GENOTYPING OF HUMAN IMMUNODEFICIENCY VIRUS TYPE-1 IN RELATION TO ANTIRETROVIRAL DRUG RESISTANCE IN WESTERN PROVINCE, KENYA

ADUNGO FEDINARD ONYAKASIT, BED SC
I56/10603/2006

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS OF THE DEGREE OF MASTER OF SCIENCE IN IMMUNOLOGY OF THE SCHOOL OF PURE AND APPLIED SCIENCES OF THE KENYATTA UNIVERSITY

APRIL 2011
DECLARATION

I declare that this thesis is my original work and has not been presented for degree or other awards in any other university.

Adungo Fedinard Onyakasit.
I56/10603/2006

Sign ………………………… Date …….5TH MAY 2011

Supervisor’s approval

This thesis has been submitted for examination with our approval as university supervisors.

Dr Michael Muita Gicheru
Department of Zoological Sciences
Kenyatta University
P. O. Box 43844 00200 Nairobi, Kenya.

Sign ………………………… Date ………………………

Dr Samoel Ashimosi Khamadi
Kenya Medical Research Institute
Walter Reed Project
P. O Box 1357 Kericho, Kenya.

Sign ………………………… Date ………………………
DEDICATION

I dedicate this thesis to my parents, notably to my mother Flavia Adungo for her support, prayers, love and devotion. In addition, my dedications go to all my brothers and sisters for supporting me during the study. Lastly, my dedications go to my loving wife Vivian Musungu without whose unflinching support I could not be writing this thesis. Above all praise and glory goes to the Almighty God.
ACKNOWLEDGEMENTS

My thanks go to my supervisors: Dr. Michael Muita Gicheru of Kenyatta University and Dr. Samoel Ashimosi Khamadi of Kenya Medical Research Institute currently of the Walter Reed Project for their mentorship, valuable time, guidance and encouragement during the entire study period.

Also, my gratitude goes to Dr. Nicholas Ikol Adungo of Kenya Medical Research Institute, formerly Director of Centre for Infectious and Parasitic Diseases Control Research for his mentorship and encouragement for me to pursue higher studies leading to this thesis. I would like also to thank Dr. Matilu Mwau of Kenya Medical Research Institute, Director Centre for Infectious and Parasitic Diseases Control Research for his support in facilitating the completion of this project.

More thanks go to staff members of Centre for Virus Research, KEMRI especially HIV laboratory for their valuable time, support and advice. Thanks to Dr. Elijah Songok, Dr. Michael Kiptoo, Joyceline Kinyua and Nancy Lagat for their support. My special gratitude goes to my wife Vivian Musungu for her inspiration; prayers and moral support. Lastly, thanks to my family members, brothers and sisters, friends and all the people who supported me during the entire study. Above all, thanks to Almighty God for His guidance and grace during the entire study period.
TABLE OF CONTENTS

Declaration  ii
Dedication  iii
Acknowledgements iv
Table of Contents v
List of Tables vii
List of Figures viii
List of Appendices ix
List of Abbreviations x
Abstract xii

CHAPTER ONE: INTRODUCTION

1.1 Background 1
1.2 Problem Statement 5
1.3 Study Justification 6
1.4 Research Questions 7
1.5 Null hypothesis 7
1.6 Objectives 7

CHAPTER TWO: LITERATURE REVIEW

2.1 Epidemiology and Magnitude of HIV/AIDS 8
2.2 HIV Subtypes 9
2.3 HIV Subtypes Circulating in Kenya 14
2.4 Subtype diversity in the Face of Antiretroviral Therapy 14
2.5 Mechanism of Action of Antiretroviral Drugs 16
2.6 Antiretroviral Resistance Profiles 21
2.7 Evolution and Mechanisms of Drug Resistance 23
2.8 Challenges to expanded ART in Resource-limited Settings 26
2.9 Drug Resistance Monitoring 26

CHAPTER THREE: MATERIALS AND METHODS

3.1 Study Sites and Study Population 27
3.2 Study Design 28
3.3 Inclusion and Exclusion Criteria 28
3.4 Sampling and Sample Size Determination 28
3.5 Laboratory Methods

CHAPTER FOUR: RESULTS

4.1 Confirmation of HIV pol Region PCR 34
4.2 Sequence Analysis 35

CHAPTER FIVE: DISCUSSION

5.1 Study Overview 44
5.2 HIV-1 Subtypes 45
5.3 Mutations and Their Significance 46
5.4 Transmitted Drug Resistance Mutations 48
5.5 Susceptibility Profiles in Relations to Subtype Diversity 49

CHAPTER SIX: CONCLUSIONS, RECOMMENDATIONS AND LIMITATIONS

6.1 Conclusions 50
6.2 Recommendations 50
6.3 Study limitations 51

REFERENCES 52
APPENDICES 65
LIST OF TABLES

Table 2.1: HIV-1 subtypes with their respective estimated worldwide prevalence 13
Table 2.2: Antiretroviral agents currently in clinical use against HIV infection 18
Table 4.1: BLAST results of generated sequences 35
Table 4.2: Categories of protease-associated mutations detected 42
Table 4.3: Protease associated mutations and drug susceptibility patterns 43
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Structure of HIV virion particle</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>HIV and its genes</td>
<td>2</td>
</tr>
<tr>
<td>2.1</td>
<td>Phylogenetic tree illustrating HIV-1 group M subtypes</td>
<td>10</td>
</tr>
<tr>
<td>2.2</td>
<td>Geographic distribution of major HIV-1 subtypes</td>
<td>12</td>
</tr>
<tr>
<td>3.1</td>
<td>Map of Kenya showing location of sampling sites</td>
<td>27</td>
</tr>
<tr>
<td>4.1</td>
<td>Photomicrograph of Agarose gel electrophoresis</td>
<td>34</td>
</tr>
<tr>
<td>4.2</td>
<td>Bar graph of isolated HIV-1 subtypes and countries of close origin</td>
<td>38</td>
</tr>
<tr>
<td>4.3</td>
<td>Phylogenetic tree analysis of generated sequences</td>
<td>40</td>
</tr>
<tr>
<td>4.4</td>
<td>Pie chart of HIV-1 subtypes isolated</td>
<td>41</td>
</tr>
</tbody>
</table>
LIST OF APPENDICES

APPENDIX I: Information for Participants 65
APPENDIX II: Informed Consent 66
APPENDIX III: Data Collection Form 69
APPENDIX IV: Primer Sequences (Protease Region) 71
APPENDIX V: Genotypic Resistance Interpretation Algorithm 72
APPENDIX VI: Copy of Ethical Approval permission 73
APPENDIX VII: Copy of Results (Generated Sequences) 74
LIST OF ACRONYMS AND ABBREVIATIONS

3TC  Lamivudine
AIDS  Acquired immunodeficiency syndrome
ART  Antiretroviral therapy
ATV/r  Atazanavir
AZT  Zidovudine
BLAST  Basic local alignment search tool
bp  base pair
CIPDCR  Centre for infectious and parasitic diseases control research
CVR  Centre for virus research
CRFs  Circulating recombinant forms
ddC  Zalcitabine
ddI  Didanosine
DNA  Deoxyribonucleic acid
DRV/r  Darunavir
EDTA  Ethylene diamine tetra-acetic acid
EFV  Efavirenz
FPV/r  Fosamprenavir
FTC  Emtricitabine
HAART  Highly active antiretroviral therapy
HIV  Human immunodeficiency virus
IDV/r  Indinavir
KDHS  Kenya demographic and health survey
KEMRI  Kenya Medical Research Institute
LPV/r  Lopinavir
MgCl₂  Magnesium chloride
NFV  Nelfinavir
NNRTI  Non-nucleoside reverse transcriptase inhibitor
NRTI  Nucleoside reverse transcriptase inhibitor
NVF  Nevirapine
PBMC  Peripheral blood mononuclear cell
PCR  Polymerase chain reaction
PMTCT  Prevention of mother to child transmission
PLWHA  People living with HIV/AIDS
POL  Polymerase
PR  Protease
rpm  Revolutions per minute
RT  Reverse transcriptase
RT-PCR  Reverse transcriptase polymerase chain reaction
SQV/r  Saquinavir
SSA  Sub-Saharan Africa
STIs  Sexually transmitted infections
TAMs  Thymidine analogue mutations
TPV/r  Tipranavir
U/L  Units per litre
UNAIDS  United Nations programme on HIV/AIDS
VCT  Voluntary counselling and testing
WHO  World health organization
ABSTRACT

Human Immunodeficiency Virus/Acquired Immunodeficiency Syndrome (HIV/AIDS) is the main pandemic in the world today. Sub-Saharan Africa bears the brunt of the catastrophe. Although it harbours less than 10% of the world’s population, it comprises 60% of all people living with HIV/AIDS globally. Kenya has a National adult HIV/AIDS prevalence of 7.4%. In recent times, expansion of access to antiretroviral therapy in resource-limited settings has gained prominence. However, it is estimated that in Africa, only one million people out of the over four million are receiving ART. Highly active antiretroviral therapy has markedly improved the prognosis of HIV-infected patients by controlling HIV replication. However, it has been observed that HAART fails to control HIV replication in an increasing number of patients as a result of a complex array of causes. One of these is the emergence of drug resistance leading to antiretroviral therapy failure. More seriously, HIV drug resistance can be transmitted and lead to treatment failure. The main objective of this study was to determine the circulating HIV-1 subtypes in Western Kenya, and the prevalence of mutations in the protease gene that confer antiretroviral resistance in the various subtypes. A total of 75 patients were sampled randomly from a cohort of 1000 clients receiving ARVs in Busia District Hospital. A single blood draw (5ml) was collected, packaged and shipped to KEMRI HIV laboratories in Nairobi for genotyping and analysis of drug associated mutations. Polymerase chain reaction, sequencing and phylogenetic analysis of HIV-1 partial protease showed that majority of strains were of subtype A1 (39/75) followed by D (21/75), C (5/75) and G (5/75) with possible recombinants of AC (1/75) and AD (1/75). Other prevalent subtypes included, A2 (2/75) and B (1/75). The generated sequences were further analysed using the sequence analysis Stanford HIV drug resistance database for drug-associated resistance mutations. Resistance associated mutations were detected in 12 (16.0%) isolates as follows: Subtypes A1 (n=10), subtype G (n=1) and subtype C (n=1). The analysis further indicated that major protease mutations such as D30N, V32I, M46I, and I47L were circulating in this region among patients unexposed to this class of ARVs. This is contrary to the earlier observations in which these mutations have been documented mainly in patients on protease-based regimens, an aspect attributed to transmitted-resistance mutations. A total of 4 subtypes A1 isolates were found to possess V32L and T74S mutations that result to reduced susceptibility to Nelfinavir. One subtype C isolate had D30N, G48R and G73S mutations, which are known to cause high-level resistance to Nelfinavir. Other minor protease mutations such as M36I and K20I were also detected. Such mutations are common to some patients not previously exposed to ARVs, especially in non-B subtypes. There is need for regular surveillance of drug resistance and monitoring of mutations conferring resistance to ARVs in both treated and untreated patient population in order to achieve successful treatment outcomes.
CHAPTER ONE: INTRODUCTION

1.1 Background

HIV-1 belongs to the family of Retroviridae and subfamily Lentivirinae. Infections with lentiviruses typically show a chronic course of disease, a long period of clinical latency, persistent viral replication and involvement of the central nervous system (Levy, 1993; Mehandru et al., 2004) Using electron microscopy, HIV-1 and HIV-2 resemble each other. However, they differ with regard to the molecular weight of their proteins as well as in their accessory genes (HIV sequence compendium, 2008). HIV-2 is genetically more closely related to the Simian Immunodeficiency Virus (SIV) found in sooty mangabeys (SIVsm) than it is to HIV-1. Both HIV-1 and HIV-2 replicate in CD4+ cells and are regarded as pathogenic in infected persons, although the actual immune deficiency is less severe in HIV-2-infected individuals (Reeves and Doms, 2002)

HIV-1 viral particles have a diameter of 100 nm and are surrounded by a lipoprotein membrane. Each viral particle contains 72 glycoprotein complexes, which are integrated into the lipid membrane, and are each composed of trimers of an external glycoprotein gp120 and a trans-membrane spanning protein gp41 (Figure 1.1; Kwong et al., 1998). The bonding between gp120 and gp41 is not strong and therefore gp120 may be shed spontaneously within the local body environment. Glycoprotein gp120 may also be detected in the serum (Kwong et al., 1998) as well as within the lymphatic tissue of HIV-infected patients (Sunila, 1997). During the process of budding, the virus may also incorporate different host proteins from the membrane of the host cell into its lipoprotein layer, such as HLA Class I and II proteins, or adhesion proteins such as ICAM-1 that may facilitate adhesion to other target cells. The matrix protein p17 is anchored to the inside of the viral lipoprotein membrane. The p24 core antigen contains two copies of HIV-1 RNA, which is part of a protein nucleic acid complex,
which is composed of the nucleoprotein $p7$ and the reverse transcriptase $p66$ (RT). (HIV sequence compendium) The viral particle contains all the enzymes that are necessary for replication (Figure 1.2): a reverse transcriptase (RT), an integrase $p32$ and a protease $p11$. The genetic material of HIV, an RNA molecule 9 kilobases in length, contains 9 different genes encoding 15 proteins. Considerable insights have been gained into the function of these different gene products (Doms and Trono, 2000). The HIV virus is spiked. It has a pair of RNA strands within the matrix of the virus. The soft surface of the virus has the $gp120$ and $gp41$ proteins (Chrystie et al., 1998).

![Figure 1.1: Structure of HIV virion particle](adopted from Gallo et al., 1988)

![Figure 1.2: HIV and its genes](adopted from Gallo et al., 1988)
Two major types of HIV are currently recognised, HIV type 1 (HIV-1) and HIV type 2 (HIV-2). HIV-1 is widely distributed worldwide. It can be further divided into four genetic groups: group M (major or main), group O (outlier), group N (new or non-M, non-O) and more recently group P described in West Africa (Plantier et al., 2009; Robertson et al., 2000). While groups O and N are restricted to countries of Central Africa, notably Cameroon; HIV-1 group M is worldwide, being responsible for the AIDS pandemic, which accounts for over 90% of HIV infections worldwide (Hemelaar et al., 2006). HIV-2 is restricted to countries in West Africa, where it also represents a minority of viral infections (3% of total HIV infections). Fortunately, it is decreasing in prevalence over time (Eholie and Anglaret, 2006).

Treatment of Human Immunodeficiency Virus type 1 (HIV-1) infected individuals with antiretroviral drugs has been highly effective in prolonging and improving quality of life. Currently, the standard care involves combinations of protease (PR) and reverse transcriptase (RT) inhibitors (Carpenter et al., 1998). These combination regimens often reduce the amount of virus circulating in the plasma of patients to below the limits of detection by currently available assays. This impedes the development of resistance to antiretroviral drugs by reducing the number of replicative events in which mutants can be generated (Cane, 2009). The development of resistant viral strains is one of the main reasons for failure of antiretroviral therapy (Kozal, 2009). This treatment failure, typically defined as a significant rise in the level of circulating virus, is often associated with the emergence of virus strains resistant to antiretroviral drugs (Hirsch et al., 1998). ART is expected to lower HIV loads in plasma and genital secretions because plasma HIV load is correlated with the risk of sexual transmission (Quinn et al., 2000). Antiretroviral drug resistance is an important cause of treatment failure in persons infected with HIV-1 and has been associated with increased mortality (Weinstock et al., 2004).
Development of drug-resistant strains has been linked to a number of factors: the dynamics of HIV replication, patient and physician failure to adhere to therapy conditions on the one hand; and drug-related factors on the other hand, which leads to emergence of HIV resistant strains in most patients. If mutations occur in one of the HIV proteins that are a target of antiretroviral drugs, the result may be decreased susceptibility or development of resistance, and lack of inhibition of viral replication by the drug. Transmission of drug resistance from treatment-experienced patients to newly infected persons has been observed in countries with expanded access to ART (Puig et al., 2000). It is estimated that 8-30% of people who contract HIV infection in Europe or the United States of America acquire viruses with drug resistance conferring mutations (Wensing and Boucher, 2003; Yerly et al., 2001). It is also evident that transmission of drug-resistant strains increases despite all prevention efforts, thus raising major public health concerns (Wensing and Boucher, 2003). Although antiretroviral therapy can have a significant impact on the HIV epidemic by reducing the probability of transmission per contact (Blower et al., 2000), transmission of drug-resistant HIV can reduce the positive impact of such therapy.

Since strains harbouring resistance-associated mutations to a single or multiple antiretroviral agents can be transmitted both horizontally and vertically, it was imperative to undertake this study in order to determine the prevalence of resistant strains in the community, hand in hand with the widespread access to antiretroviral drugs. The study also involved genotypic analysis of blood samples obtained from eligible patients on HAART through the Medecins Sans Frontieres (MSF) supported programs in western Kenya. Blood samples collected were amplified by nested polymerase chain reaction and directly sequenced to determine circulating HIV-1 subtypes and drug-resistance associated mutations in the protease (PR) gene of the virus.
Transmitted resistance may unfavourably affect clinical management of patients who have never been exposed to therapy. This poses one of the major obstacles to the long-term efficacy of ART (Durant et al., 1999). Despite the increasing access to treatment and care, most developing countries adopting large scale HAART suffer from health service associated problems, which are likely to impact negatively on implementation and validation of this programme. Our health systems are under-staffed and over-burdened. As a result of most governments being preoccupied with the problems mentioned above, monitoring development and spread of ARV drug resistance is receiving very little or no attention.

1.2 Problem Statement

HIV-1 displays extensive genetic diversity globally and different genetic subtypes with mosaic recombinants have continued to emerge in different geographic locations. There is need for continuous monitoring of emerging HIV subtypes/recombinants in country border points such as Busia, since this region is the Kenyan gateway to central Africa where all reported HIV subtypes have been documented. Antiretroviral therapy has led to improved quality of life and enhanced survival of persons infected with HIV. However, under therapy individuals develop spontaneous mutations in the target genes of HIV that could adversely affect the efficacy of therapy. In Kenya, in spite of the rapidly expanding access to treatment, there have been very limited attempts to describe the presence and extent of drug resistance mutations in ARV-naïve or experienced patients.
1.3 Justification

In the context of envisaged increase of HIV-1 subtype diversity along country border points as a result of migration, it would be necessary to monitor the transmission routes from country to country, to document circulating subtypes and advise on best treatment options based on prevalent subtypes. This will serve to advice on effective testing methodologies and regimens should those infected seek treatment.

Treatment failure is already becoming a major problem in ARV therapy in Kenya. Quite often it is associated with failure to adhere to therapy, low potency of some antiretroviral regimens, and sub-optimal drug pharmacokinetics. On several occasions, HIV drug resistance has been described in ARV-naïve patients. Presumably these clients have contracted resistance strains. ARV resistance surveillance could prove to be a vital component of rational ARV drug use. For these reasons, this study was conducted to determine circulating HIV-1 subtypes and to describe the presence and extent of drug resistance mutations in western Kenya town of Busia, which is a crossing border point between Kenya and Uganda. Such information is useful in updating treatment guidelines and to determine the best drug combinations for patients developing antiretroviral drug resistance.
1.4 Research questions

a) What are the HIV-1 subtypes circulating along the Kenya-Uganda border?

b) What is the prevalence rate of mutations in protease gene of HIV-1 that can confer resistance to specific antiretroviral drugs in western Kenya?

1.5 Null Hypothesis

Mutations on the protease gene of circulating HIV-1 subtypes are non-existent in western Kenya.

1.6 Objectives

1.6.1 General Objective

To determine the circulating HIV-1 subtypes and to describe the presence and extent of drug resistance mutations in the protease gene of HIV-1 that can confer resistance to specific antiretroviral drugs in western Kenya.

1.6.2 Specific Objectives

a) To determine HIV-1 subtypes circulating along the Kenya/Uganda border.

b) To determine the prevalence rates of mutations in the protease gene of HIV-1 that can confer resistance to specific antiretroviral drugs in western Kenya.
CHAPTER TWO: LITERATURE REVIEW

2.1 Epidemiology and Magnitude of HIV/AIDS

Human Immunodeficiency Virus (HIV) infection is today one of the most devastating infections globally (UNAIDS, 2005 and WHO, 2005). Ever since the discovery of the virus some two decades ago, there has been a dramatic increase in infections and deaths due to this virus. For example, in the recent times: 2008, there were 2.0 million deaths due to AIDS while approximately 2.7 million people were newly infected with HIV (UNAIDS, 2009 and WHO, 2009). At the same time, some 22.4 million (71%) of the infected people were living in sub-Saharan Africa, where the pandemic is particularly more devastating. The average HIV prevalence among the adult population across sub-Saharan Africa is 5.2% (UNAIDS 2009; UNICEF, 2009; WHO, 2009). It is estimated that Kenya has approximately 1.4 million individuals of ages 15-64 living with HIV/AIDS giving adult prevalence of 7.4% according to the results of 2007 Kenya AIDS Indicator Survey (KAIS, 2007). Women continue to be disproportionately infected with HIV (8.7%) compared to men (5.6%). Young women between ages 15 and 34 are more likely to have HIV compared to young men in the same age group. The survey further indicated that there is a high HIV infection rate among older adults of ages 50-64. KAIS showed that only 12% of the 1.4 million HIV infected adults who need cotrimoxazole (Septrin) were accessing medication in 2007. Although 450,000 HIV patients need ARVs, only 190,000 (42.2%) infected adults were accessing antiretroviral therapy (ART) by the end of June 2008.

HIV-1 epidemic in Kenya like the rest of the world shows regional heterogeneity. While Nyanza province has the highest overall prevalence of 15.3% followed by Nairobi (9%), Coast (7.9%), Western province (6.7%) and Rift Valley (7%). Other provinces have levels
between 4% and 6% overall; except for North Eastern that indicates a very low prevalence of only 1% (KAIS, 2007).

2.2 HIV Subtypes

HIV is a member of the Retroviridae family, which has a single stranded RNA genome. As illustrated in Figure 2.1, nine pure subtypes of HIV-1 group M are currently known (A-D, F-H, J and K). Some of these subtypes are further subdivided into sub-subtypes e.g. subtypes F (F1 and F2); A (A1, A2 and A3). Intrasubtype divergence of up to 20% also occurs. Furthermore, intersubtype divergence of about 25% to 35% occurs, for the env amino acid sequences (Achkar et al., 2004; Robertson et al., 2000; Triques et al., 2000). This intermixture of HIV-1 variants that circulate together within a geographical region provides an opportunity for recombination of virus strains within dually or multiply infected individuals (Burke, 1997). Some of these recombinant forms may further achieve epidemic relevance, giving rise to unknown circulating recombinant forms (CRFs). To date, over 40 CRFs are recognized in diverse parts of the world (Peeters and Sharp, 2000). It is currently believed that HIV-1 M subtypes and CRFs are the result of founder effects in different geographic localities, followed by localised evolution. As a consequence, such HIV-1 forms are heterogeneously spread out worldwide (Hemelaar et al., 2006).
Figure 2.1: The phylogenetic tree illustrating the different HIV-1 groups, subtypes and sub-subtypes. Source: (Robertson et al., 2000).

The phylogenetic tree shows the different HIV-1 groups, subtypes and sub-subtypes. Each of the internal branches defines a subtype or sub-subtype.

Evolutionary relationships among non-recombinant HIV-1 strains from the HIV-1/SIVcpz lineage based on neighbour joining phylogenetic analysis of near full-length genome sequences. Other group M subtypes of HIV-1 are subdivided into subtype A, B, C, D, F, G, H, J and K. Only subtype B, however, is of more significance in that following it being widely studied, it forms the basis of antiretroviral drugs. Additionally, it also forms the basis of
information on drug resistance studies. A number of studies point to the fact that this subtype B is predominant among homosexuals especially in Europe, *vis a vis* the rest of the subtypes which are found predominantly in heterosexuals (Ortiz *et al.*, 2000).

### 2.2.1 Geographical Distribution of HIV-1 Subtypes

HIV-1 subtype diversity is the highest in the world. Figure 2.2 illustrates major HIV-1 subtypes and circulating recombinant forms in Africa. It shows that Africa has all HIV-1 subtypes although A and C predominate (Nasioulas *et al.*, 1999). For example, East Africa has majorly subtypes A and D, which have dominated Uganda from mid 1980s (Hu *et al.*, 2000). As mentioned elsewhere, the majority of the circulating HIV-1 subtypes in Kenya belong to group M.

While HIV-1 subtype B is predominant in Europe, it is also found in Indonesia, the Philippines, and Taiwan (Ortiz *et al.*, 2000). Subtype F occurs in Romania, and subtype C is found in a small proportion in Brazil (McCutchan *et al.*, 2000). Subtype G has been documented in many West and Central African countries. Subtype H, however, is found only in Central Africa (Peeters and Delaporte, 1999).
Figure 2.2: Distribution of Major HIV-1 Subtypes and Circulating Recombinant Forms in Africa. Source: Essex and Mboup, 2002; UNAIDS and WHO, 2005.

2.2.2 Circulating Recombinant Forms

In certain populations and regions where multiple HIV-1 subtypes co-circulate, many combinations of intersubtype recombinant viruses have been documented: A/C, A/D, B/F, A/G/I/J (Montavon et al., 1999). Even an intergroup M/O recombinant virus has been recently isolated from a Cameroonian patient (Peeters and Delaporte, 1999). In Southern Tanzania town of Mbeya, the predominant subtype is C (Hoelscher et al., 2001). This probably reflects introduction from Southern African countries where subtype C is predominant (De Baar et al., 2000). HIV-1 subtypes A and D have been isolated from other parts of Tanzania (Blackard et al., 1999).
These recombinant viruses however, may have some advantages over their parental strains, because of eventual modifications on tropism and replication efficiency (fitness). To this end, recombination may play a significant role in global HIV evolution, creating novel viral genotypes within populations (Peeters and Sharp, 2000); thus impacting on the efficacy of ARVs. The differential distribution of HIV-1 subtypes and CRFs appears to be concomitant with their worldwide prevalence estimates (Table 2.1). For instance, while subtype B circulates in developed countries with the lowest HIV prevalence rates, it accounts for a little less than 10% of the worldwide infections. Conversely, subtype C accounts for nearly half of worldwide infections (Hemelaar et al., 2006) as it prevails in countries with the highest HIV prevalence, such as South Africa and Botswana, and in highly populated countries such as India.

Table 2.1: Most prevalent HIV-1 subtypes, with their respective estimated worldwide prevalence. Source: Hemelaar et al. (2006).

<table>
<thead>
<tr>
<th>HIV-1 subtype/CRF</th>
<th>Global prevalence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>50</td>
</tr>
<tr>
<td>A</td>
<td>12</td>
</tr>
<tr>
<td>B</td>
<td>10</td>
</tr>
<tr>
<td>F, G, H, J and K</td>
<td>7</td>
</tr>
<tr>
<td>CRF01_AE</td>
<td>5</td>
</tr>
<tr>
<td>CRF02_AG</td>
<td>5</td>
</tr>
<tr>
<td>D</td>
<td>3</td>
</tr>
<tr>
<td>Others</td>
<td>8</td>
</tr>
</tbody>
</table>

CRF = Circulating recombinant form.
2.3 HIV Subtypes Circulating in Kenya

In Kenya, the majority of circulating strains belong to group M. Four circulating recombinant forms have so far been documented among STI patients in Nairobi (Lihana et al., 2006). A study based on the analysis of the env C2-V3 region (Neilson et al., 1999), revealed that subtype A predominates (71-87%), with significant components of subtype D (7-29%) and subtype C (7-17%). Subtype A therefore, predominates in both non-recombinants (55%) and recombinants based on near full-length sequences that have been generated (Dowling et al., 2002). Subtype C and D occur as non-recombinant (2% each) but to a much lesser extent than subtype A. Also, a full-length subtype G has been identified in Kenya (Carr et al., 2005). This has recombinants between A1, A2, and D; A2 and D; A1 and D; A1 and G; A1 and C; A1, C and D (Dowling et al., 2002).

On the other hand, previous studies on molecular epidemiology of HIV-1 in Kenya have shown that HIV-1 subtypes circulating in the Northern region of Kenya mainly comprised subtype A (50%), subtype C (39%) and subtype D (11%) on the basis of partial env sequences that is subtypes A and C may be dominant. Conversely, neighbouring Ethiopia is dominated by HIV-1 subtype C, which interestingly is also the dominant subtype in the town of Moyale, on the border of Kenya and Ethiopia. This may be indicative of cross-border movements influencing circulation of subtypes in northern Kenya (Khamadi et al., 2005).

2.4 Subtype Diversity in Relation to Antiretroviral Therapy

In many cases, new epidemics are due to CRFs that are composed of viruses of different subtypes. While combination therapy has helped to inhibit progression and rates of HIV-1 subtype B infections in many western countries, worldwide epidemics with group M (non-B, A through J) and O subtypes are expanding (Louwagie et al., 1995; Myers, 1994). Although
non-subtype B viruses have also been investigated for the development of resistance, knowledge on how the resistance pathways progress in the various subtypes is not well understood (Handema et al., 2003). So far it is known that standard ARV regimens used in the treatment of HIV-1 subtype B are effective against all group M subtypes. However, HIV-1 group O and HIV-2 show natural resistance to non-nucleoside reverse transcriptase inhibitors (Quinones-Mateu et al., 1998; Descamps et al., 1997; Descamps et al., 1995; Tantillo et al., 1994). It is observed that NNRTIs do not completely inhibit progression of HIV-2, which fortunately is found in limited areas.

The epicentres of HIV-1 infection are currently in Africa, which comprises of 71% of new infections with the following subtypes: A, A/E, A/G, C, D, G, O. South-East Asia with 17% of new infections has the subtypes A/E and C (Quinones-Mateu et al., 1998; Simon et al., 1998). Particularly troublesome are escalating rates of subtype C infections amounting to 50% of new HIV-1 infections. Moreover, this pandemic has shifted towards densely populated regions of China, India and the countries of the former USSR. Most available data on antiretroviral resistance are derived from patients with subtype B viruses albeit representing only 12% of the worldwide HIV- infected population. With globalization and immigration, over 40% of new infections in Europe are currently thought to be of non-B subtypes (Quinn et al., 2001).

The prevalence of mutations already present in treatment-naive patients differs along demographic regions. Although the transmission of drug-resistant strains of HIV has been well documented (Weinstock et al., 2004), the prevalence and characteristics of persons with or without mutations associated with drug resistance are less clear. A number of studies have examined the prevalence of mutations associated with resistance. For example, in small
samples of recently or acutely infected persons, mostly homosexuals drawn from Caucasian populations, Grant et al.(2002) showed that genotypic resistance to 2 or more classes of drugs increased from 2.5% to 13.2%. Botswana was the first African country to rollout antiretroviral therapy (ART) on a large-scale, prescribing a single standard combination regimen as the most plausible therapy. Unfortunately, results revealed that individuals under such treatment regimens developed drug-resistant strains at an annual rate of 25%, after which they became infectious again and spread resistant strains. Under therapy HIV infected individuals are assumed to be less infectious because of low viral loads (Vernazza et al., 2000). However, it must be noted that after resistance has developed, there are no benefits from treatment; even though HAART increases the life expectancy of an individual with drug-susceptible HIV.

In a European multi-centric study covering more than 1600 newly diagnosed HIV-infected patients, the prevalence of primary resistance mutations over a period of 6 years (1996-2002) was about 10%. This resistance was primarily observed in recent infections with subtype B (Wensing et al., 2005). In another study covering 40 cities in the United States of America, 14% of the 371 isolates of treatment-naive patients had at least one primary mutation (Ross et al., 2007). Other studies have shown that the prevalence of resistance among patients with acute or recent infection can even be higher. For example in San Francisco and Spain, where it was 26% and 19% respectively (Blower et al., 2003; de Mendoza et al., 2002).

### 2.5 Mechanism of Action of Antiretroviral Drugs

#### 2.5.1 Nucleoside Reverse Transcriptase Inhibitors (NRTIs)

The nucleoside analogues (Table 2.2) are comprised of a base (thymidine in the case of AZT) that attaches to a ribose sugar in which the normal 3’(OH) gets replaced by an azido group
(Martin, 2006). The presence of the 3’ OH is required for elongation of the growing DNA chain. Replacement of the OH at the 3’ position prevents disulphide bonds from being formed with this nucleoside. Thus incorporation of AZT into the growing DNA chain in place of the normal nucleoside leads to a chain termination resulting into “a dead end virus”. These anti-HIV nucleoside inhibitors are also referred to as nucleoside reverse transcriptase inhibitors (NRTIs) because they are more potent inhibitors of HIV RT than human DNA polymerases (Martin, 2006). The mode of action is that both nucleoside and non-nucleotide RT inhibitors must enter the cell and become phosphorylated in order to act as synthetic substrates for RT. However, both classes of agents can prevent infection of susceptible cells but will have no effect on cells that already harbour HIV. Likewise, this class of agents target the active site on RT that is involved in DNA polymerization (Martin, 2006). Resistance to NRTIs occurs through two mechanisms. The first mechanism involves mutation of the residues that results in reduced incorporation of the NRTI into the growing DNA chain leading to conformational changes in the enzyme. This impairs binding of the drug to the active site (Martin, 2006).
Table 2.2: Antiretroviral agents currently in clinical use to counter HIV infection

<table>
<thead>
<tr>
<th>Generic name</th>
<th>Abbreviation</th>
<th>Brand name</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nucleoside/nucleotide reverse transcriptase inhibitors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zidovudine</td>
<td>AZT, ZDV</td>
<td>Retrovir®</td>
</tr>
<tr>
<td>Didanosine</td>
<td>DDI</td>
<td>Videx®</td>
</tr>
<tr>
<td>Stavudine</td>
<td>D4T</td>
<td>Zerit®</td>
</tr>
<tr>
<td>Lamivudine</td>
<td>3TC</td>
<td>Epivir®</td>
</tr>
<tr>
<td>Emtricitabine</td>
<td>FTC</td>
<td>Emtriva®</td>
</tr>
<tr>
<td>Abacavir</td>
<td>ABC</td>
<td>Ziagen®</td>
</tr>
<tr>
<td>Tenofovir</td>
<td>TDF</td>
<td>Viread®</td>
</tr>
<tr>
<td><strong>Non-nucleoside reverse transcriptase inhibitors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nevirapine</td>
<td>NVP</td>
<td>Viramune®</td>
</tr>
<tr>
<td>Delavirdine</td>
<td>DLV</td>
<td>Rescriptor®</td>
</tr>
<tr>
<td>Efavirenz</td>
<td>EFV</td>
<td>Sustiva®</td>
</tr>
<tr>
<td>Etravirine</td>
<td>TMC-125</td>
<td>Intelence®</td>
</tr>
<tr>
<td><strong>Protease inhibitors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indinavir</td>
<td>IDV/r</td>
<td>Crixivan®</td>
</tr>
<tr>
<td>Ritonavir</td>
<td>RTV</td>
<td>Norvir®</td>
</tr>
<tr>
<td>Nelfinavir</td>
<td>NFV</td>
<td>Viracept®</td>
</tr>
<tr>
<td>Saquinavir</td>
<td>SQV/r</td>
<td>Invirase®</td>
</tr>
<tr>
<td>Amprenavir</td>
<td>APV</td>
<td>Agenerase®</td>
</tr>
<tr>
<td>Fosamprenavir</td>
<td>FPV/r</td>
<td>Lexiva®</td>
</tr>
<tr>
<td>Lopinavir</td>
<td>LPV/r</td>
<td>Kaletra®</td>
</tr>
<tr>
<td>Atazanavir</td>
<td>ATV/r</td>
<td>Reyataz®</td>
</tr>
<tr>
<td>Tipranavir</td>
<td>TPV/r</td>
<td>Aptivus®</td>
</tr>
<tr>
<td>Darunavir</td>
<td>DRV/r</td>
<td>Prezista®</td>
</tr>
<tr>
<td><strong>Fusion inhibitor</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enfuvirtide</td>
<td>T-20</td>
<td>Fuzeon®</td>
</tr>
<tr>
<td><strong>Integrase inhibitor</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raltegravir</td>
<td>MK-0518</td>
<td>Isentress®</td>
</tr>
<tr>
<td><strong>Entry inhibitor</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maraviroc</td>
<td>MRV</td>
<td>Selzentry®</td>
</tr>
</tbody>
</table>
The second mechanism of NRTI resistance is associated with enhanced removal of drug from its site of attachment at the end of the DNA chain. These RT mutations allow ATP or pyrophosphate to bind at the active site adjacent to the bound nucleoside analogue. Both of these products are in high concentration within the cell. The high energy ATP or pyrophosphate can then attack the bond that binds the drug to DNA, thereby liberating the drug and terminating its effect (Martin, 2006).

2.5.2 Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs)

The second class of ARVs are the Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs). These drugs are structurally different from the nucleoside RT inhibitors (Table 2.2). They include Nevirapine, Delavirdine and Efavirenz. The NNRTIs bind near the catalytic site of reverse transcriptase and alter the enzymes ability to change conformation (Martin, 2006). This increased enzyme rigidity prevents its normal polymerization function. Resistance to this class of agents occurs mainly through mutation of hydrophobic RT residues within the binding pocket for the NNRTIs (Martin, 2006). The adverse side effects of the NNRTIs lead to non-adherence to treatment, thereby enhancing development of drug resistance. Because of this, NNRTIs are never to be used for monotherapy of HIV infection. Since all NNRTIs bind essentially at the same region of RT, mutations in this area will affect binding of all of the agents in this class to some extent. This may in part explain the high rates of HIV cross-resistance within this class of agents (Brennar et al., 2003).

2.5.3 Protease Inhibitors

The third class of ARVs is the Protease inhibitors (Table 2.2). Protease inhibitors commonly used in HIV treatment such as Saquinavir, Indinavir, Ritonavir, Nelfinavir are designed to fit and bind at the catalytic site of the HIV protease enzyme with high affinity and thereby block
the viral enzyme from cutting the viral protein molecules into the correct sizes. This process known as maturation involves cleavage of viral protein precursors by HIV protease enzymes into small pieces that then infect new cells. Once the protease inhibitors are encoded by HIV, they offer a unique and attractive target for preventing HIV maturation (Martin, 2006). Protease enzymes within the HIV are symmetrical dimers with a central core that binds the peptides, which are to be modified by the drugs (Martin, 2006). This results to the virus being unable to make copies that can infect new cells.

2.5.4 Fusion Inhibitors

Another class of ARVs is the fusion inhibitors (Table 2.2), which act by blocking the fusion and entry of HIV into the host cells. Fusion of HIV with the host cell membrane is an essential step in viral entry into the cell (Martin, 2006). In the process, HIV attaches specifically to CD4 molecules on the host cell membrane through glycoprotein gp120 on the HIV peplomer. Once attachment to the host cell occurs, gp41, a stalk of the HIV peplomer, embeds itself in the host cell membrane (Martin, 2006).

The gp41 peplomer comprises of two adjoining subunits, HR1 and HR2. The embedding of the gp41 involves the HR1 subunit of gp41 “sliding” over the HR2 subunit to draw the HIV and host cell membranes close together. The gp41 fusion peptide then undergoes a further conformational change that brings the HIV and host cell membranes in contact with one another resulting in “Fusion pores”. These facilitate entry of the HIV nucleocapsid into the host cell. The only drug, enfuvirtide, which has just been introduced to clinical practice, is a synthetic peptide that binds directly to the HIV gp41 and prevents it from undergoing the conformational change that would lead to fusion of HIV and host cell membrane. However, varying susceptibility of different HIV strains to enfuvirtide has already been documented.
(Greenberg and Cammack, 2004). While clinical resistance to enfuvirtide has not yet been observed, amino acid mutations between residues 36 and 45, which are part of the binding site for enfuvirtide, have been identified to confer some acquired resistance to the drug. However, since the region in which the mutations occur is required for viral function, enfuvirtide-resistant mutants still replicate poorly and revert back to full-drug susceptibility once the agent is stopped (Martin 2006).

2.6 Antiretroviral Resistance Profiles

HIV drug resistance will remain one of the greatest challenges to management of HIV positive patients. Transmitted resistance has the potential to more rapidly reverse the effectiveness of first-line antiretroviral therapy at the population level. It has been observed that widespread use of antiretroviral drugs to treat HIV-1 infection might result in an increasing transmission of drug-resistant virus, thus compromising therapeutic options in newly infected individuals (Little et al., 2002). It has also been observed that several other factors, such as the dynamics of HIV replication, patients themselves, physician practise and drug-related factors could lead to the emergence of HIV resistant strains. The efficacy, therefore, of each class or individual ARV is threatened by specific mutations and resistance mechanisms.

Virologic failure is defined as insufficient decrease, increase or rebound of viral load/CD4 count and or a HIV associated disease occurring during HAART. It is often associated with resistance to ARVs. HIV replicates at the rate of approximately $10^8$ to $10^9$ viral particles per day, probably giving daily rise of about $3 \times 10^{-3}$ spontaneous changes (mutations) in its genetic sequence. HAART has dramatically reduced the morbidity and mortality associated with HIV by effectively controlling the above disease progression and prolonging survival.
However, these benefits are being compromised by the development of drug resistance, a consequence of mutations that merge in the viral proteins targeted by antiretroviral agents.
The goal of resistance testing is to identify which drugs will be helpful in management of HIV infection (Paul and Jorden, 2003). There are two techniques available in evaluating HIV drug resistance; genotypic and phenotypic techniques (Paul and Jorden, 2003). Genotypic technique involves a molecular biology approach whereby the HIV genetic material isolated from the patient is amplified by polymerase chain reaction using specific primers and then sequenced. The generated sequences are examined for the presence of specific mutations on various parts of the genome. The analysis is generally confined to the specific areas of the gene (codons) representing proteins where antiretroviral medications are known to exert their action, such as the various proteases. The assays clearly indicated that some patients HIV genome had a pattern associated with resistance to a specific antiretroviral agent. These drugs should be purposely avoided in therapy as advocated by Paul and Jorden (2003).

The other technique for evaluating antiretroviral drug resistance is known as phenotypic testing. It involves evaluating the ability of the HIV virus to multiply in culture when subjected to various antiretrovirals. The virus is grown in culture using peripheral blood mononuclear cells or similar media, after which it is mixed with differing concentrations of antiretroviral agents. Viral growth is then examined, leading, in turn, to calculations of inhibitory concentrations, such as IC 50 and IC 90. This type of testing is routinely performed for bacteria and other human pathogens. Results of phenotypic testing are drawn directly from observation of viral growth following exposure to drugs. Accordingly, phenotypic analysis is considered by many to better reflect the true response of HIV to a specific agent (Fauci et al., 2002).
2.7 Evolution and Mechanisms of Drug Resistance

The rapid development of drug resistant HIV strains has been linked to the high rate of HIV replication. The ultimate size of a viral population containing a mutation is probably determined by three concurrent factors: the forward mutation frequency, the replicative capability of the mutated virus and the ‘age’ of the viral population containing the mutation that is how long ago this population was generated (Gilks et al., 2006). Approximately 10 million new dissimilar viral particles are produced every day through replication. Some of these could be due to the exceptionally high error rate of HIV reverse transcriptase (Perelson et al., 1996). This, therefore, leads to a high mutation rate and constant production of new viral strains, especially in the absence of treatment. On the other hand, in the presence of antiretroviral drugs, resistant strains are selected for as the dominant species (Drake, 1993). Continued HIV-1 replication in the presence of selective pressure of drugs targeting the reverse transcriptase viral enzyme leads to emergence of specific point mutations in the RT genomic region of the polymerase gene. This genetic variation inevitably gives rise to mutations that confer resistance to antiretroviral drugs (Zdanowicz, 2006).

Most mutations are either lethal or neutral and do not confer drug resistance. However, under conditions in which treatment does not completely inhibit viral replication, viruses with drug resistant mutations can develop and replicate, resulting in treatment failure. For example, the M184V confers Lamivudine (3TC) resistance, L74V is associated with Didanosine (ddI) resistance, T69D is associated with Zalcitabine (ddC) resistance, and Y181C is associated with most non-nucleoside RT inhibitor resistance (Harrigan et al., 2001). The T215Y mutation is most often associated with Zidovudine (AZT) resistance, but other point mutations, such as M41L, D67N, L210W, and K219Q, also contribute to AZT resistance (Kellam et al., 1992; Larder et al., 1989). While different mutations are typically associated
with resistance to specific drugs, there is a variable amount of cross-resistance conferred by each mutation (Harrigan et al., 2001).

A single mutation can confer a high degree of resistance to one or more NNRTIs. The relatively frequent K103N mutation leads to a 20 to 30-fold increase in resistance to all available NNRTIs (Petropoulos et al., 2000). In contrast to subtype B viruses, the mutation V106M is more frequent in subtype C viruses. V106M is associated with high-level resistance not only to Nevirapine but also to Efavirenz (Grossman et al., 2004). G190S and Y188C/L/H are mutations resulting in a high degree of Nevirapine and Efavirenz resistance (de Mendoza et al., 2002).

For several Nucleoside Reverse Transcriptase Inhibitors (NRTIs), such as Lamivudine, and for Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs), a high degree of resistance can develop following only a single mutation (Havlir et al., 1996; Schuurman et al., 1995). For this reason, such drugs should only be used in highly effective regimens. Emtricitabine (FTC), an NRTI approved in 2003, has the same resistance pattern as Lamivudine (3TC). Treatment failure is also associated with the M184V mutation (Rousseau et al., 2003). Thymidine Analogue Mutations mostly referred to as “TAMs”, include the mutations M41L, D67N, K70R, L210W, T215Y, and K219Q, which were initially observed with Zidovudine therapy (Larder and Kemp, 1989). It is now known that these mutations can also be selected for by Stavudine (Loveday et al., 1999). Three or more TAMs are associated with a relevant reduction in the sensitivity to Stavudine (Lafeuillade and Tardy, 2003; Calvez et al., 2002; Shulman et al., 2001).
Resistance to PIs is mediated by the appearance of protease amino acid substitutions (at positions either in direct contact with the inhibitor or at distant sites) that reduce the binding affinity between the inhibitor and the mutant protease enzyme (Clavel and Hance 2004). These amino acid substitutions, defined as major mutations, may deeply impair the protease catalytic activity and, consequently, the replication capacity of the virus (Chen et al.; Lin et al., 1995). The immune system is affected by substitutions of even a single amino acid in viral sequences that specify immunogenic epitopes making the virus to easily escape surveillance by the immune system. Moreover, cells expressing such altered peptides can even block the lysis of cells infected with unmutated virus. It is still uncertain whether immune pressures select for mutations that allow the virus to escape from cytotoxic T lymphocytes (so-called escape mutations) or whether the evasion is simply incidental to the high rate of mutations. Restoration of the replication capacity allows the viability of virus particles due to the presence of such mutations hampering the therapeutic efficacy of the drugs (Hertogs et al., 1999; Chen et al., 1995).

Generally, single mutations confer only a modest reduction in drug susceptibility (with some exceptions, such as D30N, I47A, G48V, and I50L/V), while for many PIs, particularly if they are boosted with ritonavir, a stepwise accumulation of several protease mutations is required for the development of high-level drug resistance. This requirement of multiple mutations to overcome the activities of PIs has been referred to as a high genetic barrier to drug resistance (Condra et al., 1996; Molla et al., 1996). Furthermore, most drug resistance mutations in the protease confer cross-resistance to many PIs, so this resistance should be considered class specific rather than drug specific (Kozal, 2004; Kemper et al., 2001; Hertogs et al., 2000; Race et al., 1999). Since protease is characterized by a high degree of structural flexibility and resistance to PIs is such complex phenomenon, it is conceivable that the mutational
pathways of HIV protease are far more complex than is currently known and that some mutations (and associations of mutations) have been not yet precisely defined.

2.8 Challenges to Expanded ART in Resource-Limited Settings
The complexity and the open-ended duration of HIV treatment coupled with the urgent need to begin programs quickly to treat millions of infected individuals. Although fears have been raised that drug resistance could develop in countries with limited resources and spread rapidly, thus quickly rendering anti-HIV drugs useless (Little et al., 2002). All the same, despite the fact that drug resistance does not spread rapidly in HIV, and that improvements in treatment programs can lead to decline in transmitted resistance, minimising drug resistance on a programmatic level is important in all countries.

2.9 Drug Resistance Monitoring
Continuous testing for drug resistance could provide information to support public health bodies in designing education and prevention programs in order to minimise the development and transmission of drug-resistant viruses, and to support the rational use of antiretroviral drugs by treatment programs, clinicians and policy makers. Methods for limiting resistance, should include starting HIV treatment at the appropriate time, using standard and appropriate regimens, ensuring drug supply, supporting adherence, and regular clinical monitoring of patients to ensure effectiveness of ART program. To this end, several assays have been described to monitor/investigate HIV-1 resistance by both genotypic and phenotypic techniques (Re et al., 2001). It should be noted that the more traditional phenotypic assays, which are cell culture-based, are labour intensive and require several weeks to perform. In contrast, genotypic assays are rapid and can be performed directly on patient samples (Re et al., 2001).
CHAPTER THREE: MATERIALS AND METHODS

3.1 Study Sites and Study Population

The study was conducted in western Kenya at Busia District hospital. This facility serves residents of the Kenyan border as well as those from the neighbouring country Uganda. Busia is a crossing border point to Uganda and is characterized by rampant human traffic; cross-border movement of people, trade and it also serves as a gateway to central Africa where all HIV-1 subtypes have been documented. The study participants were drawn from a cohort of over 1,000 patients receiving antiretroviral therapy. The cohort was subjected to a set of inclusion and exclusion criteria, and random numbers were used to pick those meeting the criteria. The participants were drawn from within western Kenya (Figure 3.1) codes 801-803.

Figure 3.1: Map of Kenya showing western province (code 801-803) location of sampling site. Adapted from http://www.mapsofworld.com/kenya/kenya-political-map.html.
3.2 Study Design

This was a randomized cross-sectional study conducted in a hospital set-up using a cohort of patients on antiretroviral therapy. Data collection involved the use of a structured questionnaire by interview method.

3.3 Inclusion and Exclusion Criteria

Inclusion criteria involved selecting patients from the cohort who have been on therapy for at least 6 months. Other requirements were: CD4 cell counts at the time of recruitment to the study must have been less than the one at initiation of therapy over the past 6 months based on existing hospital records, and an increasing plasma viral load (> 5,000 copies/ml) as per the existing records during sampling. On the other hand, those patients who were newly initiated to therapy and were responding with increasing CD4 cell counts and decreasing viral loads were not included in the study.

3.4 Sampling and Sample Size Determination

Random numbers were used to pick participants during clinic visits at the time of sampling. Information about this study was extensively explained to patients (Appendix I). Those participants who were found eligible based on the set inclusion criteria had Ethical considerations reviewed by the site physician. Those who were cleared were then requested to consent to participate in the study. Consent forms (Appendix II) were filled and signed by all the patients recruited. A structured questionnaire was used to capture bio-data of the 75 selected patients (Appendix III). The patients were interviewed at the time of sample collection at the hospital. A unique participant code was assigned to each patient sample, and recorded.
The sample size was determined using the formula deduced by (Fisher et al., 1998) as follows:

\[ n = p \times (100-p) \times z^2 / d^2 \]

Where:

- \( n \) = minimum sample size required.
- \( z \) = 1.96 standard error.
- \( p \) = 0.05; (5%) a demonstrated percentage of those who were infected with drug resistance strains between 1996 and 2002 (Hackett et al., 2003).
- \( d \) = 0.05 the inverse of 95% confidence limit (the allowable error).

Substituting appropriately, the minimum sample size will be as follows:

\[ n = 0.05 \times (1-0.05) \times 1.96^2 / 0.05^2 \]

Therefore, the working sample size was 75.

Among the 1,000-cohort population, the 75 study participants were randomly selected from the attendees at the various clinic visits. Then informed consent was obtained from eligible participants.

### 3.5 Laboratory Methods

#### 3.5.1 Blood Collection

This cross-sectional study involved a one-time blood draw (5 millilitres) from the study participants. The blood was collected in 10ml EDTA laced tubes in the health facility. It was then transported to the Kenya Medical Research Institute HIV laboratories in Nairobi in ice-packed boxes for further analysis. All samples were transported by courier services overnight to the central laboratory in Nairobi.
3.5.2 Isolation of Peripheral Blood Mononuclear Cells (PBMCs) from Whole Blood

The whole blood samples were delivered to the HIV laboratories within 2 days. Here, peripheral blood mononuclear cells were isolated from whole blood following methods described by (Boyum, 1968). Briefly, all the 5 ml of whole blood were mixed with 0.84% ammonium chloride, vortexed and incubated at 37°C for 10 minutes. This was then centrifuged to obtain a white pellet of cells after washing several times. The PBMCs pellet was then stored at 4°C until further use.

3.5.3 Extraction of Deoxyribonucleic Acid from Peripheral Blood Mononuclear Cells

Deoxyribonucleic acid was extracted from peripheral blood mononuclear cells using methods described by the QIAmp RNA core kit from PE Applied Biosystems, Foster City, CA, USA. In summary, the PBMC pellet was dissolved in DNAzol genomic DNA extraction reagent. Absolute ethanol was used to precipitate DNA. The DNA obtained was then dissolved in DNAse and RNAse free water and stored at -20°C until further use.

3.5.4 DNA PCR for Genotyping

Genotypic testing of HIV pol region was done to determine the circulating HIV-1 subtypes and the occurrence of drug-associated mutations among patients on therapy. This information would also be important in establishing the potential for transmission of resistance, and of the mutations likely to be transmitted. The starting template was proviral DNA extracted from PBMCs. All reagents were thawed and put on ice.

Depending on the number of samples the following conditions were used. A region of the HIV-1 pol gene including the reverse transcriptase sequence (Pol-RT region; corresponding to nucleotide positions 2513-3209 in HIV-1_HXB2 was amplified by nested PCR with primers
(forward primer) NYUPOL7 (5’-GGGAATTTTCTTCAGAGCAG-3’), and (reverse primer)
NYUPOL8 (5’-TCTTCTGTCATGGCCATTGT-3’) in the first round, and (forward primer)
NYUPOL9 (5’-TCCTTAACTTCCCTCAAATCCT-3’) and (reverse primer) NYUPOL10
(5’-CTGGCACCGGTTCATAGGACT-3’) in the second round. Appendix IV (Konings et al., 2004).

The first round PCR was carried out in 25µl tube containing 2µl of genomic DNA, 2µl of
10X buffer, 0.4µl of 2ng each reverse and forward primer, 2µl of 10mM dNTPs, 2.8µl of
25mM MgCl₂, 10.2µl of distilled water and 0.2µl of 5u/µl Taq polymerase. The amplification
was carried out in a thermal cycler (MJ Research, Inc). First amplification was done with 1
cycle of 95°C for 10 minutes and 35 cycles of 95°C for 30 seconds, 55°C for 60 seconds, and
72°C for 1 minute, with a final extension of 72°C for 10 minutes. From the first round-PCR
products, 2µl were used as a template for the nested PCR with inner primers and the same
cycling conditions as the first round-PCR were used (Ping et al., 2003).

**3.5.5 Analysis of PCR Products by Gel Electrophoresis**

Once the 2nd PCR (nested) was done, all amplification products were analysed by
conventional 2% agarose gel electrophoresis. Electrophoresis was done at a constant voltage
of 100 volts/cm for 30 minutes. After electrophoresis, the gels were stained with 0.5µg/ml of
Ethidium Bromide. The location of PCR DNA fragments on the gels was determined by
direct examination of the gel under Ultra Violet light and size estimated by comparing with
molecular weight markers loaded alongside them.
3.5.6 Sequencing PCR

Sequencing PCR utilized the products from the nested PCR as the starting template. Depending on the number of samples to be sequenced, BigDye version 3.1 from PE Applied Biosystems, Foster City, CA, USA was diluted in appropriate buffer in the ratio of 6 parts 5X buffer to 5 parts of BigDye. Amplification of the HIV *pol* gene was done using the forward primer in the 2nd round PCR (NYUPOL 9) (5’-TCCTTAACTTCCCTCAAATCACT-3’). The sequencing PCR was carried out in 25μl tube containing 1μl of DNA, 1μl of BigDye terminator premix, 1.5μl forward primer, 3.5μl of 5X sequencing buffer, and 13μl of distilled water. The amplification was done with 1 cycle of 96°C for 5 minutes and 25 cycles of 96°C for 10 seconds, 50°C for 5 seconds, and 60°C for 4 minutes.

3.5.7 Purification of PCR Products for Sequencing

In summary, 20μl of the dye-labeled PCR products were added into an eppendorf tube containing 2μl of 3mM sodium acetate and 50μl of absolute ethanol. This was then incubated at room temperature for 15 minutes in the dark. The tube was centrifuged at 14,000g for 20 minutes at room temperature. 500μl of 70% ethanol was added and centrifuged at 14,000g rpm for 10 minutes and the supernatant discarded. The two steps above were repeated and DNA air-dried for at least 45 minutes. 25μl of formamide was added and heated at 95°C for 2 minutes. The contents were transferred into sequencing tubes ready for sequencing.

3.5.8 HIV Sequencing

Seventy-five samples were successfully sequenced targeting a region of the HIV-1 *pol* gene including the reverse transcriptase sequence (*pol*-RT; corresponding to nt 2513-3209 of HIV-1HXB2). The nucleic acid sequences were submitted directly to HIV drug resistance database.
and examined for protease inhibitors associated substitutions, by comparing to references in HIV drug resistances database using HIV seq (HIV search engine for queries Appendix V) of the Stanford HIV drug resistance database at (http://hivdb.stanford.edu/). The results generated were then analysed. Copies of the sequences generated in FASTA format are attached (Appendix VII).

### 3.5.9 Sequence Analysis

Genotypic resistance testing involved the *pol* nucleotide sequence (297bp) of the HIV-1 *pol* region where nucleic acid sequences were examined for protease inhibitor resistance-associated substitutions using HIV seq (HIV search engine for queries) of the Stanford HIV drug resistance database (http://hivdb.stanford.edu). For subtyping, sequences generated were aligned against a set of reference strains from all known HIV-1 group M subtypes (gathered from the Los Alamos HIV sequence database (http://hiv-web.lanl.gov) using CLUSTAL W profile alignment option. The possible countries of origin of the subtypes were also determined by BLAST analysis. BLAST is a computer based bioinformatics program known as Basic Local Alignment Search Tool. It is an algorithm for comparing biological sequence information of different nucleotides of DNA sequences. This program enables comparison of query sequences and library/database of sequences that resemble the query sequences above a certain threshold. BLAST is hosted by the National Library of Medicine, a branch of National Institute of Health of the United States of America. The program was accessed through: http://www.hiv.lanl.gov/content/sequence/BASIC_BLAST/basic_blast.html. The BLAST results were then summarized (Table 4.1). Genetic distances were calculated by the two-parameter method of Kimura and the phylogenetic tree constructed by neighbour-joining method with its reliability being estimated by 1000 bootstraps (Kumar *et al.*, 2001). The tree profile was visualized with Tree View PPC version 1.6.5.
CHAPTER FOUR: RESULTS

4.1 Confirmation of HIV *pol* Region PCR Products Amplified From Study Samples

The amplified protease genes in nested PCR were identified by gel electrophoresis and bands visualized for successful amplification by presence of bands of DNA of the right size (Figure 4.1).

![Figure 4.1: Agarose gel indicating the sizes of the amplified products of the PCR done](image)

**Legend:**
- Lane 1: Molecular weight marker.
- Lane 2: Test sample.
- Lane 3: Negative control (a known HIV negative sample).
- Lane 4: Test sample.
- Lane 5: Positive control (a known HIV positive sample from previous PCR tests)
- Lanes 6-8: Test samples

The PCR products were 297 base pairs in size.
4.2 Sequence Analysis

All the HIV-1 subtypes identified and their respective countries of origin were summarized in Table 4.1

<table>
<thead>
<tr>
<th>Sequence name</th>
<th>Subtype</th>
<th>Country with similar subtype*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus001</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus002</td>
<td>A1C</td>
<td>Tanzania</td>
</tr>
<tr>
<td>Bus003</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus004</td>
<td>A1D</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus005</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus006</td>
<td>A1</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus007</td>
<td>A1</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus008</td>
<td>A1</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus009</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus010</td>
<td>G</td>
<td>Congo</td>
</tr>
<tr>
<td>Bus011</td>
<td>A1</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus012</td>
<td>A1</td>
<td>Tanzania</td>
</tr>
<tr>
<td>Bus013</td>
<td>C</td>
<td>Burundi</td>
</tr>
<tr>
<td>Bus014</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus015</td>
<td>A1</td>
<td>Tanzania</td>
</tr>
<tr>
<td>Bus016</td>
<td>G</td>
<td>Congo</td>
</tr>
<tr>
<td>Bus017</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus018</td>
<td>C</td>
<td>South Africa</td>
</tr>
<tr>
<td>Bus019</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus020</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus021</td>
<td>A1</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus022</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus023</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus024</td>
<td>A1</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus025</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Sequence name</td>
<td>Subtype</td>
<td>Country with similar subtype*</td>
</tr>
<tr>
<td>---------------</td>
<td>---------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>Bus026</td>
<td>A1</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus027</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus028</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus029</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus030</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus031</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus032</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus033</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus034</td>
<td>G</td>
<td>Congo</td>
</tr>
<tr>
<td>Bus035</td>
<td>A1</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus036</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus037</td>
<td>A1</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus038</td>
<td>D</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus039</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus040</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus041</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus042</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus043</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus044</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus045</td>
<td>A1</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus046</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus047</td>
<td>G</td>
<td>Congo</td>
</tr>
<tr>
<td>Bus048</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus049</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus050</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus051</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus052</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus053</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus054</td>
<td>A1</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus055</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Sequence name</td>
<td>Subtype</td>
<td>Country with similar subtype*</td>
</tr>
<tr>
<td>---------------</td>
<td>---------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>Bus056</td>
<td>A1</td>
<td>Tanzania</td>
</tr>
<tr>
<td>Bus057</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus058</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus059</td>
<td>A1</td>
<td>Tanzania</td>
</tr>
<tr>
<td>Bus060</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus061</td>
<td>G</td>
<td>Congo</td>
</tr>
<tr>
<td>Bus062</td>
<td>A2</td>
<td>Congo</td>
</tr>
<tr>
<td>Bus063</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus064</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus065</td>
<td>C</td>
<td>Burundi</td>
</tr>
<tr>
<td>Bus066</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus067</td>
<td>A1</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus068</td>
<td>A1</td>
<td>Kenya</td>
</tr>
<tr>
<td>Bus069</td>
<td>C</td>
<td>Botswana</td>
</tr>
<tr>
<td>Bus070</td>
<td>A2</td>
<td>Congo</td>
</tr>
<tr>
<td>Bus071</td>
<td>A1</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus072</td>
<td>B</td>
<td>Brazil</td>
</tr>
<tr>
<td>Bus073</td>
<td>D</td>
<td>Uganda</td>
</tr>
<tr>
<td>Bus074</td>
<td>C</td>
<td>Burundi</td>
</tr>
<tr>
<td>Bus075</td>
<td>A1</td>
<td>Uganda</td>
</tr>
</tbody>
</table>

*Countries with very close documented isolates.*
Information generated (Table 4.1) was further analyzed to determine the most common subtypes and countries they are thought to have originated based on the BLAST analysis.

Out of the 75 samples sequenced, most isolates were of HIV-1 group M subtype A1 (n=39), and subtype D (n=21) (Figure 4.2). The rest had subtypes with not more than 5 isolates. Findings from this study indicate that most isolates originated from Uganda (n=35), more so subtype D (n=20) and AI (n=15). This could be attributed to the rampant cross-border movement of people. Over 88% (n=66) of the subtypes detected are thought to have originated from within East Africa. A total of 35 isolates resembled those previously reported from Uganda in the GenBank, while 23 isolates resembled those from Kenya and only 5 isolates resembled those from Tanzania. All the subtypes detected originated from within Africa except one isolate was of HIV-1 subtype B with its origin traced back to Brazil.

![Figure 4.2: Distribution of HIV-1 group M subtypes isolated from the study and respective countries of origin.](image)
On the basis of the part of the *pol* gene (297bp), the sequence of 64 samples were further analysed by neighbour joining method (Figure 4.3). Study samples are shown in blue while the GenBank references are shown in black for they clustered very closely. Phylogenetic relationships of newly derived viral sequences were estimated from comparisons with those of previously reported HIV-1 group M from the Los Alamos sequences database using the CLUSTAL W profile alignment option (Rhee *et al.*, 2003). All the 75 samples were used to generate the phylogenetic tree, however, only 64 samples aligned with reference sequences with a similarity of over 98%. Genetic distances were calculated by the two-parameter method of Kimura and the phylogenetic tree constructed by the neighbour-joining method with its reliability being estimated by 1000 bootstraps. The tree profile was visualized with Tree View PPC version 1.6.5 (Figure 4.3).
Figure 4.3: Phylogenetic tree showing the HIV-1 subtypes isolated from the study participants. HIV-1 protease region was amplified and sequenced.
The distribution of HIV-1 group M subtypes among the study population was very diverse (Figure 4.4). Results from the study indicate that HIV-1 subtype A1 was the most predominant subtype comprising 52% (n=39) of the isolates. Other subtypes detected included subtype D 28% (n=21), A2 2.7% (n=2), while subtypes G and C circulated within the study population at equal levels of 6.7% (n=5) for both subtypes. At very low distributions were subtypes A1C, A1D, and subtype B which were detected singly with a prevalence of 1.3% (n=1).

![Figure 4.4: Distribution of HIV-1 subtypes isolated from the study participants](image)

The generated sequences were then analyzed for protease resistance-associated substitutions using HIV seq (HIV search engine for queries) of the Stanford HIV drug resistance database available at [http://hivdb.stanford.edu/pages/algs/HIVdb.html](http://hivdb.stanford.edu/pages/algs/HIVdb.html). Raw sequences were uploaded into the database without translating them into amino acids as they can either be analysed as amino acids or directly. When translated to amino acids, artificial errors can be introduced resulting in frame shifts and deletions hence giving wrong information on the
mutations generated. Since the samples were sequenced directly without cloning, the use of raw sequences was preferred to amino acid sequence translations. The results were downloaded in the form of a HIVdb genotypic resistance interpretation algorithm (Appendix V). The possible mutation points and codon positions associated with drug resistance were determined and recorded (Table 4.2).

Table 4.2: Categories of protease-associated mutations detected in study samples

<table>
<thead>
<tr>
<th>Category of protease mutations detected</th>
<th>Codon position affected (number of samples)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major protease mutations</td>
<td>I47L (n=1), V32A (n=2), M46I (n=2), V32L (n=4), D30N (n=1)</td>
</tr>
<tr>
<td>Minor protease mutations</td>
<td>L33I (n=1), G48R (n=2), G73S (n=2), T74S (n=2)</td>
</tr>
</tbody>
</table>

Antiretroviral resistance associated mutations were detected in 12 samples (16%) of the patients (Table 4.3). There is therefore a likelihood of circulating strains of HIV-1 with ARV drug resistance associated mutations in this region. Resistance mutations to protease inhibitors were detected in most HIV-1 subtypes A1 (n=10), subtype G (n=1) and subtype C (n=1). Of the ten subtype A1 isolates with resistance mutations, 6 resembled isolates from Uganda while 4 resembled isolates from Kenya. The subtype G isolates with resistance mutations resembled isolates from Congo while that of subtype C resembled isolates from Botswana. The major protease inhibitor mutations detected include: I47L, V32A, M64I, V32L and D30N. A number of minor protease inhibitor mutations were also detected. These include: V32I, L33I, G48R, G73S and T74S (Table 4.2). There were no drug resistance associated mutations in subtypes D, A1C, A1D, A2, and B isolates. Major protease mutations V32A, V32L and M46I were more frequent appearing in nine isolates. Two subtype A1
isolates and one subtype C isolate had both major and minor mutations. Two subtypes A1 isolates had only minor mutations (Table 4.3).

Table 4.3: Protease Inhibitor Associated Mutations and Drug Susceptibility Patterns.

<table>
<thead>
<tr>
<th>Sample Identifier</th>
<th>HIV-1 Subtype</th>
<th>Main Drugs</th>
<th>Protease Inhibitor</th>
<th>Other Mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus008</td>
<td>A1</td>
<td>S</td>
<td>ATV/r</td>
<td>R41K, H69K, L89M</td>
</tr>
<tr>
<td>Bus010</td>
<td>G</td>
<td>S</td>
<td>DRV/r</td>
<td>M36I, R41K, L63Q, 164V</td>
</tr>
<tr>
<td>Bus019</td>
<td>A1</td>
<td>LR</td>
<td>FPV/r</td>
<td>T31I, E35D, M36I, R41K, D60E, Q61E, L63P, H69K, L89M</td>
</tr>
<tr>
<td>Bus030</td>
<td>A1</td>
<td>LR</td>
<td>LPV/r</td>
<td>M36I, R41K, R57K, H69K, L89M</td>
</tr>
<tr>
<td>Bus036</td>
<td>A1</td>
<td>LR</td>
<td>NFV</td>
<td>M36I, R41K, R57K, H69K, L89M</td>
</tr>
<tr>
<td>Bus044</td>
<td>A1</td>
<td>LR</td>
<td>SQV/r</td>
<td>L19V, E35D, M36I, R41K, R57K, H69K, L89M</td>
</tr>
<tr>
<td>BSA045</td>
<td>A1</td>
<td>LR</td>
<td>TPV/r</td>
<td>L19V, E35D, M36I, R41K, R57K, H69K, L89M</td>
</tr>
</tbody>
</table>

**KEY:** S: Susceptible; LR: Low-level resistance; IR: Intermediate resistance; HR: High-level resistance

The drug with the least level of resistance was Tipranavir®. Nelfinavir® had the highest level of resistance thus high chance of therapy failure (Table 4.3).
CHAPTER FIVE: DISCUSSIONS

5.1 Study Overview

In this study, genotyping was used to identify circulating HIV-1 subtypes and drug resistance codons among patients on antiretroviral therapy in Busia District hospital. The process involved amplifying the HIV-1 positive DNA using specific set of primers (Appendix IV). Genotypic technique was preferred because the assays can be performed relatively rapidly, and that results can be reported within one or two weeks of sample collection (Fauci et al., 2002), this method was found to be very useful in resource limited settings such as Busia.

HIV-1 is characterized by a high degree of genetic variability, mainly due to the intrinsic inability of HIV-1 reverse transcriptase to carry out proofreading of DNA during replication (Mansky, 1998), and is exacerbated by the high rate of HIV-1 replication in vivo, the accumulation of proviral variants, and genetic recombination (Burke, 1997). In this study, part of the pol gene (297bp), which encodes enzymes such as reverse transcriptase and protease was amplified using specific primers (Appendix IV). The HIV-1 pol region is subjected not only to natural variation pressure but also to mutations imposed by the pharmacological treatment (Clavel and Hance 2004; Shafer, 2002). The HIV-1 pol region which encodes for the protease enzyme was selected for amplification and evaluation of drug resistance-associated mutations because it is responsible for the posttranslational processing of the viral polyproteins encoded by gag (p55) and gag-pol (p160), thus yielding the mature structural core proteins (the matrix, capsid, nucleocapsid, and p6 proteins) and the essential enzymes (protease, reverse transcriptase, and integrase) required to produce mature infectious particles (Turner and Summers 1999). For its crucial role in the HIV-1 life cycle, protease represents an important target for the antiretroviral therapy. To date ten protease inhibitors (Table 2.2) have been approved by the Food and Drug Administration (FDA) and are
clinically available. Unfortunately, when antiretroviral therapy fails to be fully suppressive, viral variants with reduced susceptibilities to PIs can emerge (Miller, 2001; Hertogs et al 2000; Condra et al., 1995).

5.2 HIV-1 Subtype

Results from this study show that like other parts of the country the predominant circulating subtype in Busia is A1. This is also in conformity with the report of Lihana et al (2006), based on the pol RT region of HIV-1 isolates from patients attending the Nairobi STI clinics. Other studies show a similar pattern: a near uniform epidemic in the country except for Northern border regions where subtype C predominates due to cross-border transmission from Ethiopia (Khamadi et al., 2005). Most of the subtype A1 isolates clustered closely with isolates documented previously in Uganda. This has been linked to the rampant cross-border migration and human traffic and other socio-economic activities that may have led to the introduction of A1 subtype to Kenya. Uganda has predominantly subtype A1 (Rayfield et al., 1998). BLAST analysis was used to determine the HIV-1 subtypes circulating in Busia. Of the 75 samples successfully amplified and sequenced, 39 were HIV-1 group M subtype A1 (52%), 21 subtype D (28%), 5 subtype G (6.7%), 5 subtype C (6.7%), 2 subtype A2 (2.7%), whereas subtypes B, A1C, A1D, had each only 1 isolate (1.3%). Due to such diversity, surveillance studies on the extent of inter-subtype recombination are necessary for continuous improvement of diagnosis, treatment and vaccine development.

HIV/AIDS is characterised by enormous genetic flexibility, which gives rise to drug resistance, escape from immune system responses and failure of vaccination attempts. A number of factors are thought to contribute to viral diversity in individual infections. It is clear that the high error rate of reverse transcriptase (Ji and Loeb, 1994) and the high turnover rate in vivo (Wei et al., 1995; Ho et al., 1995) generate vast numbers of different virus
mutants. The diversity of viral quasi-species, however, is shaped by a combination of mutations and selection forces. The main selective forces that have been proposed to drive HIV diversity are the immune response, cell tropism and random activation of infected cells. Whenever there is a decrease in CD4 cell count, there is a strong positive selection for amino acid change that could lead to mutants.

5.3 Mutations and their Significance

Early in the infection, the immune system responds strongly against common viral variants and hence favours rare mutants thereby providing a strong positive selection for diversification. As the immune system declines, this selection pressure becomes weaker. The pattern for decreasing positive selection however, is also consistent with the notion that most viral diversity is caused by adaptation for various cell tropism (Boyd et al., 1993). At seroconversion, the patient carries a strong homogenous virus population, and then diversification occurs as HIV infects many different cell types and tissues in the body. Initially, this would provide a strong positive selection pressure, which would decline when the virus has generated many variants with specific cell tropism. Selection pressure is significantly lower in T-cell-line adapted sequences than the macrophage-tropic sequences. It is thought that the immune system acts strongly against the T-cell-line adapted variants (Bou-Habib et al., 1994). It is therefore plausible that immune selection has greater effect on diversity than cell tropism although the two factors may be interlinked. More data of this kind are urgently needed to enable understanding of HIV evolution in individual infections and its consequences for pathogenesis.

There is a standard numbering system for HIV-1 protease and RT based on their amino acid sequences. The most commonly used wild-type reference sequence is the subtype B consensus sequence. This sequence was originally derived from alignments in the HIV
Sequence Database at Los Alamos National Laboratory (Rhee et al., 2003). Mutations are typically described using a shorthand notation in which a letter indicating the consensus B wild-type amino acid is followed by the amino acid residue number, followed by a letter indicating the mutation (eg, T215Y). If there is a mixture of more than one amino acid at a position, the components of the mixture are written after the position, often separated by a slash (eg, K103K/N denotes that the sequence has a mixture of the wild-type residue lysine (K) and the mutant residue asparagine (N) at position 103). Because so many mutations in both the protease and RT have been associated with drug resistance, it has become customary to label some drug resistance mutations as either "primary" or "major" and other mutations as "secondary" or "minor". Primary mutations are those that reduce drug susceptibility by themselves whereas secondary mutations reduce drug susceptibility in combination with primary mutations or improve the replicative fitness of virus isolates with a primary mutation that is allowing viability of the virus particle instead of inhibiting replication.

Most of the mutations detected in this study are similar to those previously characterised in non-B subtypes (Kantor et al., 2005). It has been recognised that some minor PR mutations may confer reduced susceptibility to protease inhibitors (PI) in vitro (Parkin et al., 2005). A total of 4 Subtype A1 isolates from this study were found to have V32L and T74S mutations, thus reduced susceptibility to Nelfinavir (Cane et al., 2001; Suguiro et al., 2002; Grossman et al., 2004; Sukasem et al., 2008). One isolate of HIV-1 subtype C had D30N, G48R and G73S mutations which are known to be responsible for a high-level resistance to Nelfinavir (Grossman et al., 2004; Sukasem et al., 2008). On the other hand, a number of studies have extensively shown that some non-B subtypes select for the mutation D30N in PR, implicated in resistance to PI Nelfinavir, to a much lower extent than subtype B (Holguin et al., 2004; Parkin et al., 2005). It must be noted that D30N mutation is useful in that it impairs the replication of the virus. Other minor mutations such as M36I and K20I in protease, which
were also detected in this study, are common to some patients not previously exposed to ARVs, especially non-B subtypes (Holguin et al., 2006). These mutations are disadvantageous because they modulate the replication capacity (RC) of subtypes G and CRF02_AG viruses in a complex manner (Holguin et al., 2006; Ho et al., 2008). For example, a virus strain with M36I mutation replicates faster than its wild-type counterparts in the absence of protease inhibitors (Holguin et al., 2006).

5.4 Transmitted Drug Resistance Mutations

In Kenya, little is known about mutations conferring resistance to ARVs in either treated patients or treatment naïve patients. Transmission of drug resistance from treatment-experienced patients to newly infected persons has been observed repeatedly in countries with access to antiretroviral therapy (Puig et al., 2000). It is estimated that 8–30% of people who contract HIV infection in Europe or the United States of America acquire a virus with drug resistance conferring mutations (Wensing and Boucher, 2003; Yerly et al., 2001). Such transmission of drug-resistant strains increases despite all prevention efforts, raising major public health concerns (Wensing and Boucher, 2003). As a result, clinical management of patients who have never been exposed to therapy may be unfavourably affected (Durant et al., 1999). As stated inter alia, transmitted drug resistance ranges from 8–30%. In this study, 16% of the samples showed an aspect of transmitted drug resistance.

The commonly detected mutations in the samples analysed were D30N, L33I, G73S, T74S, V32I, and M46I. It is recognised that some major protease mutations may confer reduced susceptibility to protease inhibitors. Based on the above polymorphism frequencies of the protease regions in the HIV-1 variants isolated, different viral types, groups and subtypes in this area may harbour distinct mutations, which could alter interaction with protease drugs. For instance, V32I a substrate cleft mutation detected in subtype A1 (Bus019) is associated
with reduced susceptibility to all protease inhibitors except SQV/r and TPV/r (Baxter et al., 2006). Other mutations such as M46I detected in subtype A1 (Bus026) isolate decreases susceptibility to Indinavir, Nelfinavir, Fosamprenavir, Lopinavir and Atazanavir when present with other mutations. In addition, D30N mutation detected in subtypes C (Bus069) isolate causes high-level resistance to NFV and potential low-level resistance to ATV/r (Baxter et al., 2006; Konings et al., 2004).

5.5 Susceptibility Profiles in Relation to Subtype Diversity

Results from this study indicate that HIV-1 subtype diversity continues to vary with time. Earlier studies in western region showed just HIV-1 subtypes and possible recombinations between the circulating subtypes (Songok et al., 2003). According to Songok et al, 2003, western Kenya was noted to be a potential hot spot for HIV-1 recombination in Kenya. Results from this study tend to reveal a greater diversity of the circulating subtypes. For instance, HIV-1 subtype A2, and subtype B were not reported to circulate in this region at that time. Surprisingly, the prevalence of subtype D tends to increase making it a dominant subtype with time. In addition, a number of new subtype recombinations were documented in this study yet they were never isolated previously. Examples include subtypes A1C and A1D. Among the several HIV-1 subtypes isolated, it was noted that subtype A1 had most isolates showing occurrence of mutations to protease inhibitors It should be noted that specific drug associated mutations could either be a result of genetic variability of HIV (polymorphisms) which code for distinct amino acid residues in ARV target protein of the virus or an aspects of transmitted resistant strains of the virus from persons previously exposed to protease inhibitor-based therapy. Ultimately, this may limit treatment options for patients on first-line regimens.
CHAPTER SIX: CONCLUSIONS, RECOMMENDATIONS AND LIMITATIONS

6.1 Conclusions

a) HIV-1 Subtypes

In Busia, the most dominant subtype is A1 (52%), followed by subtype D (28%) among others. This shows that the Kenyan border point to Uganda is characterized by even more genetically diverse strains of HIV.

b) Prevalence of Protease Associated Mutations

Antiretroviral resistance associated mutations are in circulation among patients on antiretroviral therapy. Subtypes A1 isolates had most mutations that could confer drug resistance to the protease inhibitors in use today. Mutations to protease inhibitors were detected in 16% of the patients on antiretroviral therapy at Busia District hospital.

6.2 Recommendations

Findings from this study show that molecular epidemiology of HIV-1 subtypes in Kenya is diverse with many subtypes in circulation. Due to the rampant cross-border migration, human traffic and other socio-economic activities that may have led to the introduction of other HIV-1 subtypes to Kenya, there is need to update data on circulating subtypes on a regular basis.

Most of the HIV-1 isolates had mutations that could confer resistance to protease inhibitors. There is need to establish surveillance systems to detect emerging mutations in the patient population that could confer antiretroviral drug resistance, given that these mutations can be transmitted from person to person thus limiting available treatment options.
6.2.1 Applications of the Study Findings

Information accrued from this study shows that there is a high likelihood of circulating strains of HIV-1 with ARV drug resistance associated mutations in this region. Data on molecular epidemiology of HIV in this region has now been updated. This information is useful in the planning of HIV prevention and control strategies. In addition, these results form a basis for support of the rational use of antiretroviral drugs by treatment program planners and individual clinicians.

6.2.2 Future Studies

More studies should be done on the implications of HIV genetic diversity on drug resistance both by phenotypic and genotypic techniques.

6.3 Study Limitations

Genotypic assays are expensive, and require highly specialised skills to interpret. In this context, only a few samples were analyzed. Although direct sequencing using plasma is the gold standard, it has been shown that peripheral blood mononuclear cells can also be reliably used for drug resistance genotyping (Vicenti et al., 2007). Nonetheless, the study serves as a starting point for more such studies involving larger and more controlled research to fully elucidate the impact of HIV genetic variation on ARV therapy in resource-limited settings.
REFERENCES


HIV sequence compendium, 2008.


APPENDIX 1 INFORMATION FOR PARTICIPANTS


Purpose of the study

This is a research study designed to establish the circulating HIV-1 subtypes and extent of drug resistances mutations among patients on therapy.

Sample collection

a) Whole blood will be used to extract DNA for genotypic amplification and sequencing for drug resistance associated substitutions in the protease gene of HIV.

b) Full informed consent will be sought from participants

Benefits

The study subjects will not receive monetary benefits for participating, but viral susceptibility status results will be released to the clinicians for patient management. Data will be provided to ministry of health as part of evaluation of extent of HIV drug resistances levels among patients on therapy before initiating PI-based regimens.

Risk of subject

There are no known major physical, psychological or social risks associated with participating in the study. However, the study subject may feel discomfort from blood draw, which will clear in 3 days.
APPENDIX II INFORMED CONSENT

Title: Genotyping of Human Immunodeficiency Virus Type-1 Drug Resistance In Western Kenya.

Participation: Participation in the study is voluntary. Refusal to participate will involve no penalty or loss of benefits to which you are otherwise entitled. You may discontinue your participation at any time unconditionally. The principal investigator may decide to withdraw you from the study if we are unable to obtain a blood sample from you.

Purpose: We are interested in finding out how many patients currently on antiretroviral therapy may be having mutations that may lead to the development of resistance to these drugs. We will test your blood sample for the presence of mutations that may lead to antiretroviral drug resistance. Your blood will also be tested for HIV sub-types. This knowledge is useful when considering selection of antiretroviral agents for patient treatment and management.

Sample collection: If you agree to participate, or act as a legal representative for an adult to participate, we will ask you some questions about age, sex, village, history of illness and treatment currently involved. A single blood sample of 3.0ml will be drawn from your arm or potentially another site of your body e.g. wrist vein using standard sterile techniques. The blood will then be taken to KEMRI laboratories for investigations mentioned above.

Benefits: If you have mutations that may lead to the development of resistance to the current drugs, this information will be given back to the medical officer useful for your future treatment. Participation in this study will not change your medication until your medical
officer believes it is necessary. You may be referred to other government health facility for further management if need arises.

**Risks:** The risk from participation in this study is minimal. There is possibility of mild discomfort, bruising and very rarely infection at the site where blood is obtained. The technician will use utmost care to minimise any pain or trauma during the blood draw. If the site becomes infected, we can treat that with medication.

**Compensation:** There is no compensation to volunteer for participation.

**Duration of participation:** This study only requires one blood draw and the questionnaire. There is no follow-up or further information needed.

**Confidentiality:** Records relating to your participation in the study will remain confidential. Your name will not be used in any report resulting from this study. All computerized records and results will have no names of the source individuals but will be given codes for identification not your name. All the results of the tests performed will be recorded using the study number only. You will receive a signed copy of this consent form.

**Declaration**
I ----------------------------------------------- (Name optional), of age------, do hereby assent to participate in this Research study that will lead to a better understanding and appropriate management of anti-HIV agents. The study above has been explained to me and that I am free to direct any questions concerning the study now or in future to the concerned officers. I understand that I am being asked to volunteer to participate and that I agree to be interviewed
also, I understand that sample collection is a harmless process to be conducted by qualified personnel. I may withdraw or stop my participation in the study without victimization of any sort.

Subject Name: ------------------------------------------
Subject signature/Thumb print: ------------
Date: -------------------
Study Number: ------------------------------
In-charge/guardian/witness Name: ---------------------------------------------
Witness signature: ------------------------- Date: -------------------
APPENDIX III DATA COLLECTION FORM

Genotyping of Human Immunodeficiency Virus Type-1 Drug Resistance in Western Kenya

Study no: ________________ Date of specimen collection ____________ (dd-mm-yy)

Interviewer: ________________

Village ________________ Sub-location ________________

Location ________________ Division ________________

District ________________ Province ________________

Gender: Male ( ) Female ( )

Age (years/ date of birth) ( ) ( )

Occupation/ additional risk or subgroup factor ______________________________

Marital status: Single □ Married □

Number of wives □ Number of children □

Date of HIV diagnosis: ______________________________ (dd-mm-yy)

Date and result of any previous HIV test (with positive or negative result)

First pregnancy/≥second pregnancy (for pregnant women)

Lab/clinical evidence for recent infection

Lab/clinical evidence of stage of HIV infection
CD4+ T-cell count closest to diagnosis (if available)

Viral load closest to diagnosis (if available)

ARV treatment history (Yes/No/Not known)

Time on therapy:

Current treatment regimen:

Remarks:

Name of physician:

Health facility:

Date:

Signature/official stamp:
APPENDIX IV PRIMER SEQUENCES (PROTEASE REGION)

NYUPOL7 (5’GGGAATTTTCTTCAGAGCAG-3’)

NYUPOL8 (5’TCTTCTGTCAATGGCCATTGT-3’)

NYUPOL9 (5’TCCTTAACCTCCCTCAAATCACT-3’)

NYUPOL10 (5’CTGGCACGTTTCAATAGGACT-3’)

Source: (Konings et al., 2004)
APPENDIX V: SAMPLE HIV DRUG RESISTANCE DATABASE: GENOTYPIC RESISTANCE INTERPRETATION ALGORITHM.

HIVdb Interpretations of one of the sequences for drug resistance by http://hivdb.stanford.edu/

**HIVdb: Genotypic Resistance Interpretation Algorithm**

**Drug Resistance Interpretation: PR**

<table>
<thead>
<tr>
<th>PI Major Resistance Mutations:</th>
<th>D30N</th>
</tr>
</thead>
<tbody>
<tr>
<td>PI Minor Resistance Mutations:</td>
<td>None</td>
</tr>
<tr>
<td>Other Mutations:</td>
<td>None</td>
</tr>
</tbody>
</table>

**Protease inhibitors**

<table>
<thead>
<tr>
<th>Drug</th>
<th>Resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>atazanavir/r (ATV/r)</td>
<td>Potential low-level resistance</td>
</tr>
<tr>
<td>darunavir/r (DRV/r)</td>
<td>Susceptible</td>
</tr>
<tr>
<td>fosamprenavir/r (FPV/r)</td>
<td>Susceptible</td>
</tr>
<tr>
<td>indinavir/r (IDV/r)</td>
<td>Susceptible</td>
</tr>
<tr>
<td>lopinavir/r (LPV/r)</td>
<td>Susceptible</td>
</tr>
<tr>
<td>nelfinavir (NFV)</td>
<td>High-level resistance</td>
</tr>
<tr>
<td>saquinavir/r (SQV/r)</td>
<td>Susceptible</td>
</tr>
<tr>
<td>tipranavir/r (TPV/r)</td>
<td>Susceptible</td>
</tr>
</tbody>
</table>

**PR Comments**

**PI Major**

D30N causes high-level resistance to NFV and potential low-level resistance to ATV/r.

**Mutation Scoring**

<table>
<thead>
<tr>
<th>PR</th>
<th>ATV/r</th>
<th>DRV/r</th>
<th>FPV/r</th>
<th>IDV/r</th>
<th>LPV/r</th>
<th>NFV</th>
<th>SQV/r</th>
<th>TPV/r</th>
</tr>
</thead>
<tbody>
<tr>
<td>D30N</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total:</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
APPENDIX VI: COPY OF ETHICAL APPROVAL PERMISSION

KENYA MEDICAL RESEARCH INSTITUTE

KEMRI/RFS/7/3/1

Mr. Ferdinand Adungo,
CIPDCR,
BUSIA,

Through:
The Director,
CIPDCR,
BUSIA

Dear Sir,

RE: SSC No. 1127
A genotyping study of human immunodeficiency virus type-1 in relation to drug resistance in western Kenya PI: FO Adungo (CIPDCR)


The Committee acknowledges receipt of the ICDs translated into Kiswahili and Luhya.

Due consideration has been given to ethical issues and the study is granted approval. You may proceed with the study.

You are responsible for reporting to the Ethical Review Committee any changes to the protocol or in the Informed Consent Document. This includes changes to research design or procedures that could introduce new or more than minimum risk to human subjects.

Yours faithfully,

[Signature]
R. C. Kithinji,
For: Secretary,
KEMRI/NATIONAL ETHICAL REVIEW COMMITTEE

February 21, 2007

In Search of Better Health
APPENDIX VII COPY OF RESULTS: GENERATED SEQUENCES

The generated sequences were deposited in the Genbank using accession numbers HQ176975-HQ177045.

Bus001>AGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAAGAAATGAATT TACCAGGAAATGGAACCAAAAAGAATGAGGAGGAATGGATCTTTATCAA GAATACGATATCAGAAATACCTGTAGAAATTGATGAGCAGATCAGGTAGTT CAGTATTAAAGGACCTACACTGGTGCTACATAATTGGAAGAAATTGTTGACCTCA GATTGTTGcACTTTAAAATTTCCCATTAGTCCTATTTGAAACCGTGCCGAA

Bus002>AGaCATAAATTTTGCAGGAAATGGAACCAAAAATGAGATAGGGGGAATT GAGGGTTTTCATCAAGGTAAGCAGTATGATCAGGTACTTTATAGAAATTTTGGA AAAAAAGGCTATAGGTACAGTATTAGTGGACCTACACTCCGTCAACATAATTTGA GAAGAAACTGTGTTGACCTAGTACATTTCATTTTTCAATATTGCTCTATTTG AAACCGTGCCGa

Bus003>GcAGCTAAGGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTAAGAATTTGCCAGGAAAATGGAAACCAAAAAGAATGAGGAGGAATTGAGGCTTTATCAAAGTAAGACAGTATGATCAGATACCTATAGAAATTTGTGGAAGAAATTGTTGACCTCA GATTGTTGcACTTTAAAATTTCCCATTAGTCCTATTTGAAACCGTGCCGAA

Bus004>TTTGCCAGGAAAGATGGAAACCAAGAATGATAGGGGGAATTGGAGGTTC ATCAGGTAAAACAGTATGACGATACCTATTATAGAAATTTTGGAAGAAAGGCT ATAGGTACAGTCTTTAGTGGACACCTACACTGGTGCAACATAATTGGAAGAAACATGT TGACCCCCGATTGGTGTACCTTTAAAACCTTTCAATATTGCTCTATTGAAACCGTGCCGACGAGAA

Bus005>ATTTACCAGGAAATGGAACCAAAAATGAGATAGGGGGAATTGGAGGTTC TATCAGGTAAAACAGTATGACGATACCTATTATAGAAATTTTGGAAGAAAGGCT ATAGGTACAGTCTTTAGTGGACACCTACACTGGTGCAACATAATTGGAAGAAATATGT TGACCTCAGATTGGTGTACCTTTAAAACCTTTCAATATTGCTCTATTGAAACCGTGCCGACGAGAA

Bus006>TCTATTAGATACAGGAGCAGATGATACAGTATTAGAAAGATAAAATTTG CCAGGAAATGGAACCAAAAATGAgTAgGGGAAATTGGAGGTTC TATCAGGTAAAACAGTATGACGATACCTATTATAGAAATTTTGGAAGAAAGGCT ATAGGTACAGTCTTTAGTGGACACCTACACTGGTGCAACATAATTGGAAGAAATATGT TGACCTCAGATTGGTGTACCTTTAAAACCTTTCAATATTGCTCTATTGAAACCGTGCCGACGAGAA

Bus008>ATTTGCCAGGAAATGGAACCAAAAATGAGATAGGGGGAATTGGAGGTTC TATCAGGTAAAACAGTATGACGATACCTATTATAGAAATTTTGGAAGAAAGGCT ATAGGTACAGTCTTTAGTGGACACCTACACTGGTGCAACATAATTGGAAGAAATATGT TGACCTCAGATTGGTGTACCTTTAAAACCTTTCAATATTGCTCTATTGAAACCGTGCCGACGAGAA

Bus009>AGAGCTCTATTAGATACAGGAGCAGATGATACAGcACAGAAAGATAAA TTTGCCAGGAAATGGAACCAAAAATGAgTAgGGGAAATTGGAGGTTC TATCAGGTAAAACAGTATGACGATACCTATTATAGAAATTTTGGAAGAAAGGCT ATAGGTACAGTCTTTAGTGGACACCTACACTGGTGCAACATAATTGGAAGAAATATGT TGACCTCAGATTGGTGTACCTTTAAAACCTTTCAATATTGCTCTATTGAAACCGTGCCGACGAGAA
CAGTATTAGTGGACCTACACCTGTCACAAATAATTTGGAAGAACATGTTGACCCA
GATTGGTTGTACTcTAAATTCCCCAACCTATTAGTCTTATTGAAACCGTGCCAGA

Bus019> CTATTAGATACAGGAGCAGATGATACATCAaATGATAGGAGCAGATGATACATC
GAGGAAATAATGGAAACCAaATAATGGGAAATTGGGAAATTGGGAGGTTTCTACAA
AATGAAAGCAATTATGATACAGAATTHTAgAAAAATTGTGGGAAAAGGGCATATAGGCA
CAGTATTAGTGACCTACACCTGTCACAAATAATTTGGAAGAAACATGTTGACCCCA
GATTGGTTGTACTTcGAATTTCCCCAACCTATTAGTCTTATTGAAACCGTGCCAGA

Bus020> AGCTCTATTAGATACAGGAGCAGATGATACATCAaATGATAGGAGCAGATGATACATC
GAGGAAATAATGGAAACCAaATAATGGGAAATTGGGAAATTGGGAGGTTTCTACAA
AATGAAAGCAATTATGATACAGAATTHTAgAAAAATTGTGGGAAAAGGGCATATAGGCA
CAGTATTAGTGACCTACACCTGTCACAAATAATTTGGAAGAAACATGTTGACCCCA
GATTGGTTGTACTTcGAATTTCCCCAACCTATTAGTCTTATTGAAACCGTGCCAGA

Bus021> CTCTATATTAGATACAGGAGCAGATGATACATCAaATGATAGGAGCAGATGATACATC
GAGGAAATAATGGAAACCAaATAATGGGAAATTGGGAAATTGGGAGGTTTCTACAA
AATGAAAGCAATTATGATACAGAATTHTAgAAAAATTGTGGGAAAAGGGCATATAGGCA
CAGTATTAGTGACCTACACCTGTCACAAATAATTTGGAAGAAACATGTTGACCCCA
GATTGGTTGTACTTcGAATTTCCCCAACCTATTAGTCTTATTGAAACCGTGCCAGA

Bus022> CTCTATTAGATACAGGAGCAGATGATACATCAaATGATAGGAGCAGATGATACATC
GAGGAAATAATGGAAACCAaATAATGGGAAATTGGGAAATTGGGAGGTTTCTACAA
AATGAAAGCAATTATGATACAGAATTHTAgAAAAATTGTGGGAAAAGGGCATATAGGCA
CAGTATTAGTGACCTACACCTGTCACAAATAATTTGGAAGAAACATGTTGACCCCA
GATTGGTTGTACTTcGAATTTCCCCAACCTATTAGTCTTATTGAAACCGTGCCAGA

Bus023> ACATaAATTTTGCCAGGAAATATGGGAAACCAaATGGGAAATTGGGAAATTGGGAGGTTTCTACAA
AATGAAAGCAATTATGATACAGAATTHTAgAAAAATTGTGGGAAAAGGGCATATAGGCA
CAGTATTAGTGACCTACACCTGTCACAAATAATTTGGAAGAAACATGTTGACCCCA
GATTGGTTGTACTTcGAATTTCCCCAACCTATTAGTCTTATTGAAACCGTGCCAGA

Bus024> AGATTTTGCCAGGAAATATGGGAAACCAaATGGGAAATTGGGAAATTGGGAGGTTTCTACAA
AATGAAAGCAATTATGATACAGAATTHTAgAAAAATTGTGGGAAAAGGGCATATAGGCA
CAGTATTAGTGACCTACACCTGTCACAAATAATTTGGAAGAAACATGTTGACCCCA
GATTGGTTGTACTTcGAATTTCCCCAACCTATTAGTCTTATTGAAACCGTGCCAGA

Bus025> TGGAAGGGTTTCTACAAAGTAACAGTGATAGCATAGCTcTAGTAGATGAGATTT
GTTGGAAATAAGaATaTaATGATACAGTGTAGGGTAGACCTACACCTGTCACAAATAATTTG
GAAATGAATaTGTTGACCCAGATTGGGTGTACTTTAAATTCCCCAACCTATTAGTCTATT
G

Bus026> CATaAATTTTGCCAGGAAATATGGGAAACCAaATGGGAAATTGGGAAATTGGGAGGTTTCTACAA
AATGAAAGCAATTATGATACAGAATTHTAgAAAAATTGTGGGAAAAGGGCATATAGGCA
CAGTATTAGTGACCTACACCTGTCACAAATAATTTGGAAGAAACATGTTGACCCCA
GATTGGTTGTACTTcGAATTTCCCCAACCTATTAGTCTTATTGAAACCGTGCCAGA
AGTATTAGTAGGACCTACACCTATCAACATAATTGGGAGAAATATGTTGACTCAGATTGGTTGACTTTAAATTTTCACACACCTAATTGGGAAACCGTGCCAG

Bus036>CATaaATTTGGCCAGGAAATGGGAAACCACAAAtTAATAGGGGAATTTGAGGTTCATTAAGGTAACACCTATCAATTAATTGGGAAATTTGACCTACACCTATCAACATAATTGGGAGAAATATGTTGACTCAGATTGGTTGTACTTTAAATTTTCCAATTAGTCCTATTGGAAACCGTGCCAG

Bus037>AGATTTGCCAGGAAATGGGAAACCACAAATGATAGGGGAATTGGAGGGTTTATCAAATGAAACAGTATGATCAGATACTTATAGAAATTTGGAAAAGGCTATAGGATCGTAGTCGAACTACACCTGTCACACATAATTTGGAAGAAATATGTTTAGCGATTGCTTTTGACTTTAAATTTTCCAATTAGTCCTATTGGAAACCGTGCCAG

Bus038>TAAAGAAGCTCTATTAGATACAGGAGCAGATGATACACAGTATTAGAAGAAATAAATTTGCCAGGAAATGGGAAACCACAAATGATAGGGGAATTGGAGGGTTTATCAAATGAAACAGTATGATCAGATACTTATAGAAATTTGGAAAAGGCTATAGGATCGTAGTCGAACTACACCTGTCACACATAATTTGGAAGAAATATGTTTAGCGATTGCTTTTGACTTTAAATTTTCCAATTAGTCCTATTGGAAACCGTGCCAG

Bus040>TTTGCCAGGAAATGGGAAACCACAAATGATAGGGGAATTGGAGGGTTTATCAAATGAAACAGTATGATCAGATACTTATAGAAATTTGGAAAAGGCTATAGGATCGTAGTCGAACTACACCTGTCACACATAATTTGGAAGAAATATGTTTAGCGATTGCTTTTGACTTTAAATTTTCCAATTAGTCCTATTGGAAACCGTGCCAG

Bus041>AGATTTGCCAGGAAATGGGAAACCACAAATGATAGGGGAATTGGAGGGTTTATCAAATGAAACAGTATGATCAGATACTTATAGAAATTTGGAAAAGGCTATAGGATCGTAGTCGAACTACACCTGTCACACATAATTTGGAAGAAATATGTTTAGCGATTGCTTTTGACTTTAAATTTTCCAATTAGTCCTATTGGAAACCGTGCCAG

Bus042>CTCTATTAGATACAGGAGCAGATGATACACAGTATTAGAAGAAATAAATTTGCCAGGAAATGGGAAACCACAAATGATAGGGGAATTGGAGGGTTTATCAAATGAAACAGTATGATCAGATACTTATAGAAATTTGGAAAAGGCTATAGGATCGTAGTCGAACTACACCTGTCACACATAATTTGGAAGAAATATGTTTAGCGATTGCTTTTGACTTTAAATTTTCCAATTAGTCCTATTGGAAACCGTGCCAG

Bus043>GCTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACACAGTATTAGAAGAAATAAATTTGCCAGGAAATGGGAAACCACAAATGATAGGGGAATTGGAGGGTTTATCAAATGAAACAGTATGATCAGATACTTATAGAAATTTGGAAAAGGCTATAGGATCGTAGTCGAACTACACCTGTCACACATAATTTGGAAGAAATATGTTTAGCGATTGCTTTTGACTTTAAATTTTCCAATTAGTCCTATTGGAAACCGTGCCAG

Bus044>TCTATGATACAGGAGCAGATGATACACAGTATTAGAAGAAATAAATTTGCCAGGAAATGGGAAACCACAAATGATAGGGGAATTGGAGGGTTTATCAAATGAAACAGTATGATCAGATACTTATAGAAATTTGGAAAAGGCTATAGGATCGTAGTCGAACTACACCTGTCACACATAATTTGGAAGAAATATGTTTAGCGATTGCTTTTGACTTTAAATTTTCCAATTAGTCCTATTGGAAACCGTGCCAG
xciv