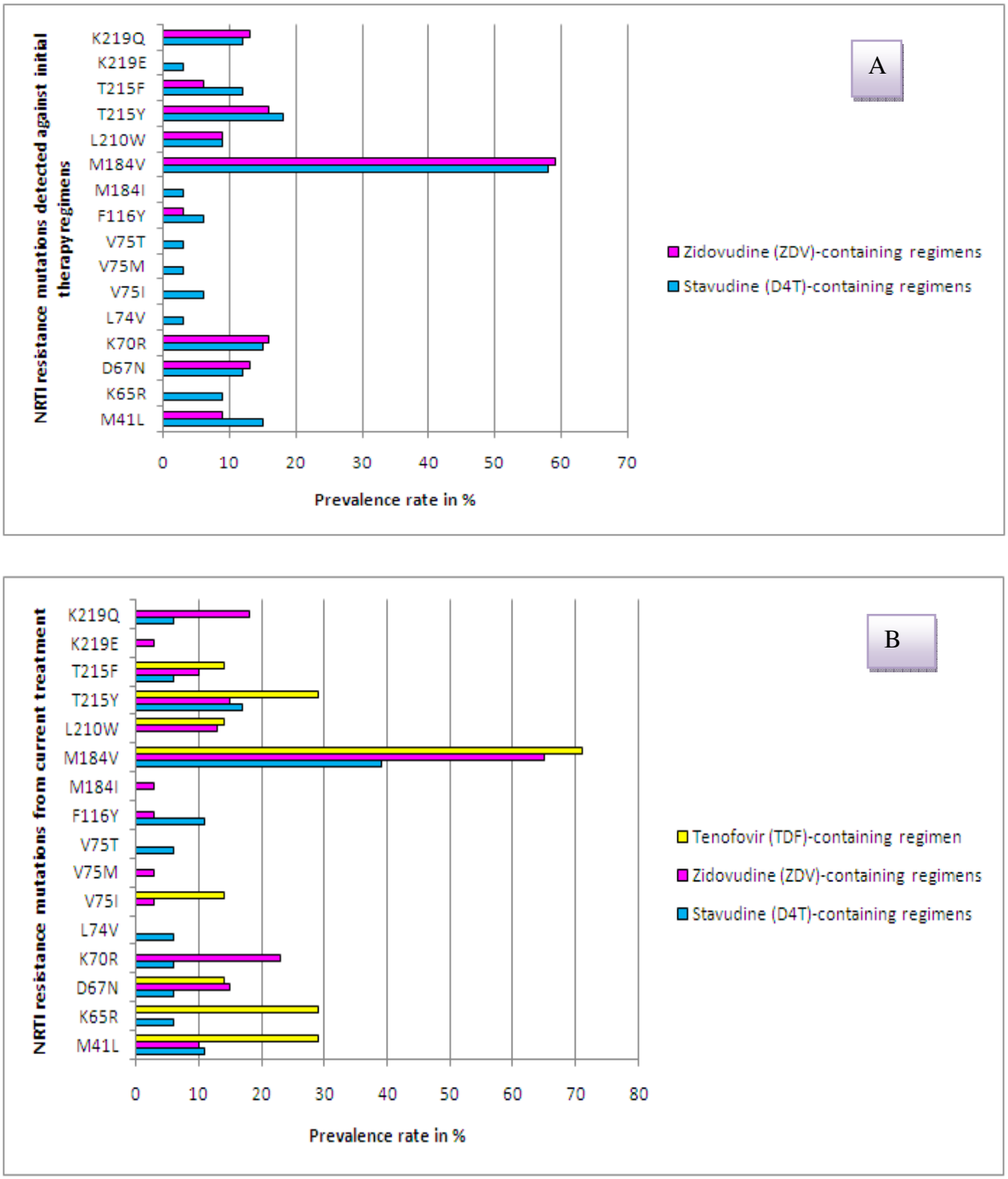


Figure 4.2.2 Prevalence rate of NRTI resistance mutations among treatment experienced patients by NRTI drugs they were exposed to in the starting regimens (Panel A) and current regimens (Panel B).

Associations of NNRTI major resistance mutations observed in the context of NNRTI drugs used in the treatment

Because of the intrinsic low genetic barrier in the RT allosteric sites to develop resistance mutations, occurrence of high level or multi-drug resistance with only one or two mutations at these sites is common in settings where NNRTIs are used as background drugs in the ART combinations (Shafer and Schapiro, 2008). In this regard, the data in this study showed disturbingly worse results for NNRTI resistance than it did for NRTI resistance (Figures 4.2.3. A and B). In the initial therapy, 37/163 (~27%) of treated patients with NVP-containing regimens or 21/38 (~55%) of NVP-exposed patients with genotyped viral isolates harbored at least one NNRTI resistance mutation. Similarly, 28/98 (~29%) of patients with EFV-containing regimens or 27/28 (96%) of EFV-exposed patients with genotyped viral isolates had at least one NNRTI resistance mutation. There was no statistically significant difference between rates of mutations selected by NVP and EFV in the initial therapy regimens as determined by paired T-test (t=0.423; p= 0.678), although EFV selection was slightly greater than that of NVP.

The status of resistance at the time of genotyping was 35/135 (~30%) for NVP-containing (or 35/36 (97%) of patients on NVP-containing regimen with successfully genotyped viral isolates, and 27/108 (25%) for EFV-containing (or 27/27 (100%) of patients currently on EFV-containing therapy with successfully genotyped viral isolates. Three persons who switched to 2^{nd}-line regimen containing 3TC-TDF-kaletra with previous exposure to EFV in the 1^{st}-line therapy, also harbored the mutations V179D, Y181I, and G190A. This may suggest an evidence for the high selection power of EFV more than NVP under the setting in this study. The difference in rates of resistance mutations selected by NPV and EFV at the time of genotyping was not statistically significant (t=0.353; p= 0.729), although EFV selection rate was slightly greater than that of

NVP. Another study conducted in Brazil also found no statistically significant difference in selecting major NNRTI resistance mutations between NVP- and EFV-containing drugs (Medeiros *et al*., 2007). However, the rates of resistance mutations against NNRTI drugs among NNRTI-exposed and genotyped patients in this study was higher than reports from other similar studies elsewhere. For example, in Iran frequency of NNRTI resistant isolates among treatment experienced patients was 62% for NVP while it was 45% for EFV (Hamkar *et al*., 2010). Similarly, the Brazilian study found that about 43% of NVP taking and 37% EFV taking patients for whom genotyping was done harbored NNRTI resistance mutations (Medeiros *et al*., 2007).

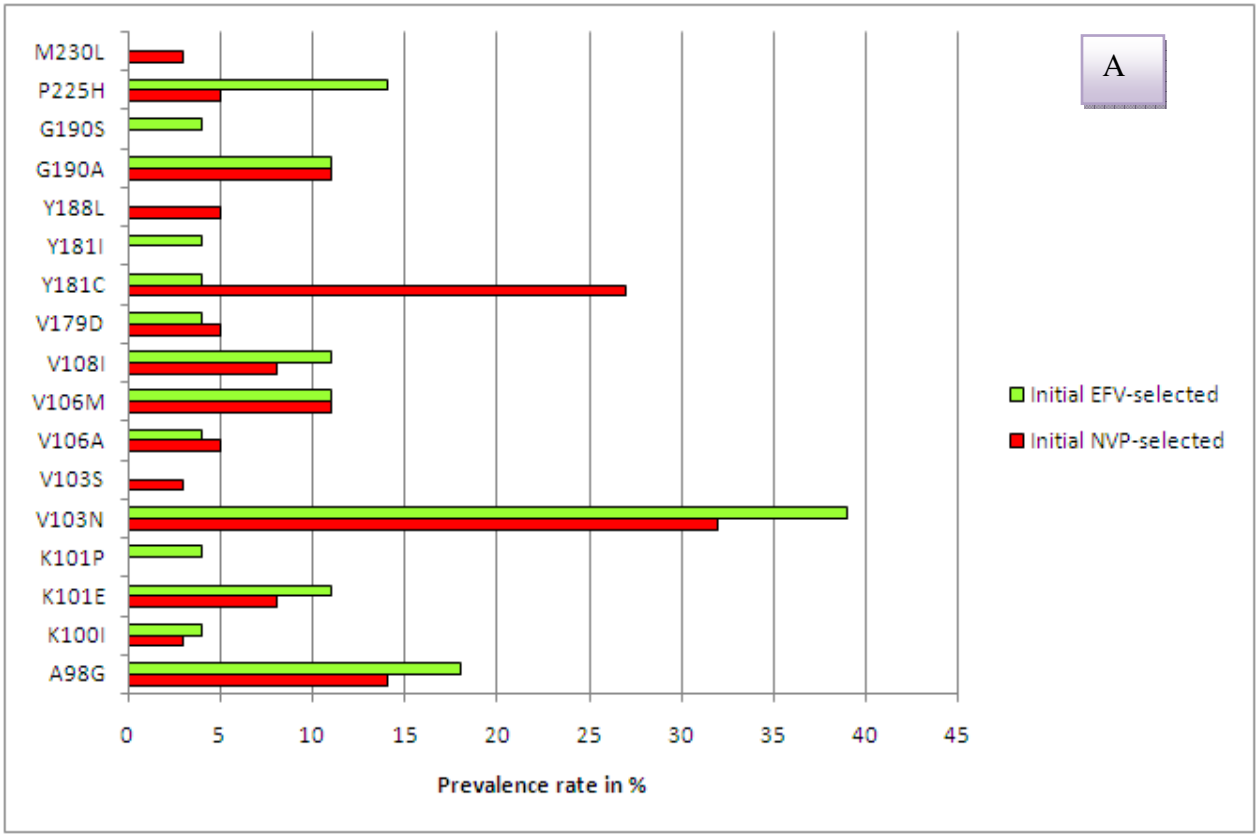
In this study, from regimens administered at the time of genotyping, NVP preferentially selected for 14 major resistance NNRTI mutations, EFV for 15, and 12 of these selections were common for both. Among the most frequent NNRTI resistance mutations at the time of genotyping, K103N was selected by NVP and EFV at the rate of 52% and 44% respectively. Other resistance mutations selected by the two NNRTI drugs included: A98G, 24% by NVP vs 19% by EFV; K101E, 19% by NVP vs 7% by EFV; Y81C, 43% by NVP vs 7% by EFV; and P225H, 5% by NVP vs 19% by EFV. There were also some mutations selected by one of the drugs but not by the other: K101P, 10% by NVP vs 0% by EFV; V103S, 0% by NVP vs 4% by EFV; Y181I, 0% by NVP vs 3% by EFV; G190S, 0% by NVP vs 4% by EFV; and M230L, 5% by NVP vs 0% by EFV.

Given documented selection of most of these mutations by both drugs (Shafer and Schapiro, 2008), it is not clear why in this study some mutations were preferentially selected by one drug but not by the other. For example, all K101P and M230L mutations were selected by NVP but none by EFV. On the other hand, V103S, Y181I, and G190S were all selected by EFV not by

NVP, although they were supposed to be selected by both (Shafer and Schapiro, 2008). The reason(s) for this discrepancy is not clear. Some unknown factors might have influenced the pattern of emergence of these resistance mutations in this study. One possibility could be interaction within or between classes of resistance mutations. For example, researchers have found that simultaneous emergence of some NNRTI resistance mutations such as G190S along with the NRTI mutation T215Y could facilitate resistance to both NNRTI and NRTI drugs (Paolucci *et al*., 2007 cited in Ghosn *et al*., 2009). Although these two mutations were not observed in this study, such co-emerging mutations whose importance was not determined so far might have contributed to the observed lack of selection by one drug in this study. Moreover, it could also be possible that some of the mutations expected to appear in this study were less abundant so that they escaped detection by the population genotyping system used, a condition common in population genotyping (Le *et al*., 2009).

Another reason could be due to the background polymorphism differences between prevalent viral isolates in Ethiopia and those from which phenotypic and genotypic data on treatment responses were generated, mostly HIV-1 clade B genotypic background. In this regard, existence of impact of background polymorphism on selection of drug resistance mutations has been ascertained by several investigations. Spira and co-workers' studies have made many enlightening observations regarding importance of genetic background in causing different pattern of drug resistance among HIV-1 isolates by comparing in vivo and in vitro phenotypic and genotypic resistance patterns between subtype B and C (Spira *et al*., 2003). It was shown that (i) there were many baseline polymorphisms and silent mutations within RT and PR at sites linked to resistance to NNRTI and NRTI in clade C (most of the work was done on clade C isolates from Ethiopia and Botswana) compared to clade B; (ii) intrinsic resistance to NVP and

EFV was observed in the C clade without resistance to another NNRTI drug Delavirine (DLV); (iii) clade C viral isolates required reduced concentration of NVP (2 µM) and EFV (0.01 µM) to develop drug resistance mutations compared to that required for clade B, 10 µM and 1 µM, respectively, showing that clade C is more prone to develop drug resistance mutations quickly under elevated concentration than clade B; (iv) several novel resistance mutations could develop in clade C when isolates were re-exposed to NVP (including A98I, A98S, K103N, V106M, etc). In addition, it was also observed that the baseline polymorphism at codon 106 in C facilitated development of the mutation V106M that confers resistance to EFV.



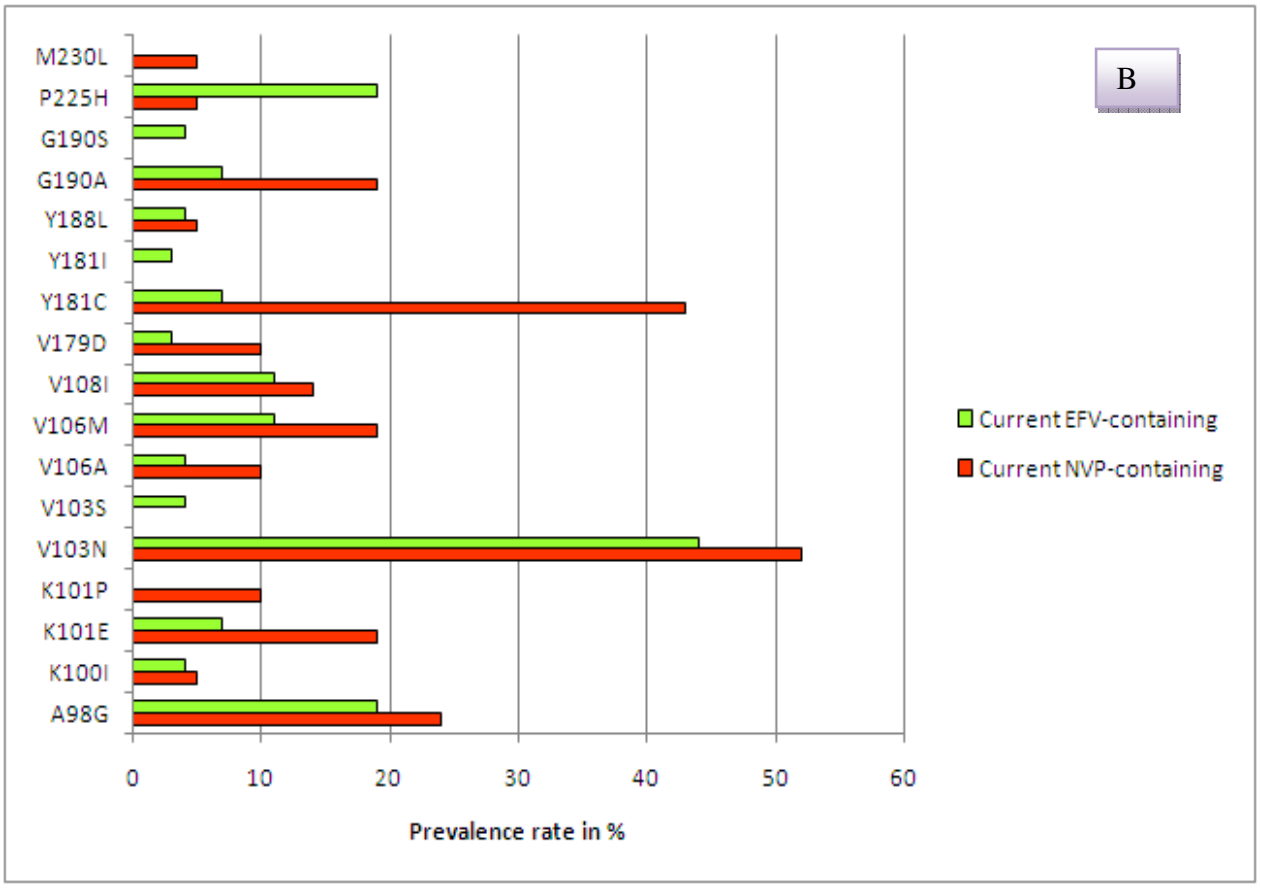


Figure 4.2.3 Prevalence rate of NNRTI resistance mutations among treatment experienced patients by NNRTI drugs they were exposed to in the starting regimens (Panel A) and current regimens (Panel B)

Patterns of co-variations of major resistance mutations

In view of the importance of double or multiple resistance mutations in compromising therapeutic efficacy of the ART drugs, simultaneous occurrence of two resistance mutations in viral isolates from each individual patient was conducted using Pearson correlation. A total of 85 statistically significant pairs of positive co-variations were detected among major resistance mutations against the three classes of ART drugs: 2 pairs within PI, 14 pairs within NNRTI, 29 pairs within NRTI, 20 pairs between PI and NRTI/NNRTI, and 30 pairs between NRTI and NNRTI resistance mutations. In the RT region, the most frequent co-variation was observed on members of thymidine analog resistance mutations. Strong correlations within NRTI resistance

mutations occurred between M41L/L210W (41%), K70R/K219Q (38%), D67N/219Q (33%), and L210/T215Y (32%), all of which happened to be thymidine analog resistance mutations. This kind of observation was also documented previously from subtype B viral isolates by Rhee *et al*. (2007).

Strong correlations within NNRTI resistance mutations were seen in V179D/Y181I and V179D/M230L, each occurring 32% of their appearances, the average being 15%. This correlation profile is different from that reported by Reuman *et al*. (2010) in that these authors found Y181C to have strong correlation with G190A, H221Y, V108I, A98G, and V179F; but in this study, Y181C had no correlation with any of its NNRTI class counter parts. They also found that L100I was strongly correlated with K103N and K101P; whereas in this study, L100I was correlated only with Y188L in only 14% of the time. Rather, the following were among notable NRTI-NNRTI correlations detected in this study: K65R/Y181C (9%), F116Y/V106A (100%), M41L/A98G (11%), T215T/A98G (15%), M184V/A98G (12%), M184V/K103N (12%), K70R/M230L (9%), etc. It is alarming to find such high level of co-variation within and between resistance mutations of the three classes since coexistence of relevant mutations might abrogate treatment efficacy in all the three classes of ART drugs.

4.2.4 Discussions on subtype prevalence

A total of 191 sequences from the three groups of study participants were analyzed for subtype assignment using five genotyping systems. There was a complete concordance between these in 184 of 191 (96%) sequences at Pol. This study clearly showed that the HIV-1 isolates genotyped in this study were overwhelmingly dominated by subtype C (183/191 [96%]) both at the PR and RT regions. Of the remaining eight sequences, two (~1%) were unanimously assigned to subtype B, whereas the rest six were assigned to different subtypes by the five algorithms used. However,

the majority of the algorithms have identified the samples to have contained recombinant sequences: Et-97-58-098 was assigned to CRFAG-02 by Stanford and phylogenetic analysis at the Pol region; to BG by jpHHM and phylogenetic analysis at RT and PR regions separately; and to ABG by SimPlot. In the same way, Et_97-58-063 was assigned to BC by jpHHM; 97-58-163, to BC by jpHHM and SimPlot; and Et_97-61-421, to BC by jpHHM. However, sequence "Et_97-61-057" should be considered as subtype C since three of the five algorithms agreed on subtype C, and the remaining two algorithms provided no different option. Two of the remaining discrepant sequences were also assigned recombinant forms by at least one algorithm: V1311DRZ2, BC by jpHMM; and V1349DRZ, A1D by jpHHM , SimPlot, and separate PR and RT phylogenetic analysis.

It was reported right from the onset of the epidemic that subtype C was the predominant HIV-1 variant in the Ethiopian epidemic (Abebe *et al*., 2001; 2000; 1997; de Wit *et al*., 2002). However, most of these subtyping studies were carried out either by sequencing short genomic regions such as C2-V3 (Abebe *et al*., 2000; de Wit *et al*., 2002) or V3 region alone or peptide serotyping (Hussein *et al*., 2000). The finding in this study regarding the predominance of subtype C on the PR and RT regions provides further evidence to the already available data. Nonetheless, detection of nine non-C subtypes including two subtype B and seven containing recombinant sequences signifies the presence of an ongoing introduction/emergence of HIV-1 subtypes not previously documented in Ethiopia. There were also two published reports on the presence of subtype A (Abebe *et al*., 1997) and D (Hussein *et al*., 2000), and another on a subcluster of subtype C called C' (Abebe *et al*., 2000), all of which detected from the envelop region, but not from the PR or RT regions.

It is true that once a particular subtype, such an expanding subtype as subtype C, sets foot in one region, the chance that it would be displaced by another externally introduced subtype would be difficult, if not impossible. Nonetheless, the presence of non-C subtypes in the surrounding countries coupled with current trend of international travel and sex trade (Perrin *et al*., 2003) make absence of non-C subtypes in Ethiopia unlikely. Just to cite few examples, in a recent assessment on subtype surveillance in Western Kenya, the region that shares boundary with Ethiopia, subtype distribution was found to be A (51.4%), D (9.4%), C (1.4%), and 4.1% recombinants including A1/C, A1/D, A1/A2, and A2/C (Oyaro *et al*., 2011). Previously subtype distribution in Northern Kenya, another region bordering Ethiopia, was C (42%), A (43%), and D (15%) (Khamadi *et al*., 2004). Another study from Nairobi on *pol/env* paired sequencing revealed subtype distribution of A1/A1 (65%), A/C (9%), A1/D (7%), C/D (2%), D/D (2%), A1/A2 (2%), G/G (2%), A2/D (2%), C/C (2%), and CRF02_ AG (7%) (Lihana *et al*., 2009b). Similarly, diverse subtypes were reported from the other countries surrounding Ethiopia: subtype D (50%), C (30%), and unspecified percentages of subtypes A and B from the Sudan (Hierholzer *et al.,* 2002); and subtypes C (73%), CRF 02_AG (18%), D (6%), and A (3%) from Djibouti (Maslin *et al*., 2005).

Thus, lack of non-C subtypes in the previous reports in Ethiopia could not be because of their absence, but could be because of total dependence of previous subtype analyses on envelope sequencing. In fact, few previously produced HIV-1 sequences which were available from the Los Alamos HIV database (http://www.hiv.lanl.gov/) were analyzed along with the sequences produced in this study using Stanford genotyping algorithm. This analysis provided evidence that some of the earlier varaints had a subtype composition different from C at the Pol region, where PR and RT proteins are coded: two of them belonged to subtype CRF02_AG (sequences

C.ET.1991. and E3099G), one to subtype CRF01_AE (sequence AC.ET.-.2_2A1), one to subtype A1 (sequence C.ET.1991.E3099G), and one to subtype D (sequence D.ET.2003.ETH_G_230). In addition, a previously conducted near-full length sequencing result identified an A/C recombinant in which identities of the *gag* and *pol* genes and the major parts of the *env* genes were of subtype A, whereas the regions flanking V3and the V3 domain itself was of subtype C (Sherefa *et al*., 1998).

These and other reports underscore the importance of using as large genomic region as possible to make correct subtype identification of isolates, suggesting that the PR and RT regions used in this study to characterize subtypes of Ethiopian isolates are not still sufficient. If some more regions such as Env and LTR were also genotyped, additional subtype profiles could have been obtained. An important concern in this subtype studies was the discordance results that different genotyping algorithms provided for the same set of sequence data. This problem looks universal. For example, Holguin *et al*. (2008a) found that Stanford, NCBI, and REGA showed discrepant results during subtyping of five non-B viral isolates from Spain which were identified by Phylogenetic analysis belonging to CRF06_cpx and CRF02_AG. The discordance of these three algorithms with each other and with phylogenetic analysis was evaluated using sequences from PR-RT regions of pure and recombinant Non-B subtypes; it was found that there were considerable variations between them (Holguin *et al*., 2008b).

In addition, availability of ART and improvement in the health status of HIV-infected patients who are on treatment coupled with reduced HIV status disclosure rate may prompt people to engage in risky sexual behaviors, as has been documented elsewhere (Elford *et al*., 2002; Shafer *et al*., 2011). Besides abetting prevalence of the already high HIV drug resistance (report in this

study) through transmission of resistant variants, this risky sexual behavior surely will lead to superinfection among the previously infected patients with HIV variants of different genetic background, which further amplifies diversity and emergence of recombinant variants (Gross *et al*., 2004). It is therefore critical to devise a strategy to regularly undertake surveillance for tracking the distribution of new HIV-1 genetic variants in the country.

4.2.5 Discussions on overall genetic diversity

In this study, it has been found that high rate of mutations were observed in all the three groups of participants. However, most of the mutations were of those types which are considered either as polymorphic in non-B subtypes in spite of their proven ability to confer resistance in subtype B, or were secondary resistance mutations which compensate for any fitness loses. For some non-resistance mutations, the rate of prevalence seems similar between the three groups. For example, PR mutations M36I, H69K, I93L, I15V, and R41K were highly prevalent among all the three groups with prevalent rate ranging between 73-95%. Similarly, RT mutations V35T, K122E, T200A, D177E, and V245Q were highly prevalent in all the three groups with prevalence rate of 79-87%.

For some others, however, there existed notable differences in mutation prevalence between the three groups, whether polymorphic or secondary resistance. For example, PR mutations K14R, G16E, E35D, and L89M were found at greater rate among recently infected drug-naïve than among the two chronically infected groups. In fact, it looks that these mutations progressively diminished with stages of the infection, in the order of recent drug naïve, chronic drug naïve, and chronic drug experienced (e.g. K14R, 24%-19%-4%; L89M, 93%-81%-77%, etc.). Conversely, some PR mutations were more prevalent in the drug-experienced group than in the two drug-naïve groups; for example, T12S (91% among drug experienced vs 76% and 39% respectively,

among chronic naïve and recent naïve groups); T74S (15% among chronic drug-experienced vs 6% among chronic naïve and 5% in recent naïve). In certain other instances, some PR mutations were observed to have occurred more among chronic naïve participants than both recent naïve and chronic experienced (e.g. I13V, 21% in chronic naïve vs 10% in recent naïve and 9% in chronic experienced; L63P, 34% in chronic naïve vs 25% in recent naïve and 23% in chronic experienced; L19I/T, 54% in chronic naïve vs 34% in recent naïve and 35% in chronic experienced).

Similar patterns of non-resistance or secondary resistance mutations with more elaborated diversity differences were observed in the RT region also. For instance, chronic drug experienced participants harbored more of the following RT mutations than chronic naïve and recent naïve participants: E36A (62% in chronic experienced vs 51% in chronic naïve and 54% in recent naïve); T39D (17% in chronic experienced vs 2% in chronic naïve, and 8% in recent naïve); S48T (44% in chronic experienced vs 34% each in the other two groups); K70G (14% in chronic experienced vs 0% in each of the other two groups); K173A (62% in chronic experienced vs 8% in chronic naive and 7% in recent naïve); and F214L (12% in chronic experienced vs 2% in chronic naïve and 1% in recent naïve). There were also instances in which some RT mutations were more prevalent among recently infected drug naïve than the other two groups (e.g. K49R, 15% in recent naïve vs 11% in chronic naïve and 8% in chronic experienced; K102Q, 12% in recent naïve vs 5% in chronic naïve and 2% in chronic experienced; D123G, 27% in recent naïve vs 22% in chronic naïve and 2% in chronic experienced; and A158S, 54% in recent naïve vs 4% in chronic naïve and 24% in chronic experienced.

In general, more genetic diversity was observed among chronically infected drug-experienced and recently infected drug-naïve participants at the RT region compared to diversity among chronically infected drug naïve persons. Two methods were utilized in this study to measure and compare genetic diversity between the three groups of participants, non-synonymous to synonymous substitution ratio and measure of Shannon entropy. It is a generally accepted assumption that synonymous substitutions (silent mutations) are neutral and that the synonymous substitution rate therefore approximates the neutral rate of evolution, as opposed to non-synonymous (amino acid altering) substitutions which are caused by positive selection and hence assumed to be diversifying selection (Doron-Faigenboim *et al*., 2005). The non-synonymous/synonymous ratio (expressed as ω= dn/ds) is taken as measure of selective pressure at the protein level, in which ω =1 meaning neutral mutations, ω <1 purifying selection, and ω> 1 diversifying positive selection (Yang e*t al*., 2000). The logic behind this assumption is that when ω>1, non-synonymous mutations offer fitness advantages to the protein (individual) and have higher fixation probabilities than synonymous mutations (Yang e*t al*., 2000).

Using SNAP software from the Los Alamos National Library website (http://www.hiv.lanl.gov/content/sequence/SNAP/SNAP.html), the non-synonymous and synonymous substitution rates were computed for the three groups separately and two more combined groups. It was shown that of the three homogenous groups, recently infected drug-naïve and chronically infected drug-experienced participants had higher rate of non-synonymous to synonymous ratio. The highest mean non-synonymous/synonymous ratio (nearly 1.3) was scored from drug-experienced sequences (Figures 3.5.14.1A-E and 4.2.4). Amino acid diversity in the recently infected group was greatest approximately in the entire PR and the last one-quarter of the RT, whereas in the drug experienced participants, the greatest non-synonymous substitutions and hence greatest

amino acid diversities were observed in the RT region, although the PR region also showed some degree of non-synonymous changes (Figure 3.5.14C). On the other hand, in the larger portion of sequences analyzed for chronically infected drug-naïve patients, it was the synonymous substitutions which were dominating throughout the sequences except in the last 50 or so codons (Figure 3.5.14B).

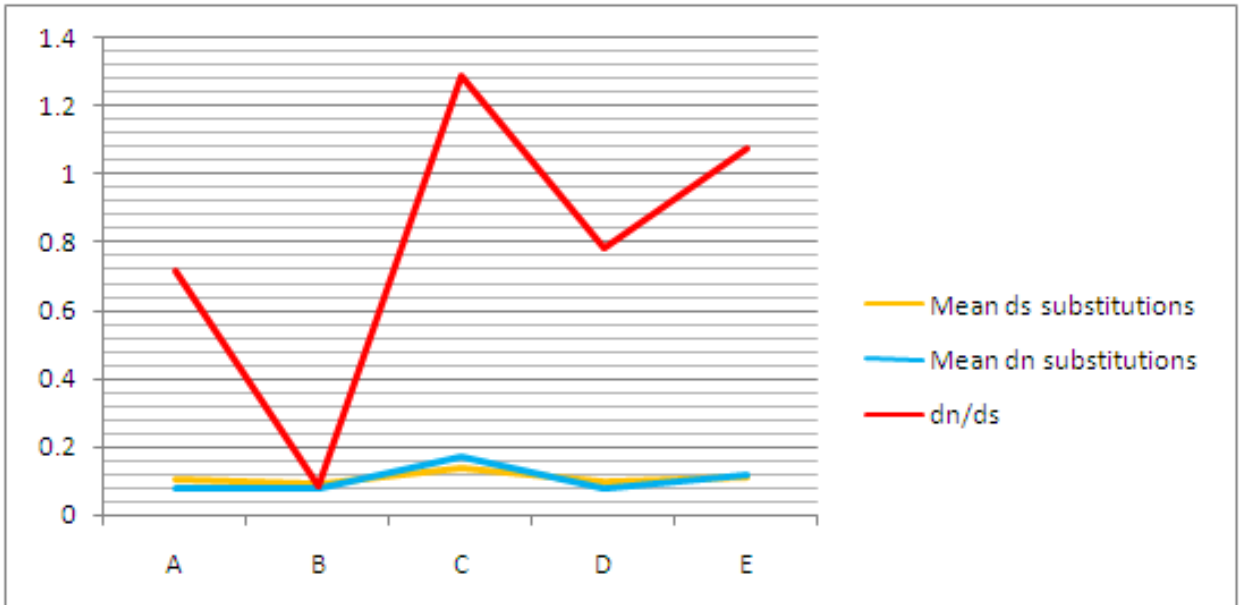


Figure 4.2.4 Overall Mean synonymous (ds) and nonsynonymous (dn) codon substitution rates and the nonsynonymous to synonymous ratio (dn/ds).
Key: A= Recently infected drug naïve; B= Chronically infected drug-naïve; C= Chronically infected drug-experienced; D= Total Drug-naïve; E= Total chronically infected

Another measure of genetic diversity used was Shannon entropy, which indicates position-specific measure of variation in sequence alignments, and which compares two sets of aligned sequences to determine if there is variability in one set (query) relative to the other (background) (Fraham *et al*., 2004). This analysis both at the nucleotide and amino acid level has revealed that entropy level was greater among recently infected drug naïve than chronically infected drug-

naïve; among chronically infected drug experienced than both recently and chronically infected drug naïve groups (Figures 3.5.15A2, B2, C2; and 4.2.5.A and B). This shows that diversity was greatest among chronically infected drug-experienced followed by recently infected drug-naïve participants, with more-or-less stabilized genetic changes among chronically infected drug-naïve patients. This is in complete agreement with a generally accepted assumption that one or a few HIV variants usually initiate infection, and that certain selective forces then drive rapid evolution of viral variants within an infected person, invoke high rate of genetic diversity within the HIV viruses (Learn *et al*., 2002).

Thus, these selective forces, particularly those related to the immune response, may have encouraged emergence of escape variants among recently infected drug naïve group in this study more than among chronically infected drug naïve participants, as viral escape is significantly more rapid (higher selective advantage) in primary infection than in chronic infection (Asquith, 2008; Asquith *et al*., 2006). It has been further shown that viral divergence stabilizes close to disease onset (Lemey *et al*., 2007), which coincides with the stage at which chronically infected drug-naïve patients in this study were found since they were seeking ARV treatment at the time of specimen drawing. Therefore, entropy level and non-synonymous substitution rates were higher and synonymous substitution rates lower among recently infected drug naïve compared to those among chronically infected drug-naïve participants, who showed the reverse of these parameters.

The high entropy level and non-synonymous to synonymous ratio in chronically infected drug-experienced patients at specific sites, largely at the RT region where nearly all drug resistance mutations were developed, showed that drug pressure (positive selection) was responsible for this

intensified genetic diversity in this group of patients. This is clearly visible from the entropy plot in Figures 11.2.2B2, D2, and E2, where statistically significant high entropy levels were localized into nearly five locations in the RT regions: around 130 in the alignment (corresponding to position 41 in the RT), positions 155-170 in the alignment (corresponding to positions 65 up to 75 in the RT), positions 200-220 in the alignment (corresponding to around 90 to 116 in the RT), just before positions 300 in the alignment (around positions 179 to 190 in the RT), and finally at the 3'-end of the alignment where positions 225 and 230 are located. All these locations contained mutations developed in response to ART drug pressure.

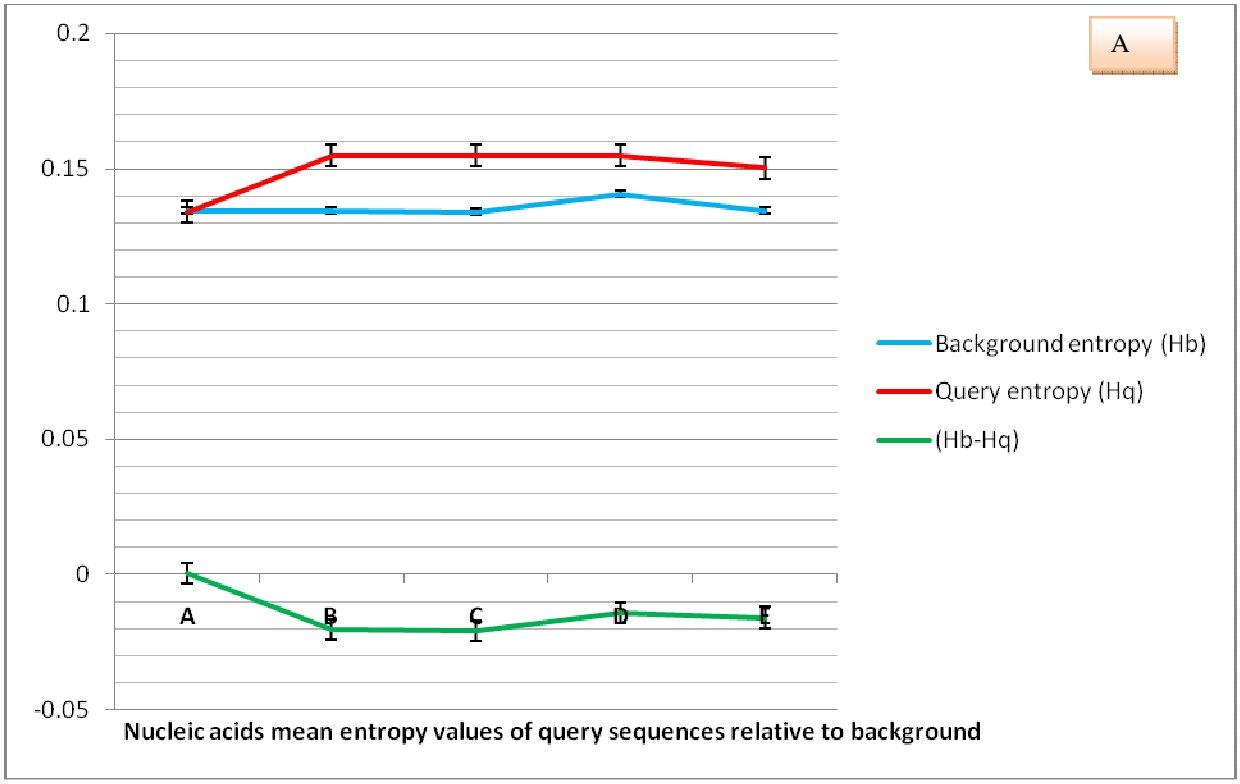


Figure 4.2.5A Statistically significant mean entropy rates
Background (blue), query (red), and differences (green) between the background and query pairs of sequence alignments at nucleotide level. **A,** recently infected drug-naïve Vs chronically infected drug-naïve; **B,** recently infected drug-naïve Vs chronically infected drug-experienced; **C,** chronically infected drug-naïve Vs chronically infected drug-experienced; **D,** all drug-naïve Vs experienced; E, recently infected Vs all chronically infected drug-experienced.

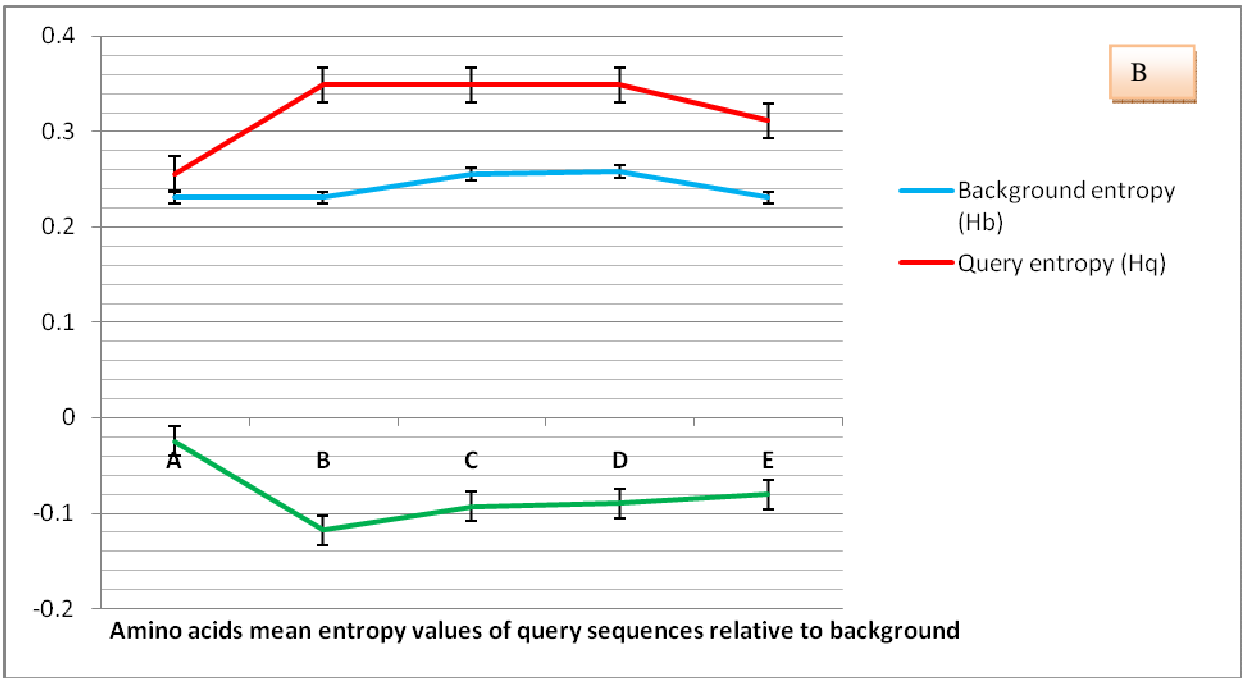


Figure 4.2.5B Statistically significant mean entropy rates
background (blue), query (red), and differences (green) between the background and query pairs of sequence alignments at amino acid level. **A,** recently infected drug-naïve Vs chronically infected drug-naïve; **B,** recently infected drug-naïve Vs chronically infected drug-experienced; **C,** chronically infected drug-naïve Vs chronically infected drug-experienced; **D,** all drug-naïve Vs experienced; E, recently infected Vs all chronically infected drug-experienced.

CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS

This study has addressed five important issues related to HIV infection and HIV isolates in Ethiopia. Data from patients in this study indicated that chronically infected drug-naïve and drug-experienced persons had long duration of sero-conversion, advanced clinical staging, very high viral RNA load, and very low CD4+ T cell count at the time of treatment initiation. These conditions predispose patients to require longer period for virological and immunological treatment gains. In this study, it was observed that 31% chronically infected heavily treated patients had virological failure with mean VL of 289,128 copies/ml, 61% of whom having viral load of WHO's virological failure threshold of $\geq$5000 copies/ml. In addition, the patients were in a poor immunological status with mean CD4+ T cell count of 237 cells/μL and poor clinical conditions with 84% of the patients staying at WHO stages III and IV after median treatment duration of 40 months (over three years). All these observed unfavorable treatment outcomes indicate that treatment starts late for most patients in Ethiopia. Perhaps the recently released WHO's treatment guideline might improve these poor treatment outcomes, which recommends initiating treatment for patients with baseline CD4+ T cell count of $\leq$350 cells/μL irrespective of clinical symptoms, and for all patients in WHO stages III and IV irrespective of CD4+ T cell count.

This study is among few conducted in Ethiopia to understand the drug resistance mutation profiles of HIV isolates from drug-naïve and the first ever from drug-experienced patients. It has produced information with immense practical importance. In the case of treatment-naïve individuals, transmitted HIV drug resistance was found in this study below the 5% threshold for categorizing it as low prevalence during 2005 when ART scaling-up began. However, several issues must be taken into consideration while interpreting this conclusion. The finding in this

study might be an underestimate of the true prevalence of transmitted drug resistance. First, the population sequencing method used in this study may have missed important resistant variants since those minor resistant quasispecies represented by 20% or lower population size could not have been detected in the sequencing. Second, the serum specimen used in the genotyping harbors only those variants currently circulating not those variant which originally infected the persons; variants with original resistance genetic background might be found in other sanctuary sites such as CNS, lymph nodes, and PBMC proviruses. Third, extrapolating this low prevalence rate to the whole country is difficult since cities/towns may have had different history of ART exposure prior to the 2005 launching of HAART scale-up. Moreover, the fact that genotyping was done five years before and that prevalence of treatment-driven resistance is currently increasing indicate that the low transmitted resistance prevalence might not be the reality today.

This therefore calls for establishment of regular surveillance system in Ethiopia to monitor prevalence of transmitted drug resistance. In view of the issues raised above, such future surveillance attempts should diversify target populations and target sampling sites, by including more individuals from private health service institutions and more sites from at least regional capital cities. Although allele-specific PCR, single genome sequencing, pyro-sequencing (ultra deep sequencing), and other more sensitive methods of detection of minority resistant variants cannot (at the moment) replace the conventional population sequencing because of the technical difficulty and financial constraints, future surveillance of transmitted drug resistance and baseline drug resistance testing may benefit from sequencing the proviruses from PBMC since archived resistant variants missed from plasma/serum could be detected from the proviruses. This only needs same primers, same amplification and sequencing protocols, except omission of the RT-PCR part of the procedure.

The other important finding in this study is the detection of eight PI- (L23I [3.53%], L33F [3.53%], M46L [3.53%], G48V [1.38%], I54V [3.53%], G73A [1.38%], I84C [1.38%], and L90M [3.53%]), eight NRTI- (M41L [3.53%], D67N [5.88%], L74V [3.53%], M184V/I [2.35%], L210W [3.53%], and T215Y [3.53%]), and six NNRTI-resistance mutations (A98G [1.38%], L100I [3.53%], K103N [4.71%], Y181C [3.53%], and Y188C/H [2.35%]) at baseline among patients who were initiating ART. The worrying part of this finding is that three (3.5%) participants already harbored type-1 Thymidine Analog Mutations (TAMs), which confer higher level reduction in clinical response and cross resistance to ABC, ddI, and TDF. In addition, double class resistance mutations against NRTIs and NNRTIs were prevalent in 2.4% of the patients, which is of concern since these two drug classes are the only ARV drugs available for initial treatment in Ethiopia, and no drug resistance testing is done before starting ART.

Regarding drug resistance mutations among treatment-experienced patients, prevalence rates of resistance mutations were observed 70% against any class of drugs, 62% against NRTIs, 68% against NNRTIs, and 61% double class NRTI/NNRTI resistance mutations. Prevalence of Thymidine analog resistance mutations (both type I and II) singly (up to 27%), in pairs (up to 41%) and all type I simultaneously (6.1%), along with non-thymidine analog resistance mutations like M184V/I, K65R and K70R are likely to narrow options for future use of NRTI drugs because these mutations are selected by all NRTI drugs currently used as component parts of both 1st line and 2nd line regimens in the country. Points that deserve special attention concerning NRTI resistance mutations are that i) M184V was the most abundant NRTI resistance mutation, which is selected largely by 3TC (This drug is the only component part of all ART combinations administered in Ethiopia both in the 1st and 2nd line regimens.); ii) disregarding other potential

confounding factors, more D4T-containing regimens than AZT-containing regimens were found to have selected for NRTI resistance mutations in the setting of this study.

The most prevalent NNRTI resistance mutations in this study (K103N, Y181C/I, V106A/M, and L190S/A) were those that are naturally selected by the two NNRTI drugs administered in Ethiopia (NVP and EFV). Unlike observations from other countries and in spite of NVP's use in Ethiopia for prevention of mother to child transmission, EFV was found in this study to have had slightly greater selection power than NVP for NNRI resistance mutations; about 55% of NVP-exposed isolates to the initial therapy regimens and 97% isolates exposed to regimens used at the time of genotyping harbored resistance mutations against NVP, while 96% of EFV-exposed isolates to the initial therapy regimens and 100% genotyped isolates which were exposed to EFV regimens at the time of genotyping harbored resistance mutations against EFV. The reason could be that more patients on EFV-containing regimens may have missed their drugs because of the toxicity problem associated with this drug. In any case, some issues about NVP- and EFV-associated mutations observed in this study deserve further investigations to elucidate with certainty reasons for their occurrences: i) for unknown reason(s) some NNRTI mutations were selected by either NVP or EFV when they should have been selected by both; ii) the mutation P225H was found in this study among isolates selected by NVP (5%) although supposed to be selected only by EFV.

The occurrence of PI resistance mutations among all study participants (~5% among recently infected drug-naïve, ~4% among chronically infected drug-naïve, and 3% among drug-experienced) with remote possibility of prior exposure to PI drugs (except in one case among those drug-experienced patients) might indicate that the genetic background of HIV isolates in

Ethiopia could favor emergence of PI resistance mutations upon widespread use of these drugs in the future. Moreover, detection of several mutations, which are considered to confer no drug resistance, on drug resistance positions require that they be closely investigated since absence of their involvement in drug resistance was tested (if at all) on other isolates under different genetic background (of both patients and viral isolates).

Taken together, NRTI and NNRTI drug resistance mutation rates observed in this study are comparable with those reported from regions of the world (like Europe and North America), where ART has been in use for longer time than in Ethiopia. Several reasons could be behind emergence of these high prevalence rates in Ethiopia. Among the possible causes is unavailability of tests such as viral load and drug resistance assays that would directly monitor treatment outcomes and hence guide the physicians on how to make individually-tailored treatment adjustments required for each patient. In addition, considering already known interpersonal and even sometimes intrapersonal differences in absorbing ARV drugs by patients, and given that the pharmacodynamics of the currently used ARV drugs in Ethiopia have not been examined under the settings of Ethiopian patients, suboptimal treatment could be one of the possible reasons for emergence of high resistance mutation rate within shorter time since the beginning of ART scale-up. Poor adherence because of toxicity-associated drug-holydays might have also contributed to suboptimal treatment, and consequently to emergence of high rates of drug resistance mutations.

In general, the treatment-driven resistance mutations observed among heavily treated individuals, together with observations from baseline and transmitted drug resistance mutations among chronically infected drug-naïve and recently infected drug-naïve persons, respectively, who participated in this study signify that: i) other treatment options (such as use of integrase

inhibitors and fusion inhibitors) that supplement PIs, NRTI and NNRTIs may need to be considered for use in the near-future; ii) genotypic drug resistance testing should be considered as part-and-parcel of HIV treatment and care in the near-future, and therefore efforts should be made to conduct the testing just before initiation of treatment, if possible, and whenever treatment failure is detected; iii) phenotypic drug resistance assessment should be conducted on Ethiopian isolates, at least within research facilities (if not for routine clinical use) in order to unequivocally conclude that mutations from Ethiopian isolates occurring at resistance positions hitherto not known to confer resistance are really non-resistance; iv) therapeutic drug monitoring (TDM) studies be conducted to establish exposure-response relationship of each drug in the Ethiopian setting among various target population under ARV treatment.

Unfortunately, with very few trained manpower available currently, only one sequencing machine at EHNRI, and over a million HIV infected patients who potentially need this service, this task would be formidably huge. It is therefore the right time now for health providing government and private institutions, and health policy makers to make the necessary preparations in terms of manpower development and acquisition of more sequencing machines to begin delivering this service. Of course, several potentially hindering factors may threaten establishment of the necessary resistance testing facilities and implementation of the testing, two of which being availability of easy-to-use and relatively inexpensive genotyping system and storage-friendly specimen source. These two issues have been successfully addressed in this study, and very promising results were produced. Regarding easy-to-use and relatively inexpensive genotyping system, the in-house genotyping system tested in this study proved to work more than satisfactorily, with mean concordance with the commercial ViroSeq™ genotyping system of ~99% at the nucleotide level. With mean relative sensitivity of 94%, and

relative predictive value of ~99% at the amino acid level under the Ethiopian setting, it is an ideal genotyping system which can be used in the routine drug resistance testing instead of the more expensive and technically more demanding ViroSeq^{TM}. Its proven high sensitivity in this study to amplify viral RNA from as low as 500 to > 6 million copies/ml will make this genotyping system promising for use when drug resistance testing is expanded in Ethiopia. Parallel to working with this already tested system, however, optimization researches have to continue to further expand regions in the viral RNA that would be amplified and sequenced beginning from the end of *gag* to end of *pol* (including the integrase region), whereby the potential targets for most of the currently used ARV drugs would be addressed.

Use of DBS which was tested in this study and elsewhere as a source of specimen for drug resistance genotyping will also prove very important when drug resistance testing becomes a necessity. Evaluation of the performance of DBS relative to that of the conventional plasma specimens in this study has shown a 96% concordance at the nucleotide level, and relative sensitivity of 97% and positive predictive value of 99% at the amino acid level. With these excellent performances, its low cost, low demand for stringent storage conditions, and convenience for transportation under ambient temperatures, it is appropriate to expand the usage of DBS as source of specimen under the Ethiopian settings. In addition, its usefulness in testing drug resistance can be extended to treatment monitoring of pediatric patients, from whom repeated and voluminous blood drawing for drug resistance testing and other treatment monitoring laboratory tests are difficult to perform. In fact, since previous studies on DBS's usefulness in CD4 count and viral load testing provided good results, additional application of DBS in monitoring drug resistance mutations will make it a multipurpose specimen.

With regards to subtype distribution among the study participants, ~96% of the sequences tested were of subtype C both at PR and RT genes, showing that subtype C is still the dominant clade in Ethiopia even after two decades of the epidemic in the country. With sequences from only two individuals (1%) identified as subtype B, the remaining 3% were subtyped at least by one genotyping algorithm as recombinant forms at the PR and RT genes: four (2%) BCs (not the BCs previously identified as CRFs since their mosaic genetic sequences of the latter are at other genes unlike at the PR and RT genes in this study), one (0.5%) CRF_02AG (or BG? or ABG?), and one (0.5%) A1D. Worth noting in this subtype assignment is that the discrepancies between the five algorithms in assigning the mosaic sequences to specific clades were so mixed-up that it was difficult to put them under previously known subtypes, and so is suggested here to recognize them as unique recombinant forms (URFs) rather than providing them with specific subtype. However, the existence of non-C subtypes in the neighboring countries and some previous reports from Ethiopia, together with the currently available high influx of people (both Ethiopians and expatriates) as tourists and for business purposes coupled with the flourishing sex trade in the capital, should entail the introduction of more non-C subtypes into the country. Therefore, molecular surveillance should continue to track subtype distribution of HIV isolates in the country. But this should be done on larger genomic regions than just the short *pol* region done in this study or *env* region done in many previous studies. The study sites should also incorporate areas of the country that share borders with the neighboring countries.

Besides detecting some unique mosaic forms, this study has also recognized different patterns of evolutionary changes in the sequenced *pol* region between isolates from acutely infected (recent infection), chronically infected needing ART treatment, and chronically infected patients who have been on ART treatment for long time. While some mutations (such as M36I, H69K, I93L,

I15V, and R41K from PR; and K122E, T200A, D177E, and V245Q from RT) were found to have high prevalence among all the three groups, others have shown notable differences in mutation prevalence between the three groups. The latter types of mutations included K14R, G16E, E35D, and L89M from PR, which progressively decreased in prevalence from the highest in acutely infected drug-naive, and chronically infected initiating ART, to the lowest in chronically infected heavily treated patients; and conversely PR mutations like T12S and T74S progressively decreased from the highest in chronically infected treatment-experienced patients, to the lowest acutely infected drug-naïve; the chronically infected ART initiating patients had prevalence rate of these mutations in between. PR mutations like I13V, L19I/T, and L63P were prevalent among isolates from advanced non-treated individuals than they were among acutely infected rug-naïve or chronically infected treated persons. Similarly, mutations of the RT such as E36A, T39D, S48T, K70G, K173A, and F214L appeared most frequently among isolates from treated individuals followed by isolates from non-treated patients with advanced infection, and least from acutely infected. Some RT mutations such as K49R, K102Q, D123G, and A158S were prevalent more among acutely infected than among advanced infections, although very much reduced among treated individuals.

Overall, Jukes-Cantor method of diversity analysis, synonymous/non-synonymous analysis, and test of Shannon Entropy all detected highest rate of diversity among isolates from chronically infected patients who were on treatment followed by recently infected patients, with the least diversity among chronically infected patients needing treatment initiation. This observation is in a total agreement with the concept that viral isolates from individuals in advanced disease (like those in this study who were chronically infected ART initiating patients), because the majority of quasispecies sprouted out during acute infection became eliminated due to various selective

pressures exerted while the long disease course was going on. Whereas diversity was enhanced among isolates from recently infected patients due to immune pressure from the host, intensified diversity among isolates from chronically infected patients who were on treatment was due to the drug selective pressure, since most of the diversities observed were on positions surrounding drug resistance sites.

Limitations of the study

1. Population sequencing was used in this study, which fails to detect and quantify minorities of drug-resistant quasi-species represented below 20%. So resistance variants which were represented in this study by a small proportion of the population might have been missed. Thus, caution is required during interpretations of the drug resistance mutaions from this study
2. The source of sequence data was the short *pol* region which codes only for whole PR and the first 250 amino acids in the RT. The data generated from this short sequence will have limited information regarding subtype profile of the isolates. It would be of great benefit if other regions such as *env, gag,* and *LTR* were sequenced, if not the whole genome.
3. Some of the data were generated long time ago, and hence, may not correctly reflect the current situation.
4. Some of the specimens used were of poor quality due either to storage, processing, or transportation problems; therefore, it was difficult to repeat some ambiguous results.
5. Contamination during amplification and sequencing occurred in some occasions that required omission of sequencing results; therefore, although all the necessary precautions were made, the possibility that some sequence results in this study might contain results from contaminant may not be nil.

List of References

Abebe A, Kuiken CL, Goudsmit J, Valk MV, Messele T, Sahlu T, Yeneneh H, Fontanet A, De Wolf F, and De Wit TFR. 1997. HIV Type 1 Subtype C in Addis Ababa, Ethiopia. *AIDS Research and Human Retroviruses*. **13(12):**1071-1075.

Abebe A, Demissie D, Goudsmit J, Brouwer M, Kuiken CL, Pollakis G, Schuitemaker H, Fontanet AL, and Rinke de Wit TF. 1999. HIV-1 Subtype C syncytium- and Nonsyncytium-Inducing Phenotypes and Coreceptor Usage among Ethiopian Patients with AIDS. *AIDS*. **13(11):**1305–1311.

Abebe A, Lukashova VV, Pollakisa G, Kliphuisa A, Fontanetb AL, Goudsmita J, and de Wit TFR. 2001. Timing of the HIV-1 Subtype C Epidemic in Ethiopia Based on Early Virus Strains and Subsequent Virus Diversification. *AIDS*. **15:**1555-1561.

Abebe A, Pollakis G, Fontanet AL, Fisseha B, Tegbaru B, Kliphuis A. Kliphuis A, Tesfaye G, Negassa H, Cornelissen M, Goudsmit J, and de Wit TFR. 2000. Identification of a Genetic Subcluster of HIV-1 Subtype C (C') Widespread in Ethiopia. *AIDS Research and Human Retrovirolgy*. **16:** 1909-1914.

Abecasis AB, Deforche K, Bacheler LT, McKenna P, Carvalho AP, Gomes P, Vandamme AM, and Camacho RJ. 2006. Investigation of Baseline Susceptibility to Protease Inhibitors in HIV-1 Subtypes C, F, G and CRF02_AG. *Antiviral. Therapy.* **11:**581-589.

Abecasis AB, Lemey P, Vidal N, de Oliveira T, Peeters M, Camacho R, Shapiro B, Rambaut A, and Vandamm AM. 2007. Recombination Confounds the Early Evolutionary History of Human Immunodeficiency Virus Type 1: Subtype G is a Circulating Recombinant Form. *Journal of Virology*. **81:**8543–8551.

Abraha A, Nankya IL, Gibson R, Demers K, Tebit DM, Johnston E, Katzenstein D, Siddiqui A, Herrera C, Fischetti L, Shattock RJ, and Arts EJ. 2009. CCR5- and CXCR4-Tropic Subtype C Human Immunodeficiency Virus Type 1 Isolates Have a Lower Level of Pathogenic Fitness than Other Dominant Group M Subtypes: Implications for the Epidemic. *Journal of Virology*. **83(11):** 5592-5605.

Adal M, Ayele W, Wolday D, Dagne K, Messele T, Tilahun T, Berkhout B, Mayaan S, Pollakis G, and Dorigo-Zetsma W. 2005. Evidence of Genetic Variability of Human Immunodeficiency Virus Type 1 in Plasma and Cervicovaginal Lavage in Ethiopian Women

Seeking Care for Sexually Transmitted Infections. *AIDS Research and Human Retroviruses*. **21(7):**649-653.

Adje-Toure C, Celestin B, Hanson D, Roels TH, Hertogs K, Larder B, Diomande F, Peeters M, Eholie S, Lackritz E, Chorba T, and Nkengasong JN. 2003. Prevalence of Genotypic and Phenotypic HIV-1 Drug-Resistant Strains among Patients Who Have Rebound in Viral Load while Receiving Antiretroviral Therapy in the UNAIDS-Drug Access Initiative in Abidjan, Cote d'Ivoire. *AIDS*.**17(Suppl 3):**S23-S29.

Ait-Khaled M, Rakik A, Griffin P, Stone C, Richards N, Thomas D, Falloo J, Tisdale M, CNA2007 International Study Team. 2003. HIV-1 Reverse Transcriptase and Protease Resistance Mutations Selected During 16-72 Weeks of Therapy in Isolates from Antiretroviral Therapy-Experienced Patients Receiving Abacavir/Efavirenz/Amprenavir in the CNA2007 study. *Antiviral Therapy*. **8(2):**111-120.

Alaeus A, Lidman K, Bjorkman A, Giesecke J, and Albert J. 1999. Similar Rate of Disease Progression among Individuals Infected With HIV-1 Genetic Subtypes A-D. *AIDS*. **13:**901–907.

Alexander CS, Dong W, Schechter MT, O'Shaughnessy MV, Strathdee SA, Mo T, Montaner JS, and Harrigan PR. 1999. Prevalence of Primary HIV Drug Resistance among Seroconverters During an Explosive Outbreak of HIV Infection among Injecting Drug Users. *AIDS*. **13**:981-985.

Al-Hashimi HM. 2009. Aerial View of the HIV Genome. *Nature*. **460:**696-698.

Allers K, Hutter G, Hofmann J, Loddenkemper C, Rieger K, Thiel E, and Schneider T. 2010. Evidence for the Cure of HIV Infection by CCR5Δ32/Δ32 Stem Cell Transplantation. *Blood*. [e-pub ahead of print. Available at: http://dx.doi.org/10.1182/blood-2010-09-309591 accessed on January 13, 2011]. doi:10.1182/blood-2010-09-309591.

Altfeld M, Addo MM, Rosenberg ES, Hecht FM, Lee PK, Vogel M, Yu XG, Draenert R, Johnston MN, Strick D, Allen TM, Feeney ME, Kahn JO, Sekaly RP, Levy JA, Rockstroh JK, Goulder PJR, and Walker BD.2003. Influence of HLA-B57 on Clinical Presentation and Viral Control During Acute HIV-1 Infection. *AIDS*. **17(18):**2581-2591

Altfeld M, and Walker BD. 2007. Acute HIV-1 Infection. In: *HIV Medicine*. 15th Edition. Christian Hoffman, and Jurgen K. Rockstroh (Editors). Flying Publisher, Paris, 2007. Pp 33-39. Available at www.HIVMedicine.com

Altfeld M, Kalife ET, Qi Y, Streec H, Lichterfeld M, Johnston MN, Burget N, Swartz ME, Yang A, Alte G, Yu XG, Meier A, Rockstroh JK, Allen TM, Jessen H, Rosenber ES, Carrington M, and Walker BD. 2006. HLA Alleles Associated with Delayed Progression to AIDS Contribute Strongly to the Initial CD8(+) T Cell Response against HIV-1. *PLoS Medicine*. **3(10):**e403. http://amedeo.com/lit.php?id= 17076553

Anderson JP, Rodrigo AG, Learn GH, Madan A, Delahunty C, Coon M, Girard M, Osmanov S, Hood L, and Mullins JI. 2000. Testing the Hypothesis of a Recombinant Origin of Human Immunodeficiency Virus Type 1 Subtype E. *Journal of Virology*. **74:**10752–10765.

Andreasson PA, Dias F, Naucler A, Andersson S, and Biberfeld G. 1993. A Prospective Study of Vertical Transmission of HIV-2 in Bissau, Guinea-Bissau. *AIDS*. **7:**989–993.

Arien KK, Abraha A, Quinones-Mateu ME, Kestens L, Vanham G, and Arts EJ. 2005. The Replicative Fitness of Primary Human Immunodeficiency Virus Type 1 (HIV-1) group M, HIV-1 Group O, and HIV-2 Isolates. *The Journal of Virology*. **79:**8979–8990.

Asquith B, Edwards C, Lipsitch M, and McLean AR. 2006. Inefficient CTL-mediated Killing of HIV-1-infected Cells *in vivo*. *PLoS Biology*. **4:**0583–0592.

Asquith B. 2008. The Evolutionary Selective Advantage of HIV-1 Escape Variants and the Contribution of Escape to the HLA-Associated Risk of AIDS Progression. *PLoS ONE.* **3(10):** e3486. doi:10.1371/journal.pone.0003486.

Assefa T, Davey G, Dukers N, Wolday D, Worku A, Messele T, Tegbaru B, Dorigo W, Sanders EJ. 2003. Overall HIV-1 Prevalence in Pregnant Women Over-Estimates HIV-1 in the Predominantly Rural Population of Afar Region. *Ethiopian Medical Journal*. **41(Suppl 1):**43-49.

Ayehunie S, Johansson B, Sönnerborg A, Zewdie DW, Britton S, and Strannegård O. 1993. Sequence Analysis of Selected Regions of the env (V3 loop and gp41) and gag (p7) Reading Frames of Ethiopian Human Immunodeficiency Virus Type 1 Strains. *Archives of Virology.* **128(3):**229-239.

Ayehunie S, Sönnerborg A, Johansson B, Fehniger TE, Zewdie DW, Yeamane-Berhan T, Petros B, Abens A, Britton S, and Strannegård O. 1990. Differences in PCR Reactivity between HIV Proviruses from Individuals in Ethiopia and Sweden. *Journal of Acquired Immunodeficiency Syndrome*. **3:**975-980.

Ayele W, de Baar MP, Goudsmit J, Kliphuis A, Tilahun T, Dorigo-Zetsma W, Wolday D, Abebe A, Mengistu Y, Pollakis G. 2005. Surveillance Technology for HIV-1 Subtype C in Ethiopia: An *env*-based NASBA Molecular Beacon Assay to Discriminate between Subcluster C and C'. *Journal of Virological Methods*. **130:**22–29.

Ayele W, Pollakis G, Abebe A, Fisseha B, Tegbaru B, Tesfaye G, Mengistu Y, Wolday D, van Gemen DB, Goudsmit J, Dorigo-Zetsma W, and de Baar MP. 2004. Development of a Nucleic Acid Sequence-Based Amplification Assay that Uses *gag*-Based Molecular Beacons to Distinguish between Human Immunodeficiency Virus Type 1 Subtype C and C' Infections in Ethiopia. *Journal of Clinical Microbiology*. **42(4):**1534–1541.

Ayele W, Schuurman R, Messele T, Dorigo-Zetsma W, Mengistu Y, Goudsmit J,. Paxton WA, de Baar MP, and Pollakis G. 2007. Use of Dried Spots of Whole Blood, Plasma, and Mother's Milk Collected on Filter Paper for Measurement of Human Immunodeficiency Virus Type 1 Burden. *Journal of Clinical Microbiology*. **45(3):**891-896.

Badiger SB, Thapar R, Mithra P, Kumar SG, Jain A, Bhaskaran U, Subramanya J. 2010. A Profile of Patients Attending an Anti Retroviral Therapy (ART) Centre at a Tertiary Care Hospital in South India. *Australasian Medical Journal.* **3(6):**344-348.

Baeten JM, Chohan B, Lavereys L, Lavreys L, Chohan V, Scott McClelland RS, Certain L, Mandaliya K, Walter Jaoko W, and 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. *The Journal of Infectious Diseases*. **195(8):**1177-1180.

Bailes E, Gao F, Bibollet-Ruche F, Courgnaud, V, Peeters M, Marx PA, Hahn BH, and Sharp BM. 2003. Hybrid Origin of SIV in Chimpanzees. *Science*. **300:**1713.

Baldrich-Rubio E, Anagonou S, Stirrups K, Lafia E, Candotti D, Lee H and Allain JP. 2001. A Complex Human Immunodeficiency Virus Type 1 A/G/J Recombinant Virus Isolated from a Seronegative Patient with AIDS from Benin, West Africa. *Journal of General Virology*. **82:** 1095–1106.

Ball SC, Abraha A, Collins KR, Marozsan AJ, Baird H, Quinones-Mateu ME, Penn-Nicholson A, Murray M, Richard N, Lobritz M, Zimmerman PA, Kawamura T, Blauvelt A, and Arts EJ. 2003. Comparing the Ex Vivo Fitness of CCR5-Tropic Human Immunodeficiency Virus Type 1 Isolates of Subtypes B and C. *Journal of Virology*. **77(2):** 1021–1038.

Bansode V, Drebert ZJ, Travers SAA, Banda E, Molesworth A, Crampin A, Ngwira B, Glynn JR, and McCormack GP. 2010. Drug Resistance Mutations in Drug-Naïve HIV-1 Type-1 Subtype C-Infected Individuals from Rural Malawi. *AIDS Research and Human Retroviruses*. **26:** In Print DOI: 10.1089/aid.2010.0203

Barbour JD, Wrin T, Grant RM, Martin JN, Segal MR, Petropoulos CJ, and Deeks SG. 2002. Evolution of Phenotypic Drug Susceptibility and Viral Replication Capacity During Long-Term Virologic Failure of Protease Inhibitor Therapy in Human Immunodeficiency Virus-Infected Adults. *Journal of Virology.* **76(21):**11104-11112.

Bar-Magen T, Donahue DA, McDonough EI, Kuhl BD, Faltenbache rVH, Xu H, Michaud V, Sloan RD and Wainberg MA. 2010. HIV-1 Subtype B and C Integrase Enzymes Exhibit Differential Patterns of Resistance to Integrase Inhibitors in Biochemical Assays *AIDS*. **24:** DOI:10.1097/QAD.0b013e32833cf265.

Barouch DH, and Korber B. 2010. HIV-1 Vaccine Development after STEP. *Annual Review of Medicine*. **61:**153–167.

Barre-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vezinet- Brun F, Rouzioux C, Rozenbaum W, and Montagnier L. 1983. Isolation of a T-lymphotropic Retrovirus from a Patient at Risk for Acquired Immune Deficiency Syndrome (AIDS). *Science*. **220:** 868–871.

Basavapathruni A, Karen S., and Anderson K. 2007. Reverse Transcription of the HIV-1 Pandemic. *The FASEB Journal*. **21:**3795-3808.

Batra M, Tien PC, Shafer RW, Contag CH, and Katzenstein DA. 2000. HIV Type 1 Envelope Subtype C Sequences from Recent Seroconverters in Zimbabwe. *AIDS Research and Human Retroviruses*. **16:**973–979.

Becquart P, Chomont N, Roques P, Ayouba A, Kazatchkine MD, Belec L, and Hocini H. 2002. Compartmentalization of HIV-1 between Breast Milk and Blood of HIV-Infected Mothers. *Virology*. **300:**109-117.

Beddows S, Galpin S, Kazmi SH, Ashraf A, Johargy A, Frater AJ, White N, Braganza R, Clarke J, McClure M, and Weber JN. 2003. Performance of Two Commercially Available Sequence-Based HIV-1 Genotyping Systems for the Detection of Drug Resistance against HIV Type 1 group M Subtypes. *Journal of Medical Virology*. **70:**337–342.

Bello G, Aulicino PC, Ruchansky D, Guimarães ML, Lopez-Galindez C, Casado C, Chiparelli H, Rocco C, Mangano A, Sen L, and Morgado MG. 2010. Phylodynamics of HIV-1 Circulating Recombinant Forms 12_BF and 38_BF in Argentina and Uruguay. *Retrovirology*. **7:**22 http://www.retrovirology.com/content/7/1/22.

Bennett DE, Camacho RJ, Otelea D, Kuritzkes DR, Fleury H, Kiuchi M, Heneine W, Kantor R, Jordan MR, Schapiro JM, Vandamme AM, Sandstrom P, Boucher CAB, van de Vijver D, Rhee SY, Liu TF, Pillay D, and Shafer RW. 2009. Drug Resistance Mutations for Surveillance of Transmitted HIV-1 Drug-Resistance: 2009 Update. *PLoS ONE* **4(3):** e4724. doi:10.1371/journal.pone.0004724.

Bennett DE. 2006. The Requirment for Surveillance of HIV Drug Resistance within Antiretroviral Rollout in the Developing World. *Current Opinion in Infectious Diseases*. **19:**607-614.

Bergamaschi A, and Pancino G. 2010. Host Hindrance to HIV-1 Replication in Monocytes and Macrophages. *Retrovirology* **7:**31. http://www.retrovirology.com/content/7/1/31.

Bergroth T, Ekici H, Gisslen M, Hagberg L, and Sonnerborg A. 2009. Difference in Drug Resistance Patterns between Minor HIV-1 Populations in Cerebrospinal Fluid and Plasma. *HIV Medicine*. **10:**111-115.

Bertagnolio S, and Sutherland D. 2005. WHO Approach to Track HIV Drug Resistance Emergence and Transmission in Countries Scaling up HIV Treatment. *AIDS*. **19:**1329-1330.

Bertagnolio S, Soto-Ramirez L, Pilon R, Rodriguez R, Viveros M, Fuentes L, Harrigan PR, Mo T, Sutherland D, Sandstrom P. 2007. HIV-1 Drug Resistance Surveillance Using Dried Whole Blood Spots. *Antiviral Therapy*. **12(1):**107-113.

Bertagnolio S, Soto-Ramirez L, Pilon R, Rodriguez R, Viveros M, Fuentes L, Harrigan PR, Mo T, Sutherland D, and Sandstrom P. 2007. HIV-1 Drug Resistance Surveillance Using Dried Whole Blood Spots. *Antiviral Therapy*. **12(1):**107-113.

Biti R, Ffrench R, Young J, Bennetts B, Stewart G, and Liang T. 1997. HIV-1 Infection in an Individual Homozygous for the CCR5 Deletion Allele. Nauret Medicine. **3:**252-253.

Blackard JT, Cohen DE, and Mayer KH. 2002. Human Immunodeficiency Virus Superinfection and Recombination: Current State of Knowledge and Potential Clinical Consequences. *Clinical Infectious Diseases.* **34:**1108-1114.

Blacklaws BA, and Harkiss GD. 2010. Small Ruminant Lentiviruses and Human Immunodeficiency Virus: Cousins that Take a Long View. *Current HIV Research*. **8(1):**26-52.

Blower SM, Aschenbach AN, Gershengorn HB, and Kahn JO. 2001. Predicting the unpredictable: Transmission of drug-resistant HIV. *Nature Medcine*. **7(9):**1016-120.

Boom R, Sol CJ, Salimans MM, Jansen CL, van Dillen PMEW, Noordaa V. 1990. Rapid and Simple Method for Purification of Nucleic Acids. *Journal of Clinical Microbiology*. **28:**495-503.

Boucheri S, Recordon-Pinson P, Ragnaud JM, Dupon M, Fleury H, and Masquelier B. 2006. HIV-1 Reverse Transcriptase (RT) Genotypic Patterns and Treatment Characteristics Associated with the K65R RT Mutation. *HIV Medicine*. **7:**294-298.

Bousheri S, Burke C, Sewanyana I, Harrigan R, Martin J, Hunt P, Bangsberg DR, and Cao H. 2009. Infection With Different HIV Subtypes Is Associated with CD4 Activation-Associated Dysfunction and Apoptosis. *Journal of Acquired Immune Deficiency Syndrome*. **52:**548–552.

Brennan CA, Bodelle P, Coffey R, Harris B, Holzmayer V, Luk KC, Swanson P, Yamaguchi J, Vallari A, Devare SG, Schochetman G, and Hackett Jr. J. 2006. HIV Global Surveillance: Foundation for Retroviral Discovery and Assay Development. *Journal of Medical Microbiology*. **78:**S24–S29.

Brenner B, Routy JP, Quan Y, Moisi D, Oliveira M, Turner D, Wainberg MA, and co-investigators of the Quebec Primary Infection Study. 2004. Persistence of Multidrug Resistant HIV-1 in Primary Infection Leading to Superinfection. *AIDS*. **18:**1653-1660.

Brenner B, Wainberg MA, Salomon H, Rouleau D, Dascal A, Spira B, Sekaly RP, Conway B, and Routy JP, Investigators of the Quebec Primary Infection Study. 2000. Resistance to Antiretroviral Drugs in Patients with Primary HIV-1 Infection. *International Journal of Antimicrobial Agents*. **16:**429-434.

Brookfield JFY, Sharp PM, Shaw GM, Peeters M, and Hahn BH. 2006. Chimpanzee Reservoirs of Pandemic and Nonpandemic HIV-1. *Science*. **313(5786):** 523–526.

Brown BK, Wieczorek L, Sanders-Buell E, Borges AR, Robb ML, Birx DL, Michael NL, McCutchan FE, Victoria R., and Polonis VR. 2008. Cross-Clade Neutralization Patterns

among HIV-1 Strains from the Six Major Clades of the Pandemic Evaluated and Compared in two Different Models. *Virology*.; **375:**529–538.

Brown F. 1986. Human Immunodeficiency Virus. *Science*. 232: 1486.

Buckton AJ, 2008. New Methods for the Surveillance of Drug resistance in the Resource Poor World. *Current Opinion in Infectious Diseases*. **21:**653-658.

Buckton AJ, Bissett SL, Myers RE, Beddows S, Edwards S, Cane PA, and Pillay D. 2008. Development and Optimization of an Internally Controlled Dried Blood Spot Assay for Surveillance of Human Immmunodeficiency Virus Type-1 Drug Resistance. *Journal of Antimicrobial Chemotherapy*. **62:**1191-1198.

Buckton AJ, Harris RJ, Pillay D, and Cane PA. 2010. HIV-Type-1 Drug Resistance in Treatment –Naïve Patients Monitored Using Minority Species Assay: a Systematic Review and Meta-Analysis. *Antiviral Therapy*. **16:** In Press (doi:10.3851/IMP1687).

Burgard M, Izopet J, Dumon B, Tamalet C, Descamps D, Ruffault A, Vabret A, Bargues G, Mouroux M, Pellegrin I, Ivanoff S, Guisthau O, Calvez V, Seigneurin JM, and Rouzioux C. 2000. HIV RNA and HIV DNA in Peripheral Blood Mononuclear Cells Are Consistent Markers for Estimating Viral Load in Patients Undergoing Long-Term Potent Treatment. *AIDS Research and Hunam Retroviruses*.**16:**1939-1947.

Burton GF, Keele BF, Estes JD, Thacker TC, and Gartner S. 2002. Follicular Dendritic Cell Contributions to HIV Pathogenesis. *Seminar in Immunology.* **14(4):**275–284.

Campbell GR, Loret EP, and Spector SA. 2010. HIV-1 Clade B Tat, but not Clade C Tat, Increases X4 HIV-1 Entry into Resting but not Activated CD4 T Cells. *Journal of Biological Chemistry*. **285:**1681–1691.

Campbell TB. 2006. Are all HIV Type-1 Strains Created Equal? *Clinical Infectious Diseases*. **42:**853-854.

Canducci F, Sampaolo M, Marinozzi MC, Boeri E, Spagnuolo V, Galli A, Castagna A, Lazzarin A, Clementi M, and Gianotti N. 2009. Dynamic Patterns of Human Immunodeficiency Virus Type 1 Integrase Gene Evolution in Patients Failing Raltegravir-Based Salvage Therapies. *AIDS*. **23:**455–460.

Cao H, Mani I, Vincent R, Mugerwa R, Mugyenyi P, Kanki P, Ellner J, and Walker BD. 2000. Cellular Immunity to Human Immunodeficiency Virus Type 1 (HIV-1) Clades: Relevance to HIV-1 Vaccine Trials in Uganda. *The Journal of Infectious Diseases*. **182:**1350–1356.

Carr JK, Salminen MO, Koch C, Gotte D, Artensrein AW, Hegerich PA, St. Louis D, Burke DS, and McCutchan F. 1996. Full-Length Sequence and Mosaic Structure of a Human Immunodefficiency Virus Type 1 Isolates from Thailand. *Journal of Virology*. **70(9):**5935-5943.

Cases-Gonzalez CE, Franco S, Martínez MA, and Menendez-Arias L. 2007. Mutational Patterns Associated with the 69 Insertion Complex in Multi-drug-resistant HIV-1 Reverse Transcriptase that Confer Increased Excision Activity and High-level Resistance to zZidovudine. *Journal of Molecular Biology*. **365(2):** 298-309.

Castelbranco EPAF, Souza EDS, Cavalcanti AMS, Martins AN, de Alencar LA, and Tanuri T. 2010. Frequency of Primary Resistance to Antiretroviral Drugs and Genetic Variability of HIV-1 Among Infected Pregnant Women Recently Diagnosed in Luanda-Angola. *AIDS Research and Human Retroviruses*. DOI: 10.1089/aid.2010.0111

CDC (Center for Disease Control). 19981. Kaposi's Sarcoma and *Pneumocystis* pneumonia among Homosexual Men - New York and California. *Morbidity and Mortality Weekly Report*. **30:** 305-308.

CDC (Center for Disease Control). 2008. Revised Surveillance Case Definitions for HIV Infection Among Adults, Adolescents, and Children Aged <18 Months and for HIV Infection and AIDS Among Children Aged 18 Months to <13 Years — United States, 2008. *Morbidity and Mortality Weekly Report.* **57:**RR-10.

CDC. 1989. Interpretation and Use of the Western Blot Assay for Serodiagnosis of Human Immunodeficiency Virus Type 1 Infections. *Journal of American Medical Assciaition (JAMA)*. **24:**3395-3397.

Chan D, and Kim PS. 1998. HIV Entry and its Inhibition. *Cell.* **93(5):**681–684.

Chan DC, Fass D, Berger JM, and Kim PS. 1997. Core Structure of gp41 from the HIV Envelope Glycoprotein. *Cell*. **89:**263–273.

Chaplin B, Eisen G, Idoko J, Onwujekwe D, Idigbe E, Adewole I, Gashau W, S. Meloni S, Sarr AD, Sankale JL, Ekong E, Murphy RL, and Kanki P. 2010. Impact of HIV Type 1 Subtype on Drug Resistance Mutations in Nigerian Patients Failing First-Line Therapy. *AIDS Research and Human Retroviruses.* **26:** DOI: 10.1089/aid.2010.0050

Chou R, Huffman LH, Fu R, Smits AK, and Korthuis T. 2005. Screening for HIV: A Review of the Evidence for the U.S. Preventive Services Task Force. *Annals of Internal Medicine*. **143:**55-73.

Cilliers T, Nhlapo J, Coetzer M, Orlovic D, Ketas T, Olson WC, Moore JP, Trkola A, and Morris L. 2003. The CCR5 and CXCR4 Coreceptors are Both Used by Human Immunodeficiency Virus Type 1 Primary Isolates from Subtype C. *Journal of Virology*. **77:**4449–4456.

Clavel F, and Hance AJ. 2004. HIV Drug Resistance. *The NewEngland Journal of Medicine*. **350:**1023-1035.

Clavel F, Guetard D, Brun-Vezinet F, Chamaret S, Rey M-A, Santos-Ferreira MO, Laurent AG, Dauguet C, Katlama C, Rouzioux C, Klatzmann D, Champalimaud JL, and Montagnier L. 1986. Isolation of a New Human Retrovirus from West African Patients with AIDS. *Science*; **233:**343–346.

Clemente JC, Coman RM, Thiaville MM, Janka LK, Jeung JA, Nukoolkar S, Govindasamy L, Agbandje-McKenna M, McKenna R, Leelamanit W, Goodenow MM, and Dunn BM. 2006. Analysis of HIV-1 CRF_01 A/E Protease Inhibitor Resistance: Structural Determinants for Maintaining Sensitivity and Developing Resistance to Atazanavir. *Biochemistry*. **45:** 5468-5477.

Colson P, Henry M, Tourres C, Lozachmeur D, Gallais H, Gastaut JA, Moreau J, and Tamalet C. 2004: Polymorphism and Drug-Selected Mutations in the Protease Gene of Human Immunodeficiency Virus Type 2 from Patients Living in Southern France. *Journal of Clinical Microbiology*. **42:**570-577.

Converse PJ, Kloos H, Mulatu MS, Haile Mariam D, and Haddis A.. 2009. Bibliography on HIV/AIDS in Ethiopia and Ethiopians in the Diaspora: The 2008 Update. *Ethiopian Journal of Health Development*. **23(1):**73-94.

Converse PJ, Mulatu MS, Kloos H, Haile Mariam D, Wuhib T, and Pankhurst A. 2006. Bibliography on HIV/AIDS in Ethiopia and Ethiopians in the Diaspora: The 2005 Update. *Ethiopian Journal of Health Development*. **20(1):**60-70.

Coombs RW, Reichlderfer PS, and Landay AL. 2003.Recent Observations on HIV Type-1 Infection in the Genital Tract of Men and Women. *AIDS*. **17(4):**455-480.

Costagliola D, Descamps D, Lambert Assoumou L, Morand-Joubert L, Marcelin AG, Brodard V, Pharm D, Delaugerre C, Mackiewicz V, Ruffault A, Izopet J, Plantier JC, Tamalet C, Sabine

Yerly S, Saidi S, Brun-Vezinet F, Masquelier B, and the Agence Nationale de Recherches sur le SIDA et les He´patites Virales (ANRS) AC11 Resistance Study Group. 2007. Prevalence of HIV-1 Drug Resistance in Treated Patients A French Nationwide Study. *Journal of Acquired Immune Deficiency Syndrome*. **46:**12-18.

Craigo JK, and Ross TM. 2010. Hot Topic: Animal Lentiviruses. *Current HIV Research*. **8(1):**1.

Currier JR, Dowling WE, Wasunna KM, Alam U, Mason CJ, Robb ML, Jean K. Carr JK, McCutchan FE, Birx DL and. Cox JH. 2003 Detection of High Frequencies of HIV-1 Cross-subtype Reactive CD8 T Lymphocytes in The Peripheral Blood of HIV-1-Infected Kenyans. *AIDS*. **17:**2149–2157

Daar ES, Little S, Pitt J, Santagelo J, Ho P, Harawa N, Kerndt P, Giorgi JV, Bai J, Gaut P, Richman DD, Mandel S, Nicholas S, and Los Angeles County Primary HIV Infection Recruitment Network. 2001. Diagnosis of Primary HIV-1 Infection. *Annals of Internal Medicine* **134(1):** 25–29.

Dam E, Quercia R, Glass B, Descamps D, Launay O, Duval X, Krausslich HG, Hance AJ, Clavel F, and ANRS 109 Study Group. 2009. Gag Mutations Strongly Contribute to HIV-Resistance to Protease Inhibitors in Highly Drug-experienced Patients Besides Compensating for Fitness Loss. *PLoS Pathogenesis*. **5(3):** e1000345. doi:10.1371/journal.ppat.1000345.

Damond F. Roquebert B, Benard A, Collin G, Miceli M, Yeni P, Brun-Vezinet F, and Descamps D. 2007. Human Immunodeficiency Virus Type 1 (HIV-1) Plasma Load Discrepancies between the Roche COBAS Amplicor HIV-1 Monitor Version 1.5 and the Roche COBAS AmpliPrep/COBAS TaqMan HIV-1 Assays. *Journal of Clinical Microbiology*. **45:**3436–3438.

De Baar MP, Abebe A, Kliphuis A, Tesfaye G, Goudsmit J, and Pollakis G. 2003. HIV-1 C and C' Sub-Clusters Based on Long Terminal Repeat Sequences in the Ethiopian HIV-1 Subtype C Epidemic. In the 24th African Health Science Congress: Challenge and Strategies in Combating Health Problems in Africa, towards Development Effort. African Union Conference Center, Addis Ababa, Ethiopia. September 28-October 2, 2003. [Abstract 15].

De Clercq E. 2009. Anti-HIV drugs: 25 Compounds Approved within 25 Years after the Discovery of HIV. *International Journal of Antimicrobial Agents* **33:**307–320.

De Mondeza C, and Soriano V. 2009. Update on HIV Viral-load Assays: New Technologies and Testing in Resource-limited Settings. *Future Virology*. **4(5):**423-430.

De Wit TFR, Tsegaye A, Wolday D, Hailu B, Aklilu M, Sanders E, Hagos M, Kliphuis A, Pollakis G, Krol A, Geskus R, Miedema F, Goudsmit J, Coutinho R, and Fontanet AL. 2002. Primary HIV-1 Subtype C Infection in Ethiopia. *Journal of Acquired Immune Deficiency Syndrome*. **30:**463-470.

Delgado E, Ampofo WK, Sierra M, Torpey K, Pérez-Alvarez L, Bonney EY, Mukadi YD, Lartey M, Nyarko C, Amenyah RN, Thomson M, and Nájera R. 2008. High Prevalence of Unique Recombinant Forms of HIV-1 in Ghana: Molecular Epidemiology from an Antiretroviral Resistance Study. *Journal of Acquired Immune Deficiency Syndrome*.. **48(5):**599-606.

Desfosses Y, Solis M, Sun Q, Grandvaux N, Carine van Lint C, Burny A, Anne Gatignol A, Wainberg MA, Lin R, and Hiscott J. 2005. Regulation of Human Immunodeficiency Virus Type 1 Gene Expression by Clade-Specific Tat Proteins. *Journal of Virology*. **79(14):** 9180–9191.

DHHS. 2011. *Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents.* Department of Health and Human Services. January 10, 2011; 1–166. Available at http://www.aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL.pdf. Accessed [January, 2011].

DHHS. 2008. *Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents*. Department of Health and Human Services. November 3, 2008; 1-139. Available at http://www.aidsinfo.nih.gov/ContentFiles/ AdultandAdolescentGL.pdf. Accessed (September 19, 2009).

Diallo K, Gotte M, and Wainberg M. 2003. Molecular Impact of the M184V Mutation in HIV-1 Reverse Transcriptase. *Antimicrobial Agents and Chemotherapy*. **47:**3377-3383.

Djoko CF, Rimoin AW, Vidal N, Tamoufe U, Wolfe ND, Butel C, LeBreton M, Tshala FM, Kayembe PK, Muyembe JJ, Edidi-Basepeo S, Pike BL, Fair JN, Mbacham WF, Saylors KE, Mpoudi-Ngole E, Delaporte E, Grillo M, and Peeters M. 2010. High HIV Type 1 Group M pol Diversity and Low Rate of Antiretroviral Resistance Mutations Among the Uniformed Services in Kinshasa, DRC. *AIDS Research and Human Retroviruses*. **26:** DOI: 10.1089/aid.2010.0201.

Doron-Faigenboim A, Stern A, Mayrose I, Bacharach E, and Pupko T. 2005. Selecton: a Server for Detecting Evolutionary Forces at a single Amino Acid Site. *Bioinformatics*. **21(9):**2101-2103.

Doualla-Bell F, Gaolathe T, Avalos A, Cloutier S, Ndwapi N, Holcroft C, Moffat H, Dickinson D, Essex M, Mark A Wainberg MA, and Mine M. 2009. Five-year Follow up of Genotypic Resistance Patterns in HIV-1 Subtype C Infected Patients in Botswana after Failure of Thymidine Analogue-Based Regimens. *Journal of the International AIDS Society*. **12**:25 doi:10.1186/1758-2652-12-25.

Drexler JF, de Souza Luna LK, Pedroso C, Pedral-Sampaio DB, Queiroz AT, Brites C, Netto EM, and Drosten C. 2007. Rates of and Reasons for Failure of Commercial Human Immunodeficiency Virus Type 1 Viral load Assays in Brazil. *Journal of Clinical Microbiology*. **45:**2061–2203. 11.

Dykes C, and Demeter LM. 2007. Clinical Significance of Human Immunodeficiency Virus Type 1 Replication Fitness. *Clinical Microbiology Reviews*. **22(4):** 550-578.

Easterbrook PJ, Smith M, Mullen J, O'Shea S, Chrystie I, de Ruiter A, Tatt ID, Geretti AM, and Zuckerman M. 2010. Impact of HIV-1 Viral Subtype on Disease Progression and Response to Antiretroviral Therapy. *Journal of the International AIDS Society.* **13:**4 http://www.jiasociety.org/content/13/1/4.

Easterbrook PJ, Smith M, Mullen J, O'Shea S, Chrystie I, de Ruiter A, Tatt ID, Geretti AM, and Zuckerman M. 2010. Impact of HIV-1 Viral Subtype on Disease Progression and Response to Antiretroviral Therapy. *Journal of the International AIDS Society*. **13:**4 http://www.jiasociety.org/content/13/1/4.

Elder JH, Lin YC, Fink E, and Grant CK. 2010. Feline Immunodeficiency Virus (FIV) as A Model for Study of Lentivirus Infections: Parallels with HIV. *Current HIV Research*. **8(1):**73-80.

Elford J, Bolding G, and Sherr L. 2002. High-Risk Sexual Behavior Increases among London Gay Men between 1998 and 2001: What is the Role of HIV Optimism? *AIDS*. **16:**1537-1544.

El-Khatib Z, Ekstrom AM, Ledwaba J, Mohapi L, Laher F, Karstaedt A, Charalambous S, Petzold M, Katzenstein D, and and Morris L. 2010. Viremia and Drug Resistance among HIV-1 Patients on Antiretroviral Treatment: a Cross-Sectional Study in Soweto, South Africa. *AIDS*. **24***:*1679–1687.

Esbjörnsson J, Månsson F, Martínez-Arias W, Vincic E, Biague AJ, da Silva ZJ, Fenyö EM, Norrgren H, and Medstrand P. 2010. Frequent CXCR4 Tropism of HIV-1 Subtype A and CRF02_AG During Late-Stage Disease – Indication of an Evolving Epidemic in West Africa. *Retrovirology*. **7:**23 http://www.retrovirology.com/content/7/1/23.

Escoto-Delgadillo M, Vazquez-Valls E, Ramırez-Rodrıguez M, Corona-Nakamura A, Amaya-Tapia G, Quintero-Perez N, Panduro-Cerda A, and Torres-Mendoza BM. 2005. Drug-Resistance Mutations in Antiretroviral-Naıve Patients with Established HIV-1 Infection in Mexico. *HIV Medicine*. **6:** 403–409.

Eshleman SH, Crutcher G, Petrauskene O, Kunstman K, Cunningham SP, Trevino C, Davis C, Kennedy J, Fairman J, Foley B, and Kop JA. 2005b. Sensitivity and Specificity of the ViroSeq Human Immunodeficiency Virus Type 1 (HIV-1) Genotyping System for Detection of HIV-1 Drug Resistance Mutations by Use of an ABI Prism 3100 Genetic Analyzer. *Journal of Clinical Microbiology*. **43:**813–817.

Farmer P. 1996. Social Inequalities and Emerging Infectious Diseases. *Emerging Infectious Diseases*. **2(4):**259-269.

Fauci A. 2003. HIV and AIDS: 20 Years of Science. *Nature Medicine*. **9(7):**839-843.

Ferrantelli F, Kitabwalla M, Rasmussen RA, Cao C, Chou TC, Katinger H, Stiegler G, Cavacini LA, Bai Y, Cotropia J, Ugen KE, and Ruprecht RM. 2004. Potent Cross-Group Neutralization of Primary Human Immunodeficiency Virus Isolates with Monoclonal Antibodies – Implications for Acquired Immunodeficiency Syndrome Vaccine. *Journal of Infecioust Diseases*. **189:**71-74.

Fischer W, Perkins S, Theiler J, Bhattacharya T, Yusim K, Funkhouser R, Kuiken C, Haynes B, Letvin NL, Walker BD, Hahn BH, and Korber BT. 2007. Polyvalent Vaccines for Optimal Coverage of Potential T-cell Epitopes in Global HIV-1 Variants. *Nature Medicine*. **13(1):** 100-106.

Fleury HJ, Toni T, Lan NT Hung PV, Deshpande A, Recordon-Pinson P, Boucher S, Lazaro E, Jauvin V, Lavignolle-Aurillac V, Lebel-Binay S, Cheret A, And Masquelier B. 2006. Susceptibility to Antiretroviral Drugs of CRF01_AE, CRF02_AG, and Subtype C Viruses from Untreated Patients of Africa and Asia: Comparative Genotypic and Phenotypic Data. *AIDS Research and Human Retroviruses.* **22:**357-366.

Fontanet AL, Messele T, Dejene A, Enquselassie F, Abebe A, Cutts FT, de Wit TR, Sahlu T, Bindels P, YenenehH, Coutinho RA, and Nokes DJ. 1998. Age- and Sex-specific HIV-1 Prevalence in the Urban Community Setting of Addis Ababa, Ethiopia. *AIDS*. **12:**315–322.

Fraham N, B Korber, C Adams, J Szinger, R Draenert, M Addo, M Feeney, K Yusim, K Sango, N Brown, D SenGupta, A Piechocka-Trocha, T Simonis, F Marincola, A Wurcel, D Stone, C J Russell, P Adolf, D Cohen, T Roach, A St John, A Khatri, K Davies, J Mullins, P Goulder, B Walker, and C Brander. 2004. Consistent CTL Targeting of Immunodominant Regions in HIV Across Multiple Ethnicities. *Journal of Virology.* 7**8(5):**2187-2000.

Frater AJ, Beardall A, Ariyoshi K, Churchill D, Galpin S, Clarke JR, Weber JN, and McClure MO. 2001. Impact of Baseline Polymorphisms in RT and Protease on Outcome of Highly Active Antiretroviral Therapy in HIV-1-Infected African Patients. *AIDS* **15:**1493-1502.

Frater AJ, Dunn DT, Beardall AJ, Ariyoshi K, Clarke JR, McClure MO, and Weber JN. 2002. Comparative Response of African HIV-1-Infected Individuals to Highly Active Antiretroviral Therapy. *AIDS*. **16:**1139–1146.

Freed EO. 2001. HIV-1 Replication. Somatic Cell Molecular Genetics. **26(1-6):**13-33.

Froland SS, Jenum P, Lindboe CF, Wefring KW, Linnestad, PJ, and Bohmer T. 1988. HIV-1 Infection in Norwegian Family before 1970. *The Lancet*. p.1344-1345.

Gallo RC, Sarin PS, Gelmann EP, Robert-Guroff M, Richardson E, Kalyanaraman VS, Mann D, Sidhu GD, Stahl RE, Zolla-Pazner S, Leibowitch J, and Popovic M. 1983. Isolation of Human T-cell Leukemia Virus in Acquired Immune Deficiency Syndrome (AIDS). *Science*. **220:**865–867.

Gandhi N, Saiyed Z, Thangavel S, Rodriguez J, Rao KV, Nair MP. 2009. Differential Effects of HIV Type 1 Clade B and Clade C Tat Protein on Expression of Proinflammatory and Anti-inflammatory Cytokines by Primary Monocytes. *AIDS Research and Human Retroviruses* **25:**691–699.

Gandhi RT, Wurcel A, Rosenberg ES, Johnston MN, Hellmann N, Bates M, Hirsch MS, and Walker BD. 2003. Progressive Reversion of Human Immunodeficiency Virus Type 1 Resistance In Vivo after Transmission of a Multi Drug-Resistance Virus. *Clinical Infectious Diseases*. **37:**1693-1698.

Gao F, Robertson DL, Carruthers CD, Li YY, Bailes E, Kostrikis LG, Salminen MO, Bibollet-Ruche F, Peeters M, Ho DD, Shaw GM, Sharp PM, and Hahn BH. 1998. An Isolate of

Human Immunodeficiency Virus Type 1 Originally Classified as Subtype I Represents a Complex mosaic Comprising Three Different Group M subtypes (A, G, and I). *Journal of Virology*. **72:**10234–10241.

Gao Y, Paxinos E, Galovich J, Troyer R, Baird H, Abreha M, Kityo C, Mugyenyi P, Petropoulos C, and Arts EJ. 2004. Characterization of a Subtype D Human Immunodeficiency Virus Type 1 Isolate that Was Obtained from an Untreated Individual and that Is Highly Resistant to Non-nucleoside Reverse Transcriptase Inhibitors. *The Journal of Virology*. **78:**5390–5401.

Gao, F; Bailes, E; Robertson, DL; Chen, Y; Rodenburg CM, Michael SF, Cummins LB, Arthur LO, Peeters M, Shaw GM, Sharp PM, and Hahn BH. 1999. Origin of HIV-1 in the Chimpanzee *Pan troglodytes troglodytes*. *Nature*. **397:**436-444.

Garber DA, Silvestri G, and Feinberg MB. 2004. Prospects for an AIDS Vaccine: Three Big Questions, No Easy Answers. *Lancet Infectious Disease* **4:**397–413.

Garrido C, Zahonero N, Fernandes D et al. 2008. Subtype Variability, Virological Response and Drug Resistance Assessed on Dried Blood Spots Collected from HIV Patients on Antiretroviral Therapy in Angola. *Journal of Antimicrobial Chemotherapy.* **61:**694-698.

Gaschen B, Taylor J, Yusim K, Foley B, Gao F, Lang D, Novitsky V, Haynes B, Hahn BH, Bhattacharya T, and Korber B. 2002. Diversity Considerations in HIV-1 Vaccine Selection. *Science*. **296:**2354–2360.

Gaudy C, MoreauA,Brunet S, Descamps J-M, Deleplanque P, Brand D, Barin F. 2004. Subtype B Human Immunodeficiency Virus (HIV) Type 1 Mutant that Escapes Detection in a Fourth-generation Immunoassay for HIV Infection. *Journal of Clinical Microbiology.* **42:**2847–2849.

Geelen S, Lange J, Borleffs J, Wolfs T, Weersink A, and Schuurman R. 2003. Failure to Detect a Non-B HIV-1 Subtype by the HIV-1 Amplicor Monitor Test, Version 1.5: a Case of Unexpected Vertical Transmission. *AIDS.* **17:**781–782.

Gelderblom H. 1997. Fine Structure of HIV and SIV. pp. IV-37-50 in HIV Molecular Immunology Database 1997. Edited by: Korber B, Brander C, Haynes B, Koup R, Moore J, Walker B. Published by: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM.

Geretti AM, Harrison L, Green H, Caroline C, Hill T, Fearnhill E, Deenan Pillay D, and Dunn D, and on behalf of the UK Collaborative Group on HIV Drug Resistance and the UK

Collaborative HIV Cohort Study. 2009. Effect of HIV-1 Subtype on Virologic and Immunologic Response to Starting Highly Active Antiretroviral Therapy. *Clinical Infectious Diseases*. **48:**1296–1305.

Geretti AM. 2006. HIV-1 Subtypes: Epidemiology and Significance for HIV Management. *Current Opinion in Infectious Diseases*. **19:**1-7.

Ghosn J, Chaix ML, and Delaugerre C. 2009. HIV-1 Resistance to First- and Second-Generation Non-nucleoside Reverse Transcriptse Inhibitor. *AIDS Reviews*. **11:**165-173.

Gibb DM, Walker AS, Kaye S, De Rossi A, Ait-Khaled M, Pillay D, Munoz-Fernandez MA, Loveday C, Compagnucci A, Dunn DT, and Babiker AG. 2002. Evolution of Antiretroviral Phenotypic and Genotypic Drug Resistance in Antiretroviral-naive HIV-1-Infected Children Treated with Abacavir/Lamivudine, Zidovudine/Lamivudine, or Abacavire/-Zidovudine, with or without Nelfinavire (the PENTA 5 trial). *Antiviral Therapy*. **7:**293-303.

Gomez C, and Hope TJ. 2005. The ins and outs of HIV Replication. *Cellular Microbiology*. **7(5):**621-626.

Gonzalez LM, Aguiar RS, Afonso A, Brindeiro PA, Arruda MB, Soares MA, Brindeiro RM, and Tanuri A. 2006. Biological Characterization of Human Immunodeficiency Virus Type 1 Subtype C Protease Carrying Indinavir Drug-Resistance Mutations. *Journal of General Virology.* **87:**1303-1309.

Gonzalez LM, Brindeiro RM, Tarin M, Calazans A, Soares MA, Cassol S, and Tanuri A. 2003. *In Vitro* Hypersusceptibility of Human Immunodeficiency Virus Type 1 Subtype C Protease to Lopinavir. *Animicrobial Agents and Chemotherapy*. **47(9):**2817-2822.

Gordon M, De Oliveira T, Bishop K, Coovadia HM, Madurai L, Englebrecht S, van Rensburg EJ, Mosam A, Smith A, and Cassol S. 2003. Molecular Characterization of Human Immunodeficiency Virus Type 1 Subtype C Viruses from KwaZulu-Natal, South Africa: Implications for Vaccine and Antiretroviral Control Strategies. *Journal of Virology*. **77:**2587-2599.

Gorry PR, Sterjoski J, Churchill M, Witlox K, Gray L, Cunningham A, and Wesselingh S. 2004. The Role of Viral Coreceptors and Enahanced Macrophage Tropism in Human Immunodeficiency Virus Type 1 Disease Progression. *Sexual Health*. **1:**23-34.

Gottesman BS, Grossman Z, Lorber M, Levi I, Shitrit P, Katzir M, Shahar E, Gottesman G, and Chowers M. 2006. Comparative Performance of the Amplicor HIV-1 Monitor Assay Versus

NucliSens EasyQ in HIV Subtype C-Infected patients. *Journal of Medical Virology.* **8(7):**883-887.

Göttlinger HG. 2001. HIV-1 Gag: a Molecular Machine Driving Viral Particle Assembly and Release. pp. 2-28 in HIV Sequence Compendium 2001. Edited by: Kuiken C, Foley B, Hahn B, Marx P, McCutchan F, Mellors JW, Wolinsky S, Korber B. Published by: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM, LA-UR 02-2877.

Gottwein E, and Krausslich HG. 2005. Analysis of Human Immunodeficiency Virus Type 1 Gag Ubiquitination. *Journal of Virology*. **79(14):**9134–9144.

Gousset K, Ablan SD, Coren LV, Ono A, Soheilian F, Nagashima K, Ott DE, and Freed EO. 2008. Real-Time Visualization of HIV-1 GAG Trafficking in Infected Macrophages. *PLoS Pathog*enes **4(3):** e1000015. doi:10.1371/journal.ppat.1000015.

Graham SM. 2007. Clinical Impact of HIV-1 Subtype: Are there Important Differences? *Future HIV Therapy*. **1(3):**273-290.

Gray RR, Tatem AJ, Lamers S, Wei Hou W, Laeyendecker O, Serwadda D, Serwankambo N, Gray RH, Wawer M, Thomas C. Quinn TC, Goodenow MM, and Salemi M. 2009. Spatial Phylodynamics of HIV-1 Epidemic Emergence in East Africa. *AIDS*. **23:** DOI:10.1097/QAD.0b013e32832faf61.

Greenwald JL, Burstein GR, Pincus J, and Branson B. 2006. A Rapid Review of Rapid HIV Antibody Tests. *Current Infectious Disease Reports.* **8:**125-131.

Gross KL, Porco TC, and Grant RM. 2004. HIV-1 Superinfection and Viral diversity. *AIDS*. **18:**1513-1520.

Grossman Z, Vardinon N, Chemtob D, Alkan ML, Bentwich Z, Burke M, Gottesman G, Istomin V, Levi I, Maayan S, Shahar E, and Schapiro JM for the Israel Multi-Center Study Group. 2001. Genotypic Variation of HIV-1 Reverse Transcriptase and Protease: Comparative Analysis of Clade C and Clade B. *AIDS*. **15:**1453–1460.

Gurtler LG, Zekeng L, Simon F, Eberle J, Tsague JM, Kaptue L, Brust S, and Knapp S. 1995. Reactivity of Five Anti-HIV-1 Subtype O Specimens with Six Different Anti-HIV Screening ELISAs and Three Immunoblots. *Journal of Virological Methods.* **51:**177–184.

Hahn BH, Shaw GM, De Cock KM, Sharp PM. 2000. AIDS as a Zoonosis: Scientific and Public Health Implications. *Science*. **287:**607-414.

Hall TA. 1999. BioEdit: a User-Friendly Biological Sequence Alignment Editor and Ananlysis Program for Windows 95/98/NT. *Nucleic Acids Symposium Series*. **41:**95-98.

Hallack R, Doherty LE, Wethers JA, and Parker MM. 2008. Evaluation of Dried Blood Spot Specimens for HIV-1 Drug-Resistance Testing Using the Trugene® HIV-1 Genotyping Assay. *Journal of Clinical Virology*. **41:**283-287.

Halvas EK, Wiegand A, Boltz VF, Kearney M, Nissley D, Wantman M, Hammer SM, Palmer S, Vaida F, Coffin JM, and Mellors JW. 2010. Low Frequency Nonnucleoside Reverse Transcripatse Inhibitor-Resistant Variants Contribute to Failure of Efavirenz-Containing Regimens in Treatment-Experienced Paients.*The Journal of Infectious Diseases*. **201:**672-680.

Hamers RL, Siwale M, Wallis CL, Labib M, van Hasselt R, Stevens WS, Schuurman R, Wensing MJ, Vugt MV, and de Wit TFR, on Behalf of the PharmAccess African Studies to Evaluate Resistance. 2010. HIV-1 Drug Resistance Mutations Are Present in Six Percent of Persons Initiating Antiretroviral Therapy in Lusaka, Zambia. *Journal of Acquired Immune Deficiency Syndrome*. **55(1):**95-101.

Hamers RL, Smit PW, Stevens W, Schuurman R, de Wit TFR. 2009. Dried Fluid Spots for HIV Type-1 Viral Load and Resistance Genotyping: a Systematic Review. *Antiviral Therapy*. **14(5):**619-629.

Hamkar R, Mohraz M, Lorestani S, Aghakhani A, Truong HAM, McFarland W, Banifazl M, Eslamifar A, Foroughi F, Pakfetrat A, and Ramezani A. 2010. Assessing subtype and drug-resistance-associated mutations among antiretroviral-treated HIV-infected patients *AIDS*. **24 (suppl 2):**S85–S91.

Hansasuta P, and Rowland-Jones SL. 2001. HIV-1 Transmission and Acute HIV-1 Infection. *British Medical Bulletin*. **58:**109-127.

Hantano H, Delwart EL, Norris PJ, Lee TH, Neilands TB, Kelley CF, Hunt PW, Hoh R, Linnen JM, Martin JN, Busch MP, and Deeks SG. 2010. Evidence of Persistent Low-Level Viremia in Long-Term HAART-Suppressed, HIV-Infected Individuals. *AIDS*. **24:** DOI:10.1097/QAD.0b013e32833dba03.

Harris ME, Maayan S, Kim B, Zeira M, Ferrari G, Birx DL, And Mccutchan FE. 2003. A Cluster of HIV Type 1 Subtype C Sequences from Ethiopia, Observed in Full Genome Analysis, Is

Not Sustained in Subgenomic Regions. *AIDS Research and Human Retroviruses*. **19(12):**1125-1133.

Haubrich R and Demeter L. 2001. International Perspectives on Antiretroviral Resistance. Clinical Utility of Resistance Testing: Retrospective and Protective Data Supporting Use and Current Recombinations. *Journal of Acquired Immune Deficiency Syndrome*. **26(suppl 1):**S51-S59.

Hawes SE, Sow PS, Stern JE, Critchlow CW, Gottlieb GS, and Kiviat NB. 2008. Lower Levels of HIV-2 than HIV-1 in the Female Genital Tract: Correlates and Longitudinal Assessment of Viral Shedding. *AIDS*. **22:**2517–2525.

Hawkins C, Chaplin B, Idoko J, Idehen C, Adewole I, Gashau W, Murphy R, Kanki P, and APIN Plus/Harvard PEPFAR Team. 2007. Clinical and Genotypic Findings in HIV-Infected Patients with the K65R Mutations Failing Stavudine, Lamivudine, and Nevirapine First Line Antiretroviral Therapy. *Antiviral Therapy*. **12:**S78.

Hemelaar J, Gouws E, Ghys P, Osmanov S, and WHO/UNAIDS Network for HIV Isolation and Characterization. 2010. Global Trends in Molecular Epidemiology of HIV-1 during 2000-2007. XVIII International AIDS Conference July 18-23, 2010 Vienna, Austria. Abstract number THAA0201.

Hemelaar J, Gouws E, Ghys PD, and Osmanov S. 2006. Global and Regional Distributionof HIV-1 Genetic Subtypes and Recombinants. *AIDS*. **20:**W13-W23.

Hierholzer M, Graham RR, Khidir IE, Tasker S, Darwish M, Chapman GD, Fagbami AH, Soliman A,. Birx DL, McCutchan F, and Carr JK. 2002. HIV Type 1 Strains from East and West Africa Are Intermixed in Sudan. *AIDS Research and Human Retroviruses*. **18(15):**1163-1166. doi:10.1089/088922202320567923.

Hirsch MS. 2005. Initiating Antiretroviral Treatment for Patients with HIV Infection. *Advanced Studies in Medicine*. **5(2A):**S97-S104.

Hiscott J, Kwon H, and Genin P. 2001. Hostile Takeovers: Viral Appropriation of the NF-κB. *The Journal of Clinical Investigation*. **107(2):**143-151.

Hoffman NG, Seillier-Moiseiwitsch F, Ahn JH, Walker JM, and Swanstrom R. 2002. Variability in the Human Immunodeficiency Virus Type 1 gp120 Env Protein Linked to Phenotype-Associated Changes in the V3 Loop. *Journal of Virology*. **76(8):**3852-3864.

Hogg RS, Bangsberg DR, Lima VD, Alexander C, Bonner S, Yip B, Wood E, Dong WWY, Montaner JSG, and Harrigan PR. 2006. Emergence of Drug Resistance Is Associated with an Increased Risk of Death among Patients First Starting HAART. *PLoS Medicine*. **3(9):**e356. DOI: 10.1371/journal.pmed.0030356.

Holguın A, Lopez M, Molinero M, and Soriano V. 2008b. Performance of Three Commercial Viral Load Assays, Versant Human Immunodeficiency Virus Type 1 (HIV-1) RNA bDNA v3.0, Cobas AmpliPrep/Cobas TaqMan HIV-1, and NucliSens HIV-1 EasyQ v1.2, Testing HIV-1 Non-B Subtypes and Recombinant Variants. *Journal of Clinical Microbiology*. **46(9):**2918-2923.

Holguin A, Lospitao E, Lopez M, de Arellano ER, Pena MJ, del Romero J, Martın C, and Soriano V. 2008a. Genetic Characterization of Complex Inter-Recombinant HIV-1 Strains Circulating in Spain and Reliability of Distinct Rapid Subtyping Tools. *Journal of Medical Virology*. **80:**383–391.

Holguin A, Paxinos E, Hertogs K, Womac C, Soriano V: 2004. Impact of Frequent Natural Polymorphisms at the Protease Gene on the *in vitro* Susceptibility to Protease Inhibitors in HIV-1 Non-B Subtypes. *Journal of Clinical. Virology.* **31:**215-220.

Holguin A, Sune C, Hamy F, Soriano V, and Klimkait T. 2006. Natural Polymorphisms in the Protease Gene Modulate the Replicative Capacity of Non-B HIV-1 Variants in the Absence of Drug Pressure. *Journal of Clinical Virology*. **36:**264–271.

Holumberg SD, Hamburger ME, Moorman AC, Wood KC, and Palella FJ,Jr. for the HIV Outpatient Study Investigators. 2003. Factors Associated with Maintenance of of Long Term Plasma Human Immunodeficiency Virus RNA Supression. *Clinical Infectious Diseases*. **37:**702-707.

Hu DJ, Buvé A, Bagg J, van der Groen G, and Dondero TJ. 1999. What Role Does HIV-1 Subtype Play in Transmission and Pathogenesis? An Epidemiological Perspective. *AIDS*. **13:**873–881.

Hu DJ, Vanichseni S, Mastro TD, Raktham S, Young NL, Mock PA, Subbarao S, Parekh BS, Srisuwanvilai L, Sutthent R, Wasi C, Heneine W, and Choopanya K. 2001. Viral Load Differences in Early Infection with Two HIV-1 Subtypes. *AIDS*. **15:**683-691.

Huang W, Toma J, Fransen S, Stawiski E, Reeves JD, Whitcomb JM, Parkin N, and Petropoulos CJ. 2008. Coreceptor Tropism Can Be Influenced by Amino Acid Substitutions in the gp41

Transmembrane Subunit of Human Immunodeficiency Virus Type 1 Envelope Protein. *Journal of Virology*. **81(11):** 5584–5593.

Hue S, Gifford RJ, Dunn D, Fernhill E, and Pillay D, on Behalf of the UK Collaborative Group on HIV Drug Resistance. 2009. Demonstration of Sustained Drug-Resistant Human Immunodeficiency Virus Type 1 Lineages Circulating among Treatment-Naïve Individuals. *Journal of Virology*. **83(6):**2645-2654.

Hupfeld J, and Efferth T. 2009. Drug Resistance of Human Immunodeficiency Virus and Overcoming it by Natural Products. *In Vivo*. **23:**1-6.

Hurley JH, Boura E, Carlson LA, and Rozycki B. 2010. Membrane Budding. *Cell*. **143:**875-887.

Hussain A, Wesley C, Khalid M, Chaudhry A, and Jameel S. 2008. Human Immunodeficiency Virus Type 1 Vpu Protein Interacts with CD74 and Modulates Major Histocompatibility Complex Class II Presentation. *Journal of Virology*. **82(2):** 893–902.

Hussein M, Abebe A, Pollakis G, Brouwer M, Petros B, Fontanet AL, and de Wit TFR. 2000. HIV-1 Subtype C in Commercial Sex Workers in Addis Ababa, Ethiopia. *Journal of Acquired Immune Deficiency Syndrome*. **23:**120-127.

Hutter H, Nowak D, Mossner M, Ganepola S, Mußig A, Allers K, Schneider T, Hofmann J, Kücherer C, Blau O, Blau IW, Hofmann WK, and Thiel E . 2009. Long-Term Control of HIV by *CCR5* Delta32/ Delta32 Stem-Cell Transplantation. *The New England Journal of Medicine*. **360:**692-698.

Ibe S, Yokomaku Y, Shiino T, Tanaka R, Hattori J, Fujisaki S, Iwatani Y, Mamiya N, Utsumi M, Kato S, Hamaguchi M, and Sugiura W. 2010. HIV-2 CRF01_AB: First Circulating Recombinant Form of HIV-2. *Journal of Acquired Immune Deficiency Syndrome*. **54:**241-247.

Imamichi H, Koita O, Dabitao D, Dao S, Ibrah M, Sogoba D, Dewar RL, Berg SC, Jiang MK, Parta M, Washington JA, Polis MA, Lane HC, Tounkara A. 2009. Identification and Characterization of CRF02_AG, CRF06_cpx, and CRF09_cpx Recombinant Subtypes in Mali, West Africa. *AIDS Research and Human Retroviruses*. **25(1):**45-55.

Iversen AKN, Learn GH, Skinhoj P, Mullins JI, McMichael AJ, and Rambaut A. 2005. Preferential Detection of HIV Subtype C' over Subtype A in Cervical Cells from a Dually Infected Woman. *AIDS*. **19(9):**990–993.

Jaffe HW, Bregman DJ, and Selik LM. 1983. Acquired Immune Deficiency Syndrome in the United States: the first 1,000 cases. *The Journal of Infectious Diseases.* **148:**339-345.

Jagodzinski LL, Cooley JD, Weber M, Michael NL. 2003. Performance Characteristics of Human Immunodeficiency Virus Ype 1 (HIV-1) Genotyping Systems in Sequence-Based Analysis of Subtypes Other than HIV-1 Subtype B. *Journal of Clinical Microbiology*. **41:**998–1003.

Janssen W, Buve A, and Nkengasong JN. 1997. The Puzzle of HIV-1 Subtypes in Africa. *AIDS*. **11:**705-712.

Jeeninga RE, Hoogenkamp M, Armand-Ugon M, De Baar M, Verhoef K, and Berkhout B. 2000. Functional Differences between the Long Terminal Repeat Transcriptional Promoters of Human Immunodeficiency Virus Type 1 Subtypes A through G. *Journal of Virology*. **74(8):** 3740–3751.

Jegede O, Babu J, Santo RD, McColl DJ, Weber J, and Quinoes-Mateu. 2008. HIV Type 1 Integrase Inhibitors: From Basic Research to Clinical Implications. *AIDS Reviews*. **10:**172-189.

Jenny-Avital ER, and Beatrice ST. 2001. Erroneously Low or Undetectable Plasma Human Immunodeficiency Virus Type 1 (HIV-1) Ribonucleic Acid Load, Determined by Polymerase Chain Reaction, in West African and American Patients with Non-B Subtype HIV-1 Infection. *Clinical Infectious Disease*s. **32:**1227–1230.

Johannessen A, Trøseid M, and Calmy A. 2009. Dried Blood Spots Can Expand Access to Virological Monitoring of HIV Treatment in Resource-Limited Settings. *Journal of Antimicrobial Chemotherapy*. **64:**1126–1129.

Johanson JA, Li JF, Wei X, Lipscomb J, Irlbeck D, Craig C, Smith A, Bennett DE, Monsour M, Sandsorm P, Lanier ER, and Heneine W. 2008b. Minority HIV-1 Drug Resistance Mutations are Present in Antiretroviral Treatment-Naïve Populations and Associate with Reduced Treatment Efficacy. *PLoS Medicine*. **5(7):**e158. doi:10.1371/journal.pmed.0050158.

Johnson VA, Brun-Vezinet F, Clotet B, Gunthard HF, Kuritzkes DR, D, Schapiro JM, and Richman DD. 2008a. Update of the Drug Resistance Mutations in HIV-1: *Topics in HIV Medicine.* 16:62-68.

Johnson VA, Brun-Vezinet F, Clotet B, Gunthard HF, Kuritzkes DK, Pillay D, Schapiro JM, and RichmanDD. 2009. Update of the Drug Resistance Mutations in HIV-1: December 2009. *Topics in HIV Medicine*. **17(5):**138-145.

John-Stewart GC, Nduati RW, Rousseau CM, Mbori-Ngacha DA, Richardson BA, Rainwater S, Panteleeff DD, and Overbaugh J. 2005. Subtype C Is Associated with Increased Vaginal Shedding of HIV-1. *The Journal of Infectious Diseases.* **192:**492-496.

Jukes TH, and Cantor CR. 1969. Evolution of Protein Molecules. In: Munro HN, editor, *Mammalian Protein Metabolism*, pp. 21-132, Academic Press, New York.

Kahn JO, and Walker BD. 1998. Acute Human Immunodeficiency Virus Type 1 Infection. The *New England Journal of Medicine*. **331**(1):33–39.

Kaleebu P, French N, Mahe C, Yirrell D, Watera C, Lyagoba F, Nakiyingi J, Rutebemberwa A, Morgan D, Weber J, Gilks C, and Whitworth J. 2002. Effect of Human Immunodeficiency Virus (HIV) Type 1 Envelope Subtypes A and D on Disease Progression in a Large Cohort of HIV-1-Positive Persons in Uganda. *The Journal of Infectious Diseases*. **185:**1244–1250.

Kaleebu P, Ross A, Morgan D, Yirrell D, Oram J, Rutebemberwa A, Lyagoba F, Hamilton L, Biryahwaho B, and Whitworth J. 2001. Relationship between HIV-1 Env Subtypes A and D and Disease Progression in a Rural Ugandan Cohort. *AIDS*. **15:**293-299.

Kamps BS, and Hoffmann C. 2007. *Introduction. In: HIV Medicine 2007.* Edited by Christian Hoffman, JK Rockstro, and BS Kamps. 15th Edition. Flying Publisher. Available at: www.HIVMedicine.com

Kandathil AJ, Kannangai R, Abraham OC, Pulimood SA, Jensen MA, and Sridharan G. 2009. A Comparison of Interpretation by Three Different HIV Type 1 Genotypic Drug Resistance Algorithms Using Sequences from Non-Clade B HIV Type 1 Strains. *AIDS Research and Human Retroviruses* **25(3):**315-318.

Kanki PJ, Hamel DJ, Sankale JL, Hsieh CC, Thior I, Barin F, Woodcock SA, Gue`ye-Ndiaye A, Zhang E, Montano M, Siby T, Marlink R, NDoye I, Essex ME, and MBoup S. 1999. Human Immunodeficiency Virus Type 1 Subtypes Differ in Disease Progression. *The Journal of Infectious Diseases.* **179:**68-73.

Kanki PJ, Travers KU, Mboup S, Hsieh CC, Marlink RG, Gueye-NDiaye A, Siby T, Thior I, Hernandex-Avila M, Sankale JL, *et al*. 1994. Slower Heterosexual Spread of HIV-2 than HIV-1. *Lancet*. **343:**943–946

Kantor R, and Katzenstein D. 2003. Polymorphism in HIV-1 Non-Subtype B Protease and Reverse Transcriptase and its Potential Impact on Drug Susceptibility and Drug Resistance Evolution. *AIDS Review*. **5(1):**25-35.

Kantor R, and Katzenstein D. 2004. Drug Resistance in Non-Subtype B HIV-1. *Journal of Clinical Virology.* **29:**152-159.

Kantor R. 2006. Impact of HIV-1 *pol* Diversity on Drug Resistance and its Clinical Implication. *Current Opinion in Infectious Diseases.* **19:**594-606.

Kassu A, Fujino M, Matsuda M, WA MN, Ota F, and Sugiura W. 2007. Molecular Epidemiology of HIV Type 1 in Treatment-Naïve Patients in North Ethiopia. *AIDS Rsearch and Human Retroviruses*. **23:**564–568.

Kebede D, Aklilu M, and Sanders E. 2000. The HIV Epidemic and the State of its Surveillance in Ethiopia. *Ethiopian Medical Journal*. **38:**283-302.

Keele BF, Van Heuverswyn F, Li Y, Bailes E, Takehisa J, Santiago ML, Bibollet-Ruche F, Chen Y, Wain LV, Liegeois FL, Loul S, Ngole EM, Bienvenue Y, Delaporte E, Brookfield JFY, Sharp PM, Shaw GM, Peeters M, and Hahn BH. 2006. Chimpanzee Reservoirs of Pandemic and Nonpandemic HIV-1. *Science*. **313(5786):** 523–526.

Keller M, Lu Y, Lalonde RG, and Klein MB. 2009. Impact of HIV-1 Viral Subtype on CD4+ T-cell Decline and Clinical Outcomes in Antiretroviral Naive Patients Receiving Universal Healthcare. *AIDS*. **23:** DOI:10.1097/QAD.0b013e328326f77f

Khamadi SA, Lihana RW, Ochieng W, Lwembe R, Kiptoo MK, Kinyua J, Lagat N, Muriuki J, Mwangi J, Oishi I, Ichimura H, Mpoke S, Songok EM. 2004. HIV-1 Subtypes in the Northern Border Region of Kenya. International Conference on AIDS (15th : 2004 : Bangkok, Thailand). *Int Conf AIDS.* 2004 Jul 11-16; 15: abstract no. TuPeA4373.

Kijak GH, Rubio AE, Pampuro SE, Zala C, Galli R, Montaner JS, and Salomon H. 2003. Discrepant Results in the Interpretation of HIV-1 Drug-Resistance Genotypic Data among Widely Used Algorithms. *HIV Medicine*. **4:**72–78.

Kilmarx PH. 2009. Global Epidemiology of HIV. *Current Opinion in HIV and AIDS*. **1:**240-246.

King JE, Eugenin EA, Buckner CM, and Berman JW. 2006. HIV tat and Neurotoxicity. *Microbes and Infection*. **8:**1347–1357.

Kiwanuka N, Laeyendecker O, Quinn TC, Wawer MJ, Shepherd J, Robb M, Kigozi G, Kagaayi J, Serwadda D, Makumbi FE, Reynolds SJ, and Gray RH. 2009. HIV-1 Subtypes and

Differences in Heterosexual HIV Transmission among HIV-Discordant Couples in Rakai, Uganda. *AIDS*. **23:**2479-2484.

Kiwanuka N, Laeyendecker O, Robb M, Kigozi G, Arroyo MA, and McCutchan F, Eller LA, Eller M, Makumbi F, Birx D, Wabwire-Mangen W, Serwadda D, Sewankambo NK, Quinn TC, Wawer M, and Gray R. 2008. Effect of HIV-1 Subtypes on Disease Progression in a Cohort of Incident Cases Infected with Pure Subtypes, Recombinants, and Multiple Subtypes, Rakai, Uganda. *The Journal of Infectious Diseases*. **197:**707–713.

Kiwanuka N, Robb M, Laeyendecker O, Kigozi G, Wabwire-Mangen F, Makumbi FE, Nalugoda F, Kagaayi J, Eller M, Eller LA, Serwadda D, Sewankambo NK, Reynolds SJ, Quinn TC, Gray RH, Wawer MJ, andWhalen CC. 2010. HIV-1 Viral Subtype Differences in the Rate of CD4+ T-Cell Decline Among HIV Seroincident Antiretroviral Naive Persons in Rakai District, Uganda. *Journal of Acquired Immune Deficiency Syndrome*. **54(2):**180–184. doi:10.1097/QAI.0b013e3181c98fc0.

Klatte EC. 2010. Pathology of AIDS, Virsion 21. Mercer University, School of Medicine Savannah. April 2010. Available at http://library.med.utah.edu/WebPath/AIDS2010.PDF

Kloos H, Assefa Y, Haile Mariam D, Mulatu MS, and Converse PJ. 2008. Bibliography on HIV/AIDS in Ethiopia and Ethiopians in the Diaspora: The 2007 Update, with an Emphasis on the ART Scale-up. *Ethiopian Journal of Health Development*. **22(1):**71-95.

Koch WH, Sullivan PS, Roberts C, Francis K, Downing R, Mastro TD, Nkengasong J, Hu D, Masciotra S, Schable C, and Lal RB. 2001. Evaluation of United States-Licensed Human Immunodeficiency Virus Immunoassays for Detection of Group M Viral Variants. *Journal of Clinical Microbiology*. **39:**1017–1020.

Kolata Gina. 1987. Boy's 1969 Death Suggests AIDS Invaded U.S. Several Times. *New York Times*. 28 October 1987.

Korber B, Gaschen B, Yusim K, Thakallapally R, Kesmir C, and Detours V. 2001. Evolutionary and Immunological Implications of Contemporary HIV-1 Variation. *British Medical Bulletin.* **58:** 19–42.

Korber B. 2000. HIV Signature and Sequence Variation Analysis. Computational Analysis of HIV Molecular Sequences, Chapter 4, pages 55-72. Allen G. Rodrigo and Gerald H. Learn, eds. Dordrecht, Netherlands: Kluwer Academic Publishers.

Korber BT, Letvin NL, and Haynes BF. 2009. T-Cell Vaccine Strategies for Human Immunodeficiency Virus, the Virus with a Thousand Faces. *Journal of Virology*. **83:**8300–8314.

Krebs FC, Hogan TH, Quiterio S, Gartner S, and Wigdahl B. 2001. *Lentiviral LTR-Directed Expression, Sequence Variation, and Disease Pathogenesis*. pp. 29-70. In: HIV Sequence Compendium 2001. Edited by: Kuiken C, Foley B, Hahn B, Marx P, McCutchan F, Mellors JW, Wolinsky S, Korber B. Published by: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM, LA-UR 02-2877.

Kuiken C, Foley B, Leitner T, Apetrei C, Hahn B, Mizrachi I, Mullins J, Rambaut A, Wolinsky S, and Bette Korber editors. 2010. *HIV Sequence Compendium 2010*. Publisher: Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, New Mexico. LA-UR 10-03684.

Kumarasamy N, Madhavan V, Venkatesh KK, Saravanan S, Kantor R, Balakrishnan P, Devaleenal B, Poongulali S, Yepthomi T, Solomon S, Mayer KH, Benson C, and Schooley R. 2009. High Frequency of Clinically Significant Mutations after First-Line Generic Highly Active Antiretroviral Therapy Failure: Implications for Second-Line Options in Resource-Limited Settings. *Clinical Infectious Diseases*. 49:306–309.

Langford, SE, Ananworanich J, and Cooper DA. 2007. Predictors of Disease Progression in HIV Infection: a Reveiew. *AIDS Research and Therapy*. **4:11** doi:10.1186/1742-6405-4-11.

Larke N, Im E-J, Wagner R, Williamson C, Williamson A-L, McMichael AJ, and Hanke T. 2007. Combined Single-Clade Candidate HIV-1 Vaccines Induce T cell Responses Limited by Multiple forms of *In Vivo* Immune Interference. *European Journal of Immunology*. **37:**566–577. **6**:59 doi:10.1186/1471-2334-6-59.

Lawn SD, Myer L, Bekker LG, and Wood R. 2006. CD4 Cell Count Recovery among HIV-infected Patients Very with Adavnced Immunodeficiency Commencing Antiretroviral Treatment in Sub-Saharan Africa. *BMC Infectious Diseases*. **6:**59.

Le T, Chiarella J, Simen BB, Hanczaruk B, Egholm M, Landry ML, Dieckhaus K, Rosen MI, and Kozal MJ. 2009. Low Abundance Drug –Resistant Viral Variants in Treatment Experienced Persons Correlate with Historical Antiviral Use. *PLoS One*. **4(6):**e6079. doi:10.1371/journal.pone.0006079.

Learn GH, Muthui D, Brodie SJ, Zhu T, Diem K, Mullins JI, and Corey L. 2002. Virus Population Homogenization Following Acute Human Immunodeficiency Virus Type 1 Infection. *Journal of Virology.* **76:**11953–11959.

Lee JK, Stewart-Jones G, Dong T, Harlos K, Di Gleria K, Dorrell L, Douek DC, van der Merwe PA, Jones EY, and McMichael AJ. 2004. T Cell Cross-Reactivity and Conformational Changes During TCR Engagement. *Journal of Experimental Medicine*. **200:**1455–1466.

Lee S, Wood O, Tang S, Hu J, Machuca A, Kerby S, Awazi B, Vockley C, And Hewlett I. 2007. Detection of Emerging HIV Variants in Blood Donors from Urban Areas of Cameroon. *AIDS Research and Human Retroviruses*. **23(10):**1262-1267.

Lemey P, Kosakovsky Pond SL, Drummond AJ, Pybus OG, Shapiro B, Barroso H, Taveira N, and Rambut A. 2007. Synonymous Substitution Rates Predict HIV Disease Progression as a Result of Underlying Replication Dynamics. *PLoS Computational Biology*. **3(2):**e29. doi:10.1371/journal.pcbi.0030029

Letourneau S, Im EJ, Mashishi T, Brereton C, Bridgeman A, Yang H, Korber B, McMichael AJ, and Hanke T. 2007. Design and Preclinical Evaluation of a Universal HIV-1 Vaccine. *PLoS One*. **2(10):**e984.

Letvin NL, and Walker BD. 2003. Immunopathogenesis and Immunotherapy in AIDS Virus Infection. *Nature Medicine*. **9(7):**861-866.

Levy JA, Hoffman AD, Kramer SM, Landis JA, and Shimarukuro JM. 1984. Isolation of Lymphocytopathic Retroviruses from San Fransesco Patients with AIDS. *Science*. **225:**840-842.

Levy JA. 2009. HIV Pathogenesis: 25 Years of Progress and Persistent Challenges. *AIDS.* **23:**147–160.

Liu SL, Mittler JE, Nickle DC, Mulvania TM, Shriner D, Rodrigo AG, Kosloff B, He X, Corey, and Mullins I. 2002. Selection for immunodeficiency virus type 1 recombinants in a patient with rapid progression to AIDS. *Journal of Virology*. **76(21):**10674-10684.

Li JC, Lee DC, Cheung BK, Lau AS. 2005. Mechanisms for HIV Tat Upregulation of IL-10 and Other Cytokine Expression: Kinase Signaling and PKR-mediated Immune Response. *FEBS Letter*.;**579:**3055–3062.

Li JCB, HYim CH, and Lau ASY. 2010. Role of HIV-1 Tat in AIDS Pathogenesis: its Effects on Cytokine Dysregulation and Contributions to the Pathogenesis of Opportunistic Infection. *AIDS*. **24:**1609–1623.

Li Y, Svehla K, Louder MK, Wycuff D, Phogat S, Tang M, Migueles SA, Wu X, Phogat A, Shaw GM, Connors M, Hoxie J, Mascola JR, and Wyatt R. 2009. Analysis of Neutralization Specificities in Polyclonal Sera Derived from Human Immunodeficiency Virus Type 1-Infected Individuals. *Journal of Virology*. **83:**1045–1059.

Lihana RW, Khamadi SA, Lubano K, Lwembe R, Kiptoo MK, Lagat N, Kinyua JG, Okoth FA, Songok EM, Makokha EP, and Ichimura H. 2009a. HIV Type 1 Subtype Diversity and Drug Resistance among HIV Type 1-Infected Kenyan Patients Initiating Antiretroviral Therapy. *AIDS Research and Human Retroviruses*. **25(12):**1211-1217.

Lihana RW, Khamadi SA, Lwembe RM, Ochieng W, Kinyua JG, Kiptoo MK, Muriuki JK, Lagat N, Osman S, Mwangi JM, Okoth FA, and Songok EM. 2009b. The Changing Trend of HIV Type 1 Subtypes in Nairobi. *AIDS Research and Huan Retroviruses*. **25(3):**337-342.

Little SJ, Holte S, Routy JP, Daar E, Markowitz M, Collier AC, Koup RA, Mellors JW, Connick E, Conway B, Kilby M, Wang L, Whitcomb JM, Hellmann NS, and Richman DD. 2002. Antiretroviral-Drug Resistance among Patients Recently Infected with HIV. *The New England Journal of Medicine*. **342:**359-394.

Loannidis JPA, Havlir DV, Tebas P, Hirsch MS, Collier AC, and Richman DD. 2000. Dynamics of HIV-1 Viral Load Rebound among Patients with Previous Suppression of Viral Replication. *AIDS*. **14:**1481-1488.

Loemba H, Brenner B, Parniak MA, Ma'ayan S, Spira B, Moisi D, Oliveira M, Detorio M, Wainberg MA. 2002. Development of Resistance against Nonnucleoside Inhibitors of RT. *Antimicrobial Agents and Chemotherapy*. **46(7):**2087-2094.

Lofgren SM, Morrissey AB, Chevallier CC, Malabeja AI, Edmonds S, Amos B, Sifuna DJ, von Seidlein L, Schimana W, Stevens WS, Bartlett JA, and Crump JA. 2009. Evaluation of a Dried Blood Spot HIV-1 RNA Program for Early Infant Diagnosis and Viral Load Monitoring at Rural and Remote Healthcare Facilities. *AIDS*. **23:**2459–2466.

Lok JJ, Bosch RJ, Benson CA, Collier AC, Robbins GK, Shafer RW, and Hughes MD, for ALLRT team. 2010. Long-Term Increase in CD4+ T-cell Counts during Combinantion Antiretroviral Therapy for HIV-1 Infection. *AIDS*. **24:**1867-1876.

Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, and Ray SC. 1999. Full-Length Human Immunodeficiency Virus Type 1 Genomes from Subtype C-Infected Seroconverters in India, with Evidence of Intersubtype Recombination. *Journal of Virology*. **73(1):**152-160.

Loong TW. 2003. Understanding Sensitivity and Specificity with the Right Side of the Brain. *British Medical Journal*. **327:**716-719.

Loussert-Ajaka I, Ly TD, Chaix ML, Ingrand D, Saragosti S, Courouce AM, Brun-Ve´zinet F, and Simon F. 1994. HIV-1/HIV-2 Seronegativity in HIV-1 Subtype O Infected Patients. *Lancet*. **343:**1393–1394.

MacDonald KL, Jackson B, Bowman RJ, Polesky HF, Rhame FS, Balfour Jr. HH, and Osterhom MT. 1989. Performance Characteristics of Serologic Tests for Human Immunodeficiency Virus Type 1 (HIV-1) Antibody among Minnesota Blood Donors: Public Health and Clinical Implications. *Annals of Internal Medicine.* **110:**617-621.

Maes B, Schrooten Y, Snoeck J, Derdelinckx I, Ranst MV, Vandamme AM, Laethem KV.2004. Performance of ViroSeq HIV-1 Genotyping System in Routine Practice at a Belgian Clinical Laboratory. *Journal of Virological Methods*. **119:**45–49.

Maga G, Radi M, Gerard MA, Botta M, and Eric Ennifar E. 2010. HIV-1 RT Inhibitors with a Novel Mechanism of Action: NNRTIs that Compete with the Nucleotide Substrate. *Viruses.* **2:**880-899.

Mani I, Gilbert P, Sankalé JL, Eisen G, Mboup S, and Kanki PJ. 2002. Intrapatient diversity and its correlation with viral setpoint in human immunodeficiency virus type 1 CRF02_A/G-IbNG infection. *Journal of Virology*. **76(21):**10745-10755.

Marison L. 2001. The Global Epidemiology of HIV/AIDS. *British Medical Bulletin*. **58:**7-18.

Marozsan AJ, Moore DM, Lobritz MA, Fraundorf E, Abraha A, Reeves JD, and Art EJ. 2005. Differences in the Fitness of Two Diverse Wild-Type Human Immunodeficiency Virus Type 1 Isolates are Related to the Efficiency of Cell Binding and Entry. *Journal of Virology*. **79(11):** 7121-7134.

Martinez-Cajas JL, Pai NP, Klein MB, Wainberg MA. 2009. Differences in Resistance Mutations among HIV-1 Non-Subtype B Infections: A systematic Review of Evidence (1996–2008). *Journal of Internatonal AIDS Society*. **12:** 11. doi:10.1186/1758-2652-12-11

Martinez-Cajas JL, Pant-Pai N, Klein MB, Wainberg MA. 2008. Role of Genetic Diversity amongst HIV-1 Non-B Subtypes in Drug Resistance: A Systematic Review of Virologic and Biochemical Evidence. *AIDS Review.* **10:**212–223.

Maserati R, Silvestri AD, Uglietti A, Colao G, Biago AD, Bruzzone B, Pietro MD, Re MA, Tinelli C, and Zazzi M, on Behalf of the ARC Collaborative Group. 2010. Emeriging Mutations at Virological Failure of HAART Combinantions Containing Tenofovir and Lamivudine or Emitricitabine. *AIDS*. **24:**1013-1018.

Masciotra S, Garrido C, Youngpairo AS, McNulty A, Zahonero N, Corral A, Heneine W, de Mendoza C, and Garcia-Lerma JG. 2007. High Concordance between HIV-1 Drug Resistance Genotypes Generated from Plasma and Dried Blood Spots in Antiretroviral-Experienced Patients. *AIDS*. **21:**2503-2511.

Maslin J, Rogier C, BergerF, Khamil MA, Mattera D, Grandadam M, Caron M, and Nicand E. 2005. Epidemiology and Genetic Characterization of HIV-1 Isolates in the General Population of Djibouti (Horn of Africa). *Journal of Acquired Immune Deficiency Syndrome*. **39(2):**129-132.

Mastro TD, Kunanusont C, Timothy J. Dondero TJ, and Wasi C. 1997. Why Do HIV-1 Subtypes Segregate Among Persons with Different Risk Behaviors in South Africa and Thailand? *AIDS.* **11:**113-116.

Matamoros T, Nevot M, Martínez MA, and Menendez-Arias L. 2009. Thymidine Analogue Resistance Suppression by V75I of HIV-1 Reverse Transcriptase: Effects of Substituting Valine 75 on Stavudine Excision and Discrimination. *The Journal of Biological Chemistry*. **284(47):**32792–32802.

Matheron S, Pueyo S, Damond F, Simon F, Lepretre A, Campa P, Salamon R, Chene G, Brun-Vezinet F, and French HIV-2 Cohort Study Group. 2003. Factors Associated with Clinical Progression in HIV-2 Infected Patients: the French ANRS Cohort. *AIDS*. **17(18):**2593-2601.

May M, Boulle A, Phiri S, Messou E, Myer L, Keiser O, Sterne JAC, Dabis F, and Egger M, for leDEA Southern Africa and West Africa. 2010. Prognosis of Patients with HIV-1 Infection Starting Antiretroviral Therapy in Sub-Saharan Africa: a Collaborative Analysis of Scale-up Programs. *The Lancet*. DOI:10.1016/S0140-6736(10)60666-6.

McCutchan FE. 2006. Global Epidemiology of HIV. *Journal of Medical Virology*. **78:**S7-S12.

McMichael AJ. 2006. HIV Vaccines. *Annul Reviews of Immunology*. **24:**227–255.

McNulty A, Jennings C, Bennett D, Fitzgibbon J, Bremer JW, Ussery M, Kalish ML, Heneine W, and Garcıa-Lerma JG. 2007. Evaluation of Dried Blood Spots for Human Immunodeficiency Virus Type 1 Drug Resistance Testing. *Journal of Clinical Micobiology*. **45(2):**517-521.

Medeiros MS, Arruda EAG, Guerrant RL, Brown CC, and Lima AM. 2007. Impact of the Number of Failed Therapeutic Regimens on the Development of Resistance Mutations to HIV-1 in North Brazil. *The Brazilian Journal of Infectious Diseases*. **11(5):**451-455.

Mehret M, Khodakevich L, Zewdie D, Ayehunie S, Gizaw G, Shanko B, *et al*. 1990a. HIV-1 Infection and Related Risk Factors among Female Sex Workers in the Urban Areas of Ethiopia. *Ethiopian Journal of Health Development*. **4:**163-170.

Mehret M, Khodakevich L, Zewdie D, Gizaw G, Ayehunie S, Shanko B *et al*. 1990b. HIV-1 Infection among eEmployees of the Ethiopian Freight Transport Corporation. *Ethiopian Journal of Health Development*. **4:**177-182.

Mekonnen Y, Geskus RB, Hendriks JCM, Messele T, Borghans J, Miedema F, Wolday D, Coutinho RA, and Dukers NHTM. 2005b. Low CD4 T Cell Counts before HIV-1 Seroconversion Do Not Affect Disease Progression in Ethiopian Factory Workers. *Journal of Infectious Diseases*. **192:**739–748.

Mekonnen Y, Sanders E, Messele T, Wolday D, Dorigo-Zestma W, Schaap A, Mekonnen W, Meless H, Mihret W, Fontanet A, Coutinho RA, and Dukers NHTM. 2005a. Prevalence and Incidence of, and Risk Factors for, HIV-1 Infection among Factory Workers in Ethiopia, 1997-2001. *Journal of Health, Population and Nutrition.* **23(4):**358-368.

Metzner KJ, Rauch P, Walter H, Boesecke C, Zollner B, Jessen H, Schewe K, Fenske S, Gellermann H, and Stellbrink HJ. 2005. Detection of Minority populations of Drug-Resistant HIV-1 in Acute Seroconversion. *Antiviral Therapy*. **10:**S123.

Migulues SA, and Connors M. 2010. Long-term Nonprogressive Disease Among Untreated HIV-Infected Individuals: Clinical Implications of Understanding Immune Control of HIV. *Journal of American Medical Association (JAMA).* **304(2):**194-201.

Mild M, Kvist A, Esbjornsson J, Karlsson I, Fenyo¨ EM, Medstrand P. 2010. Differences in Molecular Evolution between Switch (R5 to R5X4/X4-tropic) and Non-Switch (R5-tropic only) HIV-1 Populations During Infection. *Infection, Genetics and Evolution*. **10:**356-364.

Miller RJ, Cairns JS, Bridges S, and Sarver N. 2000. Human Immunodeficiency Virus and AIDS: Insight from Animal Lentiviruses. *Journal of Virology*. **74(16):**7187-7195.

Miller V, Stark T, Loeliger A, and Lange J. 2002. The Impact of the M184V Substitution in HIV-1 Reverse Transcriptase on Treatment Response. *HIV Medicine*. **3:**135-145.

Ministry of Health (MOH). 2005. *Guideline for Implementation of Antiretroviral Therapy in Ethiopia*. January 2005. Ministry of Health. Addis Ababa. Available at: http://www.etharc.org/ arvinfo/ethARTgiude.pdf.

MOH/HAPCO. 2010. *ART monthly updates*. Available at http://www.etharc.org/resources/download/ Accessed on December, 2010.

MOH. 2004. *National Guidelines for ANC-Based HIV Surveillance*. HIV/AIDS and Other STIs Prevention and Control Team, Disease Prevention and Control Department, Ministry of Health, Addis Ababa.

MOH/FHAPCO. 2007a. *Single Point HIV Prevalence Estimate*. June 2007. Addis Ababa, Ethiopia. Available at: http://www.etharc.org/resources/download/

MOH/HAPCO. 2006. *AIDS in Ethiopia. 6th Report*. September 2006. Addis Ababa, Ehiopia. Availalble at: http://www.etharc.org./aidsinEth6th_en.pdf.

MOH/HAPCO. 2007b. *Guidelines for Management of Opportunistic Infections and Antiretroviral Treatment in Adolescents and Adults in Ethiopia*. July 2007. Addi Ababa, Ethiopia.Availalble at: http://www.etharc.org/resources/download/

Molla A, Korneyeva M, Gao Q, Vasavanonda S, Schipper PJ, MO HM, Markowitz M, Chernyavskiy T, Niu P, Lyons N, Hsu A, Granneman GR, Ho DD, Boucher CAB, Leonard JM, Norbeck DW, and Kempf DJ. 1996. Ordered Accumulation of Mutations in HIV Protease Confers Resistance to Ritonavire. *Nature Medicine*. **2(7):**760-766.

Monforte AA. Sabin CA, Phillips A, Sterne J, May M, Dabis F, Grabar S, Ledergerber B, Gill J, Reiss P, and Egger M. 2005. The Changing Incidence of AIDS Events in Patients Receiving Highly Ative Antiretroviral Therapy. *Archives of Internal Medicine*. **165:**416-423.

Montano MA, Nixon CP, Ndung'u T, Bussmann H, Novitsky VA, Dickman D, and Essex M. 2000. Elevated Tumor Necrosis Factor-alpha Activation of Human Immunodeficiency Virus Type 1 Subtype C in Southern Africa is Associated with an NF-kappaB Enhancer Gain-of-function. *The Journal of Infectious Diseases*. **181(1):**76–81.

Moore JP, Cao Y, Leu J, Qin L, Korber B, and Ho DD. 1996. Inter- and Intraclade Neutralization of Human Immunodeficiency Virus type 1: Genetic Clades Do Not Correspond to Neutralization Serotypes But Partially Correspond to gp120 Antigenic Serotypes. *Journal of Virology*. **70:**427-444.

Moroni M, and Antinori S. 2003. HIV and Direct Damage of Organs: Disease Specrum Before and During the Highly Active Antiretroviral Therapy Era. *AIDS*. **17(suppl 1):**S51-S64.

Mosebi S, Morris L, Dirr HW, and Sayed Y. 2008. Active-Site Mutations in the South African Human Immunodeficiency Virus Type 1 Subtype C Protease Have a Significant Impact on Clinical Inhibitor Binding: Kinetic and Thermodynamic Study. *Journal of Virologgy*. **82(22):** 11476-11479.

Motomura K, Chen J, and Hu WS. 2008. Genetic Recombination between Human Immunodeficiency Virus Type 1 (HIV-1) and HIV-2, Two Distinct Human Lentiviruses. *Journal of Virology*. **82(4):**1923-1933.

Moutouch L, Corbeil J, and Richman DD. 1996. Recombination Leads to the Rapid Emergence of HIV-1 Dually Resistant Mutants under Selective Drug Pressure. *Proceedings of National Academy of Science (PNAS)*. **93:**6106-6111.

Moyle G. 200. Resistance and Cross-Resistance to Abacavir. *HIV Medicine*. **2:**154-162.

Mulatu MS, Kloos H, Converse P, Wuhib T, and Pankhurst A. 2005. Bibliography on HIV/AIDS in Ethiopia and Ethiopians in the Diaspora: The 2004 Update. *Ethiopian Journal of Health Development*. **19(1):**65-85.

Mulatu MS, Kloos H, Converse PJ, and Haile Mariam D. 2007. Bibliography on HIV/AIDS in Ethiopia and Ethiopians in the Diaspora: The 2006 Update. *Ethiopian Journal of Health Development*. **21(1):**70-91.

Myatt M, and Bennett D. 2008. A Novel Sequencial Sampling Technique for the Surveillance of Transmitted HIV Drug Resistance by Cross-Sectional Survey for Use in Low Resource Settings. *Antiviral Therapy*. **13(Suppl 2):**37-48.

Naghavi MH, Schwartz S, Sonnerborg A, and Vahlne A. 1999. Long Terminal Repeat Promoter/Enhancer Activity of Different Subtypes of HIV Type 1. *AIDS Research and Human Retroviruses*. **15:**1293-303.

Nijhuis M, van Manarseveen N, Shhipper B Glass, Dulude D, de Jong D, Goedegebuure I, Brakier-Gingras L, Kraeusslich HG, and Boucher C. 2005. Novel HIV gag Based Protease

Drug-Resistnce Mechanism Caused by an Increased Processing of NC/p1 Cleavage Site. *Antiviral Therapy*. **10:**S117.

Nikolenko GN, Delviks-Frankenberry KA, and Pathak KA. 2010. A Nobel Molecular Mechanism of Dual Resistance to Nucleoside and Nonnucleoside Reverse Transcriptase Inhibitors. *Journal of Virology*. **48(10):**5238-5249.

Nikolenko GN, Frankenberry KA, Palmer S, Maldarelli F, Mellors JW, Coffin JM, and Pathak VK. 2005. RNase H Domains Obtained from Treatment Experienced Patients Increase Resistance to AZT. *Antiviral Therapy.* **10:**S89.

Novitsky V, Gaolathe T, Woldegabriel E, Makhema J, and Essex M. A. 2007. Seronegative Case of HIV-1 Subtype C Infection in Botswana. *Clinical Infectious Diseases*. **45:**e68–71.

Novitsky V, Smith UR, Gilbert P, McLane MF, Chigwedere P, Williamson C, Ndung'u T, Klein I, Chang SY, Peter T, Thior, I, Foley BT, Gaolekwe S, Rybak N, Gaseitsiwe S, Vannberg F, Marlink R, Lee TH, and Essex M. 2002. Human immunodeficiency virus type-1 subtype C molecular phylogeny: consensus sequence for an AIDS vaccine design? *Journal of Virol*ogy **76(11):**5435-5451.

Nyambi PN, Nadas A, Mbah HA, Burda S, Williams C, Gorny MK, and Zolla-Pazner. 2000. Immunoreactivity of Intact Virions of Human Immunodeficiency Virus Type 1 (HIV-1) Reveals the Existence of fewer HIV-1 Immunotypes than Genotypes. *Journal of Virology*. **74:**10670–10680.

Oh J, McWilliams J, Julias JG, and Hughes SH. 2008. Mutations in the U5 Region Adjacent to the Primer Binding Site Affect tRNA Clavage by Human Immunodeficiency Virus Type 1 Reverse Transcriptatse *In Vivo*. *Journal of Virology*. **82(20):**719-727.

Osmanov S, Pattou C, Schwardlander B, Esparza J, and the WHO-UNAIDS Network for HIV Isolation and Characterization. 2002. Estimated Global Distribution and Regional Spread of HIV-1 Genetic Subtypes in the Year 2000. *Journal of Acquired Immune Deficiency Syndrome*. **29:**184-190.

Owen SM, Yang C, Spira T, Ou CY, Pau CP,Parekh BS, Candal D, Kuehl D, Kennedy MS, Rudolph D, Lou W, Delatorre N, Masciotra S, Kalish ML, Cowart F, Barnett T, Lal R, and McDougal JS. 2008. Alternative Algorithms for Human Immunodeficiency Virus Infection Diagnosis Using Tests That Are Licensed in the United States. *Journal of Clinical Microbiology*. **46(5):**1588–1595.

Oyaro M, Mbithi J, Oyugi F, Laten A, Anzala O, and Engelbrecht S. 2011. Molecular Characterization of HIV Type 1 Among HIV-Infected Respondents in a Cohort Being Prepared for HIV Phase III Vaccine Clinical Trials, Western Kenya. *AIDS Research and Human Retroviruses*. **27(3):**257-264. doi:10.1089/aid.2010.0061.

Palmer S, Alaeus A, Albert J, and Cox S. 1998. Drug Susceptibility of Subtypes A, B, C, D, and E Human Immunodeficiency Virus Type 1 Primary Isolates. *AIDS Research and Human Retroviruses.* **14:**157-162.

Palmer S, Kearney M, Maldarelli F, Halvas EK, Bixby CJ, Bazmi H, Rock D, Falloon J, Davey RT, Dewar RL, Metcalf JA, Hammer S, Mellors JW, and Coffin JM. 2005. Multiple, Linked Human Immunodeficiency Virus Type 1 Drug Resistance Mutations in Treatment-Experienced Patients are Missed by Standard Genotype Analysis. *Journal of Clinical Microbiology*. **43(1):**406-413.

Pantaleo G, Demarest JF, Schacker T, Vaccarezza M, Cohen OJ, Daucher M, Graziosi C, Schnittman SS, Quinn TC, Shaw GM, Perrin L, Tambussi G, Lazzarin A, Sekaly RP, Soudeyns H, Corey L, and Fauci AS. .1997. The Qualitative Pature of the primary immune Response to HIV Infection is a Prognosticator of Disease Progression Independent of the Initial Level of Plasma Viremia. *Proceedings of National Academy of Science U S A.* **94(1):** 254–258. *Journal of Acquired Immune Deficiency Syndrome*.Parczewski M, Leszczyszyn-Pynka M, Bander D, Urbańska A, Stańczak G, and Boroń-Kaczmarska A. 2010. Characteristics of HIV-1 Non-B Subtype Infections in Northwest Poland. *Journal of Medical Virology*. **82:**1306–1313.

Parkin N, Chappey C, Lam E, and Petropoulos C: 2005. Reduced Susceptibility to Protease Inhibitors (PI) in the Absence of Primary PI-Resistance-Associated Mutations. *Antiviral. Therapy.* **10:** S118 Abstract 108.

Parreira R, Padua E, Piedade J, Teresa Venenno T, Paixao MT, and Esteves A. 2005. Genetic Analysis of Human Immunodeficiency Virus Type 1 nef in Portugal: Subtyping, Identification of Mosaic Genes, and Amino Acid Sequence Variability. *Journal of Medical Virology*. **77:**8–16.

Pastore C, Nedellec R, Ramos A, Pontow S, Ratner L, and Mosier DE. 2006. Human Immunodeficiency Virus Type 1 Coreceptor Switching: V1/V2 Gain-of-Fitness Mutations Compensate for V3 Loss-of-Fitness Mutations. *Journal of Virology*. **80(2):**750–758.

Patel MB, Hoffman NG, and Swanstrom R. 2008. Subtype-Specific Conformational Differences within the V3 Region of Subtype B and Subtype C Human Immunodeficiency Virus Type 1 Env Proteins. *Journal of Virology*. **82(2):**903–916.

Peeters M, Aghokeng AF, and Delaporte E. 2010. Genetic Diversity among Human Immunodeficiency Virus-1 Nnon-B Subtypes in Viral Load and Drug Resistance Assays. *Clinical Microbiology and Infection*. **16:**1525–1531.

Peeters M. 2000. *Recombinant HIV Sequences: Their Role in the Global Epidemic. pp. I-39-54 in HIV Sequence Compendium 2000.* Edited by: Kuiken CL, Foley B, Hahn B, Korber B, McCutchan F, Marx PA, Mellors JW, Mullins JI, Sodroski J, and Wolinsky S. Published by: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM.

Perrin L, Kaiser L, and Yerly S. 2003. Travel and the Spread of Genetic Variants. *Lancet Infectious Diseases*. **3:**22-27.

Pilcher CD, J. J. Eron Jr. JJ, Galvin S, Gay C, and Cohen MS, 2004. Acute HIV Revisited: New Opportunities for Treatment and Prevention. *Journal of Clinical Investigation.* **113(7):**937–945.

Pilger D, Hauser A, Kuecherer C, Mugenyi K, Kabasinguzi R, Somogyi S, Harms G, and Kunz A. 2010. Minor Drug-Resistanct HIV Type-1 Variants in Breast Milk and Plasma of HIV Type-1-Infected Ugandan Women after Nevirapine Single-Dose Prophylaxis. *Antivirus Therapy*. **16:doi:**in press 10.3851/IMP1698.

Plantier JC, Leoz M, Dickerson JE, Oliveira FD, Cordonnier F, Lemee V, Damond F, Robertson DL, and Simon F. 2009. A New Human Hmmunodeficiency Virus Derived from Gorillas. *Nature Medicine*. **15(8):**871-872.

Pond SLK, and Smith DM. 2009. Are all Subtypes Created Equal? The Effectiveness of Antiretroviral Therapy against Non-Subtype B HIV-1. *Clinical Infectious Disease*. **48:**1306-1309.

Pope M, and Haase AT. 2003. Transmission, Acute HIV-1 Infectiion and the Quest for Strategies to Prevent Infection. *Nature Medicine*. **9(7):**847-852.

Porter KA, Kelley LN, Goerge A, Harton JA, and Duus KM. 2010. Class II Transactivator (CIITA) Enhances Cytoplasmic Processing of HIV-1 Pr55Gag. PLoS ONE **5(6):** e11304. doi:10.1371/journal.pone.0011304.

Preiser W, and Stephen Korsman. 2007. HIV Testing. In: *HIV Medicine 2007*. Edited by Christian Hoffman, JK Rockstro, and BS Kamps. 15th Edition. Flying Publisher. Available at: www.HIVMedicine.com. Pp:41-57.

Pujades-Rodrıguez M, Balkan S, Arnould L, Brinkhof MAW, Calmy A, and for the AIDS Working Group of MSF. 2010. Treatment Failure and Mortality Factors in the Patients Receiving Second Line HIV-Therapy in Resource-Limited Countries. *Journal of American Medical Association (JAMA)*. **304(3):**303-312.

Quiñones-Mateu ME, and Arts EJ. 2001. HIV-1 Fitness: Implications for Drug Resistance, Disease Progression, and Global Epidemic Evolution. In: *HIV Sequence Compendium 2001*. Edited by: Kuiken C, Foley B, Hahn B, Marx P, McCutchan F, Mellors JW, Wolinsky S, Korber B. Published by: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM, LA-UR 02-2877. pp. 134-170.

Ravela J, Betts BJ, Brun-Vezinet F, Vandamme AM, Descamps D, Laethem KV, Smith K, Schapiro JM, Winslow DL, Reid C, and Shafer RW. 2003. HIV-1 Protease and Reverse Transcriptase Mutation Patterns Responsible for Discordances between Genotypic Drug Resistance Interpretation Algorithms. *Journal of Acquired Immune Deficiency Syndrome*. **33:**8–14.

Ravichandran S, Veerasamy R, Raman S, Krishnan PN, and Agrawal RK. 2008. An Overview on HIV-1 Reverse Transcriptase Inhibitors. *Digest Journal of Nanomaterials and Biostructures*. **3(4):** 171-187.

Reniers G, Araya T, Davey G, Nagelkerke N, Berhane Y, Coutinho R, and Sanders EJ. 2009. Steep Declines in Population-Level AIDS Mortality Following the Introduction of Antiretroviral Therapy in Addis Ababa, Ethiopia. *AIDS*. **23:**511-518.

Renjifo B, Fawzi W, Mwakagile D, Hunter D, Msamanga G, Spiegelman D, Miriam G, Charles K, Anne K, Beth C, Ellen H, and Max E. 2001. Differences in Perinatal Transmission Among Human Immunodeficiency Virus Type 1 Genotypes. *Journal of Human Virology*. **4:**16–25.

Renjufi B, Gilbert P, Chaplin B, Msamanga G, Mwakagile D, Fawzi W, Tanzanian Vitamin and HIV Study Group, and Essex M. 2004. Preferential In-Utero Transmission of HIV-1 Subtype C as Compared to HIV-1 Subtype A or D. *AIDS*. **18(12):**1629–1636.

Reuman EC, Rhee SH, Holmes SP, and Shafer RW. 2010. Constrained Pattern of Covariation and Clustering of HIV-1 Non-nucleoside Reverse Transcriptase inhibitor Resistance Mutations. *Journal of Antimicrobial Chemotherapy*. **65:**1477-1487.

Reynolds SJ, Ndongala LM, Luo CC, Mwandagalirwa K, Losoma AJ, Mwamba KJ, Bazepeyo E, Nzilambi NE, Quinn TC, and Bollinger RC. 2002. Evaluation of a Rapid Test for the Detection of Antibodies to Human Immunodeficiency Virus Type 1 and 2 in the Setting of Multiple Transmitted Viral Subtypes. *International Journal of STD and AIDS*. **13(3):**171-173.

Rhee SY, Liu TF, Holmes SP, Shafer RW. 2007. HIV-1 Subtype B Protease and Reverse Transcriptase Amino Acid Variation. *PLoS Computational Biology*. **3(5):** e87. doi:10.1371/journal.pcbi.0030087.

Richman DD, Morton SC, Wrin T, Hellman N, Berry S, Shapiro MF, and Bozzette SA. 2004. The Prevalence of Antiretroviral Drug Resistance in the United States. *AIDS.* **18:**1393–1401.

Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, and Funkhouser RK, Gao F, Hahn BH, Kuiken C, Learn GH, Leitner T, McCutchan F, Osmanov S, Peeters M, Pieniazek D, Kalish ML, Salminen M, Sharp PM, Wolinsky S, and Korber B. 2000. HIV-1 Nomenclature Proposal. *Science.* **288:**55–56.

Rodenburg CM, Li Y, Trask SA, Chen Y, Decker J, Robertson DL, Kalish ML, Shaw GM, Allen S, Hahn BH, and Gao F. 2001. Near Full-Length Clones and Reference Sequences for Subtype C Isolates of HIV Type 1 from three Different Continents. *AIDS Research and Human Retroviruses*. **17:**161-168.

Rodriguez RJ, Dayhoff DE, Chang G, Cassol SA, Birx DL, Artenstein AW, and Nelson L. Comparison of Serum and Plasma Viral RNA. 1997. Measurements in Primary and Chronic Human Immunodeficiency Virus Type 1 Infection. *Journal of Acquired Immune Deficiency Syndrome and Human Retrovirology*. **15(1):**49-53.

Rolland M, Nickle DC, and Mullins JI. 2007. HIV-1 Group M Conserved Elements Vaccine. *PLoS Pathogens*. **3(11):**e157.

Rosina M,Micheli V, Besan S, Toia C, Gismondo MR, Cargnel A. 2000. Genotyping in HIV-Drug Resistance Mutations: Epidemiology in 145 Patients. *International Journal of Antimicrobial Agents*. **16:**435-440.

Rouet F, Chaix ML, Nerrienet E, Ngo-Giang-Huong N, Plantier JC, Burgard M, Peeters M, Damond F, Ekouevi DK, Msellati P, Ferradini L, Rukobo S, Marechal V, Schvachsa N, Wakrim L, Rafalimanana C, Rakotoambinina R, Viard JP, Seigneurin JM, Rouzioux C, for the Agence Nationale de Recherches sur le SIDA AC11/AC12 Working Groups. 2007. Impact of HIV-1 Genetic Diversity on Plasma HIV-1 RNA Quantification: Usefulness of the Agence Nationale de Recherches sur le SIDA Second Generation Long Terminal Repeat-based Real-time Reverse Transcriptase Polymerase Chain Reaction Test. *Journal of Acquired Immune Deficiency Syndrome*. **45:**380–388.

Rubbert A, Behrens G, and Ostrowski M. 2007. Pathogenesis of HIV-1 Infection. In: *HIV Medicine 2007*. Edited by Christian Hoffman, JK Rockstro, and BS Kamps. 15th Edition. Flying Publisher. Available at: www.HIVMedicine.com. Pp: 59-86.

Salminen M, Johansson B, Sonnerborg A, Ayehunie A, Gotte D, Leinikki P, Burke DS, and McCutchan FE. 1996. Full-Length Sequence of an Ethiopian HIV-1 Isolate of Genetic Subtype C. *AIDS Research and Human Retroviruses* **12:**1329–1339.

Sanders EJ, Araya T, Kebede D, Schaap AJ, Nagelkerke ND, and Coutinho RA. 2003a. Mortality Impact of AIDS in Addis Ababa, Ethiopia. *AIDS*. **17(8):**1209-1216.

Sanders EJ, Messele T, Wolday D, Dorigo-Zetsma W, WoldeMichael T, Geyid A, Coutinho R. 2003b. Development of Research Capability in Ethiopia: the Ethio-Netherlands AIDS Research Project (ENARP): 1994-2002, Achievements, Scientific Findings and Project Goals. *Ethiopian Medical Journal*. **41(Suppl 1):**11-23.

Santiago ML, Rodenburg CM, Kamenya S, Bibollet-Ruche F, Gao F, Bailes E, Meleth S, Soong SJ, Kilby JM, Moldoveanu Z, Fahey B, Muller MN, Ayouba A, Nerrienet E, McClure HM, Heeney JL, Pusey AE, Collins DA, Boesch C, Wrangham RW, Goodall J, Sharp PM, Shaw GM, and Hahn BH. 2002. SIVcpz in Wild Chimpanzees. *Science*. **295:**465.

Sarafianos SG, Marchand B, Das K, Himmel DM, Michael A. Parniak MA, Hughes SH, and Arnold E. 2009. Structure and Function of HIV-1 Reverse Transcriptase: Molecular Mechanisms of Polymerization and Inhibition. *Journal of Molecular Biology.* **385:**693-713.

Saunders CJ, McCaffrey RA, Zharkikh I, Kraft Z, Malenbaum SE, Burke B, Cheng-Mayer C, and Stamatatos L. 2005. The V1, V2, and V3 Regions of the Human Immunodeficiency Virus Type 1 Envelope Differentially Affect the Viral Phenotype in an Isolate-Dependent Manner. *Journal of Virology*. **79(14):** 9069-9080.

Schaeffer E, and Greene WC. 2003. Efficiency of Host Cell Entry is the Dominant Factor Mediating *Ex Vivo* HIV-1 Fitness. *Antiviral Therapy*. .**8(suppl. 1):**S209.

Schindler M, Wurfl S, Benaroch P, Greenough TC, and Daniels R. 2003. Down-Modulation of Mature Major Histocompatibility Complex Class II and Up-Regulation of Invariant Chain Cell Surface Expression Are Well-Conserved Functions of Human and Simian Immunodeficiency Virus *nef* Alleles. *Journal of Virology*. **77(19):** 10548–10556.

Senkaali D, Muwonge R, Morgan D, Yirrell D, Whitworth J, and Kaleebu P. 2004. The Relationship between HIV Type 1 Disease Progression and V3 Serotype in a Rural Ugandan Cohort. *AIDS Research and Human Retroviruses*. **20(9):**932-937.

Shafer LA, Nsubuga RN, White R, Mayanja BN, Chapman R, O'Brien K, der Paal LV, Grosskurth H, and Maher D. 2011. Antiretroviral Therapy and Sexual Behavior in Uganda: a Cohort Study. *AIDS*. **25:**671-678.

Shafer RW, and Shapiro JM. 2008. HIV-1 Drug Resistance Mutations: an Updated Framework for the Second Decade of HAART. *AIDS Reviews*. **10(2):**67-84.

Sharp PM, Bailes E, Chaudhuri RR, Rodenburg CM, Santiago MO, and Hahn BH. 2001. The Origin of Acquired Immune Deficiency Syndrome Viruses: Where and When. *Philosophical Transactions of the Royal Society of London.* **356:**867-876.

Sharp PM, Shaw GM, and Hahn BH. 2005. Simian Immunodeficiency Virus Infection of Chimpanzees. *Journal of Virology*. **79(7):** 3891–3902.

Sherefa K, Johansson B, Salminen M, and Sonnerborg A. 1998. Full-Length Sequence of Human Immunodeficiency Virus Type 1 Subtype A, Recombined with Subtype C in the *env* V3 Domain. *AIDS Research and Human Retroviruses.* **14:**289–292.

Sherefa K, Sallberg M, and Sonnerborg A. 1994. Evidence of No Change in V3 Loop Antibody Recognition Pattern in HIV-1 Infected Ethiopians between 1988 and 1993. *AIDS Research and Human Retroviruses*. **10:**1551–1556.

Sherefa K, Sallberg M, Johansson B, Salminen M, Sand onnerborg A. 1997. Subtyping of Human Immunodeficiency Virus type 1 Strains by Using Antibodies Specific for the Third Variable Domain (V3) of gp120: Results May be Affected by Divergent V3 Sequences. *Journal of Clinical Microbiology.* **35:**2419–2421

Simon F, Mauclere P, Roques P, Loussert-Ajaka I, Muller-Trutwin MC, Saragosti S, Georges-Courbo MC, Barre-Sinouss F, and Brun-Vezinet F. 1998. Identification of a New Human

Immunodeficiency Virus type 1 Distinct from Group M and group O. *Nature Medicine*. **4:**1032–1037

Simon V, and Ho DD. 2003. HIV-1 Dynamics *In Vivo*: Implications for Therapy. *Nature Reviews Microbiolgy*. **1:**181-190

Simon V, Ho DD, and Karim QA. 2006. HIV/AIDS Epidemiology, Pathogenesis, Prevention, and Treatment. *Lancet*. **368:** 489–504.

Slobod K, Coleclough C, Brown SA, Stambas J, Zhan X, Surman S, Jones BG, Zirkel A, Freiden PJ, Brown B, Sealy R, Bonsignori M, and Hurwitz JL. 2005. Clade, Country and Region-specific HIV-1 Vaccines: Are They Necessary? *AIDS Research and Therapy*. **2**:3 doi:10.1186/1742-6405-2-3.

Snoeck J, Riva C, Steegen K, Schrooten Y, Maes B, Vergne L, Van Laethem K, Peeters M, Vandamme AM. 2005. Optimization of a Genotypic Assay Applicable to All Human Immunodeficiency Virus Type 1 Protease and Reverse Transcriptase Subtypes. *Journal of Virological Methods*. **128:** 47-53.

Soares MA. 2008. Drug Resistance Differences Among HIV Types and Subtypes: Classification and Global Distribution of HIV. *Future HIV Therapy.* http://cme.medscape.com/viewarticle/583361. Posted: 11/01/2008; Updated: 11/21/2008.

Solomon JA, and Murray CJL. 2001. Modeling HIV/AIDS Epidemics in Sub-Saharan Africa Using Seroprevalence Data from Antenatal Clinics. *Bulletin of the World Health Organization* **79:**596–607.

Sönnerborg A, and Sherefa K. 1997. HIV-1 Subtype C Epidemic in Ethiopia. *AIDS*. **11:**1897.

Spina C, Prince HE, and Richman DD. 1997. Preferential Replication of HIV-1 in the CD45RO Memory Cell Subset of Primary CD4 Lymphocytes *In Vitro*. *Journal of Clinical Investigation*. **99:**1774-1785.

Spira S, Wainberg MA, Loemba H, Turner D, and Brenner BG. 2003. Impact of Diversity on HIV-1 Virulence, Antiviral Drug Sensitivity and Drug Resistance. *Journal of Antimicrobial Chemotherapy*. **51:**229-240.

Stebbing J, Gazzard B, and Douek DC. 2004. Where Does HIV Live? *The New England Journal of Medicine*. **350:**1872-1880.

Stevenson M. 2003. HIV-1 Pathogenesis. *Nature Medicine*. **9(7):**853-860.

Strebel K. 2003. Virus-Host Interaction: Role of HIV Protiens Vif, Tat and Rev. *AIDS* **17(suppl 4):**S25-S34.

Stumptner-Cuvelette P, Morchoisne S, Dugast M, Gall SL, a Raposo G, Schwartz O, and Benaroch P. 2001. HIV-1 Nef Impairs MHC class II Antigen Presentation and Surface Expression. *Proceedings of the National Academy of Science (PNAS).* **98(21):**12144–12149.

Su L, Graf M, Zhang Y, von Briesen H, Xing H, Köstler J, Melzel H, Wolf H, Shao Y, and Wagner R. 2000. Characterization of a Virtually Full-Length Human Immunodeficiency Virus Type 1 Genome of a Prevalent Intersubtype (C/B') Recombinant Strains in China. *Journal of Virology*. **74(23):**11367-11376.

Suhasini M, and Reddy R. 2009. Cellular Protiens and HIV-1 Rev Function. *Current HIV Research*. **7:**91-100.

Swanson P, de Mendoza C, Joshi Y, Golden A, Hodinka RL, Soriano V, Devare SG, and Hackett Jr. J. 2005. Impact of Human Immunodeficiency Virus Type 1 (HIV-1) Genetic Diversity on Performance of Four Commercial Viral Load Assays: LCx HIV RNA Quantitative, AMPLICOR HIV-1 MONITOR v1.5, VERSANT HIV-1 RNA 3.0, and NucliSens HIV-1 QT. *Journal of Clinical Microbiology*. **43:**3860–3868.

Swanson P, Soriano V, Devare Sg, and Hackett Jr. J. 2001. Comparative Performance of Three Viral Load Assays on Human Immunodeficiency Virus Type 1 (HIV-1) Isolates Representing Group M (Subtypes A to G) and Group O: LCx HIV RNA Quantitative, AMPLICOR HIV-1 MONITOR Version 1.5, and Quantiplex HIV-1 RNA Version 3.0. *Journal of Clinical Microbiology*. **39(3):** 862–870.

Takehisa K, Kraus MH, Ayouba A, Bailes E, Van Heuverswyn F, Decker JM, Li Y, Rudicell RS, Learn GH, Neel C, Ngole EM, Shaw GM, Peeters M, Sharp PM, and Hahn BH. 2009. Origin and Biology of Simian Immunodeficiency Virus in Wild-Living Western Gorillas. *Journal of Virology*. **83(4):**1635-1648.

Tamura K, Dudley J, Nei M, and Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Molecular Biology and Evolution.* **24:**1596-1599.

Tapia N, Franco S, Puig-Basagoiti F, Menendez C, Alonso PL, Mshinda H, Clotet B, Saiz JC, and Martinez MA. 2003. Influence of Human Immunodeficiency Virus Type 1 Subtype on Mother-to-Child Transmission. *Journal of General Virology*. **84:**607–613.

Taylor BS, Sobieszczyk ME, McCutchan, and Hammer SM. 2008. The Challenge of HIV-1 Subtype Diversity. *The New England Journal of Medicine.***358:**1590-1602.

Tebit DM, and Arts E. 2011. Tracking a Century of Global Expansion and Evolution of HIV to Drive Understanding and to Combat Disease. *Lancet Infectious Diseases*. **11:**45-56.

Tegbaru B, Meless H, Kume A, Messele T, Hailu A, Mekonnen Y, and Fontanet AL. 2004. Nine-year Trends in HIV-1 Prevalence among Visa applicants in Urban Ethiopia. *Ethiopian Medical Journal*. **42:**41-47.

Theys K, and Vandamme AM. 2008. Undestanding the Changing Prevalence of K65R. *AIDS Reviews*. **10:**125.

Thompson MA, Aberg JA, Pedro G, Montaner JSG, Rizzadini G, Telenti A, Gatell JM, Gunthard HF, Hammer SM, Hirsch MS, Jacobsen DM, Richman DD, Volberding PA, Yeni P, and Schooley RT. 2010. Antiretroviral Treatment of Adult HIV Infection 2010 Recommendations of the International AIDS Society–USA Panel. *Journal of American Medical Assciaition (JAMA)*. **304(3):**321-333. (doi:10.1001/jama.2010.1004)

Thomson MM, Casado G, Posada D, Marı´a Sierra M, and Najera R. 2005. Identification of a Novel HIV-1 Complex Circulating recombinant form (CRF18_cpx) of Central African Origin in Cuba. *AIDS*. **19:**1155-1163.

Tirado G, Jove G, Kumar R, Noel RJ, Reyes E, Sepulveda G, Yamamura Y, and Kumar A. 2004. Differential Virus Evolution in Blood and Genital Tract of HIV-Females: Evidence for the Involvement of Drug and Non-drug Resistance-Associated Mutations. *Virology*. **324:**577-586.

Tozzi V, Zaccarelli M, Bonfigli S, Lorenzini P, Liuzzi G, Trotta MP, Fobrici F, Gori C, Bellagamba R, Narciso P, Perno CF, and Antinori A on Behalf of the Collaborative Group for Clinical Use of HIV Genotype Resistance Test.2006. Drug-class-wide Resistance to Antiretrovirals in HIV-infected Patients Failing Therapy: Prevalence, Risk Factors and Virological Outcomes. *Antiviral Therapy*. **11:**553-560.

Triques K, Bourgeois A, Vidal N, Mpoundi-Ngole E, Mulanga-Kabeya C, Nzilambi N, Torimiro N, Saman E, Delaporte E, and Peeters M. 2000. Near Full-Length Genome Sequencing of Divergent African HIV type 1 Subtype F Viruses Leads to the Identification of a New HIV type 1 Subtype Designated K. *AIDS Research and Human Retroviruses***. 16(2):**139-151.

Trotta MA, Cozzi-Lepri A, Ammassari A, Vecchiet J, Cassola G, Caramello P, Vullo V, Soscia, F, Chiodera A, Ladisa N, Abeli C, Cauda R, Buonuomi AR, Antinori A, and Monforte AA, for the ICONA Foundation Study. 2010. Rate of CD4+ Cell Count Increase Over Periods of Viral Load Suppression: Relaltionship with the Number of Previous Virological Failure. *Clinical Infectious Diseases*. **51(4):**456-464.

Troyer RM, Collins KR, Abraha A, Fraundorf E, Moore DM, Krizan RW, Toossi Z, Colebunders RL, Jensen MA, Mullins JI,Vanham G, and Arts EJ. 2005. Changes in Human Immunodeficiency Virus Type 1 Fitness and Genetic Diversity during Disease Progression. *Journal of Virology*. **79(14):** 9006-8018.

Tscherning C, Alaeus A, Fredriksson R, Bjorndal A, Deng HK, Littman DR, Fenyo EM, and Albert J. 1998. Differences in Chemokine Coreceptor Usage between Genetic Subtypes of HIV-1. *Virology*. **241:**181–188.

Tsega E, Mengesha B, Nordenfelt E, Hansson BG, and Lindberg J. 1988. Serological Survey of Human Immunodeficiency Virus Infection in Ethiopia. *Ethiopian Medical Journal*. **26:**179-184.

Tsegaye A, De Wit TFR, Mekonnen Y, Beyene A, Aklilu M, and Messele T. Abebe A, Coutinho R, Sanders E, and Fontanet AL. 2002. Decline in Prevalence of HIV-1 Infection and Syphilis among Young Women Attending Antenatal Care Clinics in Addis Ababa, Ethiopia: Results from Sentinel Surveillance, 1995-2001. *Journal of Acquired Immune Deficiency Syndrome*. **30:**359-362.

Tully DC, and Wood C. 2010. Chronology and Evolution of the HIV-1 Subtype C Epidemic in Ethiopia. *AIDS*. **24:**1577–1582.

Turk G, Carobene M, Monczor A, Andrea Elena Rubio AE, Gómez-Carrillo M, and Salomón H. 2006. Higher Transactivation Activity Associated with LTR and Tat Elements from HIV-1 BF Intersubtype Recombinant Variants. *Retrovirology*. **3:**14 doi:10.1186/1742-4690-3-14.

Turk G, Gundlach S, Carobene M, Schindler M, Salomon H, and Benaroch P. 2009. Single Nef Proteins from HIV Type 1 Subtypes C and F Fail to Upregulate Invariant Chain Cell Surface Expression But Are Active for Other Functions. *AIDS Research and Human Retroviruses*. **25(3):**285-296.

.

Turner D, Brenner B, Routy JP, Moisi D, Rosberger Z, Roger M, and Wainberg MA. 2004. Diminished Representation of HIV-1 Variants Containing Selected drug Resistance-Conferring Mutations in Primary HIV-1 Infections. *Journal of Acquired Immune Deficiency Syndrome*. **37:**1627-1631.

UK Collaborative Group on HIV Drug Resistance. 2005. HIV Drug Resistance in Antiretroviral-Treated Individuals in the United Kingdom. *The Journal of Infectious Diseases* **192:**967-973.

UNAIDS. 2010. *Global Report: UNAIDS Report on the Global AIDS Epidemic 2010*.

UNAIDS/WHO. 2009. Report on the Global AIDS Epidemic Update, December 2009.

Urnovitz HB, Sturge JC, and Gottfried TD. 1997. Increased Sensitivity of HIV-1 Antibody Detection. *Nature Medicine*.**3:**1258.

van Harmelen J, Wood R, Lambrick M, Rybicki EP, Williamson AL, and Williamson C. 1997. An Association between HIV-1 Subtypes and Mode of Transmission in Cape Town, South Africa. *AIDS*. **11:**81-87.

Van Heuverswyn F, Li Y, Neel C, Bailes E, Keele BF, Liu W, Loul S, Butel C, Liegeois F, Bienvenue Y, Ngolle EM, Sharp PM, Shaw GM, Eric Delaporte E, Hahn BH, Peeters M. 2006. SIV Infection in Wild Gorillas. *Nature*. **444:**164.

van Oosterhout JJ, Bodasing N, Kumwenda JJ, Nyirenda C, Mallewa J, Cleary PR, de Baar MP, Burger DM, and Zijistra. 2005. Evaluation of Antiretroviral Results in Blantyre, Malawi. *Malawi Medical Journal*. **17(4):**112-116.

van Opijnen T, and Berkhout B. 2005. The Host Environment Drives HIV-1 Fitness. *Reviews in Medical Virology*. **15:**219-233.

Velazquez-Campoy A, Todd MJ, Vega S, and Freire E. 2001. Catalytic Efficiency and Vitality of HIV-1 Proteases from African Viral Subtypes. *Proceedings of National Academy of Science (PNAS)*. **98(11):**6062–6067.

Vergne L, Malonga-Mouellet G, Mistoul I, Movoungou R, Mansaray H, Peeters M, and Delaporte E. 2002. Resistance to Antiretroviral Treatment in Gabon: Need for Implementation of Guidelines on Antiretroviral Therapy use and HIV-1 Drug Resistance Monitoring in Developing Countries. *Journal of Acquired Immune Deficiency Syndrome*. **29:**165–168.

Vessierea A, Leoz M, Brodard B, Strady C, Leme´e V, Depatureaux A, Simone, and Plantier JC. 2010. First Evidence of a HIV-1 M/O Recombinant form Circulating outside Cameroon. *AIDS*. **24(7):**1079-1082.

Viljoen J, Gampini S, Danaviah S, Valéa D, Pillay S, Kania D, Méda N, Newell ML, Van de Perre P, Rouet F; for the World Health Organization/ANRS 1289 Kesho Bora Study Group. 2010. Dried Blood Spot HIV-1 RNA Quantification Using Open Real-Time Systems in South Africa and Burkina Faso. *Journal of Acquired Immune Deficiency Syndrome*. Aug 10. [Epub ahead of print]

von Truchsess I, Harris B, Schatzl HM, and Hackett Jr J. 2006. The First B/G Intersubtype Recombinant Form of Human Immunodeficiency Virus Type 1 (HIV-1) Identified in Germany Was Undetected or Underquantitated by Some Commercial Viral Load Assays. *Journal of Medical Virology*. **78:**311–317.

Wain LV, Bailes E, Bibollet-Ruche F, Decker JM, Keele BF, Heuverswyn FV, Li Y, Takehisa J, Ngole EM, Shaw GM, Peeters M, Hahn BH, and Sharp PM. 2007. Adaptation of HIV-1 to its Human Host. *Molecular Biology and Evolution*. **24(8):**1853–1860.

Wahern B and Landy A. 2002. HIV immunology better understood and vaccination attempts started. *AIDS*. **16(suppl 4):**S85-S88.

Walker PR, Ketunuti M, Choge IA, Meyers T, Gray G, Holmes EC, and Morris L. 2007. Polymorphisms in Nef Associated with Different Clinical Outcomes in HIV Type 1 Subtype C-Infected Children. *AIDS Research and Human Retroviruses*. **23(2):**204-215.

Wallis CL, Papathanasopoulos MA, Lakhi S, Karita S, Karita E, Kamali A, Kaleebu P, Sanders E, Anzala O, Bekker LG, Stevens G, Rinke de Wit TF, and Stevens W. 2010. Affordable In-House Antiretroviral Drug Resistance Assay with Good Performance in Nonsubtype B HIV-1. *Journal of Virological Methods.* **163:**505–508.

Walter BL, Armitage AE, Graham SC, de Oliveira T, Skinhoj P, Jones EY, Staurt DI, McMichael AJ, Chesebro B, and Iversen AK. 2009. Functional Characteristics of HIV-1 Subtype C Compatible with Increased Heterosexual Transmissibility. *AIDS*. **23:**1047-1057.

Watts JM, Dang KK, Gorelick RJ, Leonard CW, Bess Jr JW, Swanstrom R, Burch CL, and Weeks KM. 2009. Architecture and Secondary Structure of an Entire HIV-1 RNA Genome. *Nature*. **460(6):**711-716.

Weber J. 2001. The Pathogenesis of HIV-1 Infection. *British Medical Bulletin*. **58:**61-72.

Wei M, Cen S, Niu M, Guo F, and Kleiman L. 2005. Defective Replication in Human Immunodeficiency Virus Type 1 When Non—$tRNA_3^{Lys}$ Primers are Used for Reverse Transcription. *Journal of Virology*. **79(14):**9081-9087.

Wei X, Decker JM, Wang S, Hui H, Kappes JC, and Wu X, Salazar-Gonzalez JF, Salazar MG, Kilby JM, Saag MS, Natalia L. Komarova, Nowak MA, Hahn BH, Kwongk PD, and Shaw GM. 2003. Antibody Neutralization and Escape by HIV-1. *Nature*. **422:**307–312.

Weiss RA. 2001. Gulliver'sTravels in HIVland. *Nature*. **410:**963-967.

Wheeler WH, Zeibell RA, Zabina H, Pieniazek D, Prejean J, Bodnar UR, Mahle KC, Heneine W, Johnson JA, Hall HI, for the Variant, Atypical, and Resistant HIV Surveillance Group. 2010. Prevalence of Transmitted Drug Resistance Associated Mutations and HIV-1 subtypes in New HIV-1 Diagnoses, U.S.–2006. *AIDS*. **24:**1203-1212.

Whitcomb J, Parkin N, Chappey C, Hellmann N, and Petropoulos CJ. 2003. Broad NRTI Cross-Resistance in HIV-1 Clinical Isolates. *Journal of Infectious Diseases*. **188:**992-1000.

White EJ, McColgan B, Kassaye S, Zijenah L, and Katzenstein D. 2010. Unusual Five Amino Acid Insert within Subtype C HIV-1 Envelope Contributes to Dual-Tropism (X4R5). *AIDS*. **24:**1063-1064.

WHO. 2006. *HIV Treatment Access Reaches Over 1 Million in Sub-Saharan Africa. 2006* WHO Reports. Geneva, Switzerland.

WHO. 2007. *WHO Case Definitions of HIV for Surveillance and Revised Clinical Staging and Immunological Classification of HIV-Related Disease in Adults and Children* Available at www.who.int/hiv/pub/guidelines/HIVstaging150307.pdf.

WHO. 2009. *Guidelines for HIV Diagnosis and Monitoring of Antiretroviral Therapy*. Regional Office for South-East Asia. Revised Version 2009.

WHO. 2010. *Antiretroviral Therapy for HIV Infection in Adults and Adolescents: Recommendations for a Public Health Approach.* 2010 Revision. Geneva.

WHO/UNAIDS/UNICEF. 2010. *Towards Universal Access: Scaling up Priority HIV/AIDS Interventions in the Health Sector*. Progress Report 2010.

Wyatt R, and Sodroski J. 1998. The HIV-1 Envelope Glycoproteins: Fusogens, Antigens, and Immunogens. *Science.* **280(5371):**1884–1888.

Yamamoto JK, Sanou MP, Abott JR, and Coleman JK. 2010. Feline Immunodeficiency Virus Model for Designing HIV/AIDS Vaccines. *Current HIV Research*. **8(1):**14-25.

Yang C, Li M, Newman RD, Shi YP, Ayisi J, van Eijk AM, Otieno J, Misore AO, Steketee RW, Nahlen BL, and Lal RB. 2003. Genetic Diversity of HIV-1 in Western Kenya: Subtype-Specific Differences in Mother-to-Child Transmission. *AIDS*.**17:**1667–1674.

Yang Z, Nielsen R, Goldman N, and Pedersen AMK. 2000. Codon-Substitution Models for Heterogeneous Selection Pressure at Amino Acid Sites. *Genetics*. **155:**431–449.

Yap SH, Sheen CW, Fahey J, Zanin M, Tyssen D, Lima VD, Wynhoven B, Kuiper M, Sluis-Cremer N, Harrigan PR, and Tachedjian G. 2007. N348I in the Connection Domain of HIV-1 Reverse Transcriptase Confers Zidovudine and Nevirapine Resistace. *PLoS Medicine*. 4(12):e335. doi:10.1371/journal.pmed.0040335.

Yim HC, Li JC, Lau JS, Lau AS. 2009. HIV-1 Tat Dysregulation of Lipopolysaccharide-induced Cytokine Responses: Microbial Interactions in HIV Infection. *AIDS*. **23:**1473–1484.

Youngpairoj AS, Masciotra, Garrido C, Zahonero N, de Mendoza C, Garcia-Lerma G. 2008. HIV-1 Drug Resistance Genotyping from Dried Blood Spot Stored for 1 Year at 4^0C. *Journal of Antimicrobial Chemotherapy*. **61:**1217-1220.

Zhang H, Hoffmann F, He J, He X, Kankasa C,West JT, Mitchell CD, Ruprecht RM, Orti G, and Wood C. 2006. Characterization of HIV-1 Subtype C Envelope Glycoproteins from Perinatally Infected Children with Different Courses of Disease. *Retrovirology.* **3:**73.

Zhang H, Tully DC, Zhang T, Moriyama H, Thompson J, and Wood C. 2010. Molecular Determinants of HIV-1 Subtype C Coreceptor Transition from R5 to R5X4. *Virology*. doi:10.1016/j.virol.2010.07.047.

Zheng YH, Lovsin N, Peterlin BM. 2005. Newly Identified Host Factors Modulate HIV Replication. *Immunology Letters*. **97:**225-234.

ANNEX I

INFORMED CONSENT FORM FOR PARTCIPATING IN THE STUDY OF GENOTYPIC ANTRETROVIRAL DRUG RESISTANCE AND GENOMIC CHARACTERIZATION

Description

As it is well known, HIV/AIDS is attacking the Ethiopian people so rampantly that millions have already been victims of the pandemic in the past two decades. Although no complete cure or preventive vaccine is developed so far, it has been possible to delay the progression of the disease by treating the patients with Highly Active Antiretroviral Therapy (HAART). However, because of its high mutation rate, the virus develops drug resistance so rapidly that it requires changing the drug regimen accordingly in order to maintain the effectiveness of the treatment. The presence of drug resistant HIV variants in a patient is known only by doing genotypic drug resistance assays using venous blood drawn from the patient. Studies conducted elsewhere revealed that with introduction of Antiretroviral drugs in a community, development and transmission of drug resistant variants is inevitable. In addition, results from some investigations are indicating that genotypic drug resistance properties of HIV variants may vary depending on their subtypes.

The genotypic drug resistance properties of HIV isolates from Ethiopian patients have not been determined thus far. Moreover, the genomes of HIV sub-types prevalent in Ethiopia are not well characterized. We are therefore trying to fill this gap of knowledge by conducting a study. The purpose of this study is to produce useful information regarding the drug resistance pattern of HIV isolates from persons who are currently participating in the HAART program. In addition to enabling to identify those first-line drugs for which resistance has already been developed, the result of the study will also furnish important information about the genetic characteristics of the isolates, which may contribute towards development of vaccine against HIV subtypes prevailing in the country. Because you are a permanent client of the HAART program in this health delivery institution, we found you to be appropriate for our study. We therefore kindly invite you to take part in this study.

Conditional to accepting our invitation to participate in the study, you will be asked to give small quantity of blood (about 10 ml) from your arm, which may cause little discomfort. You will also be asked to come back after approximately 6 months and provide smaller quantity of blood from either the arm or the tip of the finger. You will receive no incentive for your participation. However, you will be refunded Birr 30 (Thirty Birr) for the transport cost you may have expended. In addition, you will be provided with CD4 count and Viral Load Test free of charge and the results of these tests will be communicated to your physician.

Strict confidentiality will be observed regarding the information we obtain from you. Every information will be stored in a file that does not bear your name; your name will be substituted with a code number. The key that relates the code number with your name will be stored locked so that no unauthorized person can get access to it. You have the right not to participate in this study, or to terminate your participation after you started. This will by no means affect the HAART program that you are getting from this health institution or from any other institution for that matter.

If you have decided to take part in the study, we kindly ask you to read the following form and put your signature at the end.

CONSENT FORM

I, whose name is written below, have decided to take part in the said study after getting the following information explained clearly.

1. That the purposes of the study are:
    a. To study the genotypic drug resistance pattern of HIV isolates from the Ethiopian patients
    b. To study subtypes and the genomic characteristics of HIV isolates from these individuals
2. That on the first visit, I will be asked to give approximately 10 ml blood sample from my arm, and that I will also be asked to come back for the same purpose after six month

3. That I have the right not to take part in the study. And that I can terminate participating in the study whenever I wish so, and that this will never affect in any way my getting HIV treatment.
4. That I shall get no incentive for my participation in the study. However, I will be refunded Birr 30 (thirty Birr) for the cost of transportation I will have expended. In addition, I shall get CD4 count and Viral Load Test done free of charge.
5. That the confidentiality of any information about me will be observed strictly.

After giving thorough thoughts on the issues explained above, I hereby ascertain that I have accepted the invitation and am willing to take part in the study.

Name of the person giving the consent: _____

Signature: _____

Date: _____

Name of the researcher/physician accepting the consent: _____

Signature: _____

Date: _____

DECLARATION

I, the undersigned, declare that this thesis is my own original work, has not been presented for a degree in any other university, and that all sources of materials used for the thesis have been duly acknowledged.

Name: Woldaregay Erku Abegaz

Date: _____

Signature: _____

Supervisor:

Name: Dr. Yohannes Mengistu

Date: _____

Signature: _____