The Epidemiology of Dual HIV infection in the KwaZulu-Natal Anti-Retroviral Roll-out Programme

by

Anneta Frances Naidoo

Submitted in fulfillment of the requirements of the degree of
Master in Medical Science
in the

Department of Virology
Faculty of Health Sciences
University of KwaZulu-Natal
Durban

August 2007

Declaration

This study presents original work by the author and has not been submitted to this or any other university. The research described in this thesis was carried out in the Department of Virology, Faculty of Health Sciences, University of KwaZulu-Natal under the supervision of Dr. R. Parboosing. Where use is made of the work of others, it has been duly acknowledged in the text.

Naidoo

Anneta Naidoo

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Medical Research Council South Africa, National Research Foundation (NRF) and the National Health Laboratory Services (NHLS) for funding this project

My supervisor, Dr Raveen Parboosing for his undivided support, commitment and constant guidance throughout this time

Michelle Gordon, a 'godsend.'

Every staff member at the Department of Virology, IALCH their assistance, time and encouragement is invaluable (Special thanks to Pinky, Sharmayla, Logan, Koshla and Dr. Pravi Moodley)

Sunny, from Chemical Pathology and Karen Bishop for assistance with HMA

Desa and Kumari, from the Department of Neuroscience for the use of facilities in their department

Students and staff at the Molecular biology unit, DDMRI for the use of their lab for cloning

Africa Centre for the use of sequencing facilities.

South African Sugar Research Association for performing sequencing.

Friends and family especially my Mum, Dad and Elishia for love and support.

Yash, for his love

ABBREVIATIONS

µl	microliters
^{0}C	degrees celcius
3'	three prime
3TC	lamivudine
A	adenine
5'	five prime
A98S	adenine at position 98 in the HIV genome changes to serine
ABI	Applied Biosystems Incorporated
AIDS	Acquired Immune Deficiency Syndrome
APS	Ammonium Persulphate
ARV	Anti-Retroviral
ATP	Adenosine Triphosphate
AZT	Zidovudine
bDNA	branched Deoxyribonucleic acid
Bld WS	Blood Wash
bp	base pairs
C	Cytosine
CCR5	CC-chemokine receptor 5
CD4	cluster of differentiation antigen 4
CD45-FITC	cluster of differentiation antigen 45 fluorescein isothyocyanate
CD4-PE	cluster of differentiation antigen 4 Phycoerythrin
cDNA	complimentary Deoxyribonucleic acid
cm	centimetres
CMV	cytomegalovirus
CRF	circulating recombinant forms
CTK	Christ the King
D30N	aspartic acid at position 30 changes to asparagine
d4t	stavudine
ddC	dideoxycytidine
ddI	didanosine
dNTP	deoxynucleotide triphosphate
ddNTP	dideoxynucleotide triphosphate
DHHS	Department of Health and Human Services
DNA	Deoxyribose Nucleic acid
DOH	Department of Health
dsDNA	double stranded Deoxyribose Nucleic acid
E.coli.	Escherichia coli
EDTA	ethylenediaminetetraacetic acid
EIA	enzyme immunoassay
ELISA	Enzyme-Linked ImmunoSorbent Assay
env	envelope
G	guanine
g	grams
gag	Group Specific Antigen

GBV	hepatitis G virus
gp120	glycoprotein 120
gp41	glycoprotein 41
Group M	Main
Group N	Non-M
Group O	outlier
HAART	highly active antiretroviral therapy
HCV	hepatitis C virus
HIV	human immunodeficiency virus
HIV-1	human immunodeficiency virus -1
HIV-2	human immunodeficiency virus - 2
HMA	heteroduplex mobility assay
HTLV III	Human T-cell Lymphotrophic Virus III
IbNg	Ibadan Nigeria
IC50	Inhibitory concentration at 50%
IC90	Inhibitory concentration at 90%
IFA	Immuno-fluorescence assay
IVDU	intravenous drug users
kph	kilometers per hour
KEH	King Edward Hospital
KZN	KwaZulu-Natal
L	litres
L90M	leucine at position 90 changes to methionine
lacZ	structural gene of Lac operon encoding *ß-galactosidase*
LAV	lymphadenopathy-associated virus
LB	Luria Bertani
LTRs	long terminal repeats
M	molar
$MgCl_2$	magnesium chloride
MGMH	Mahatma Ghandi Memorial Hospital
MHA	multi-region hybridization assay
ml	millilitres
N	Normal
NaCl	sodium chloride
NASBA	Nucleic Acid Sequence Based Amplification
NF-κB	Nuclear Factor Kappa Beta
NNRTI	non-nucleoside reverse transcriptase inhibitors
NRTI	nucleoside reverse transcriptase inhibitors
NVP	nevirapine
p24	protein 24
PB	primer binding
PBS	primer binding site
PCR	polymerase chain reaction
PEIA	peptide enzyme immunoassay
PENTA 5	Paediatric European Network for Treatment of AIDS
pH	potential of Hydrogen

PI	protease inhibitors
PLG CD4	PanLeucogated cluster of differentiation antigen 4
PMTCT	prevention of mother to child transmission
pol	polymerase
PPT	polypurine tract
PR	protease
Psi	Pounds per square inch
Rev	Responsive element
RFLP	restriction fragment length polymorphisms
RNA	ribonucleic acid
RNAse H	Ribonuclease H
rpm	rotations per minute
RT	reverse transcriptase
RT-PCR	reverse transcriptase – polymerase chain reaction
SA	South Africa
SIV	simian immunodefieciency virus
SIV_{cpz}	Simian Immunodeficiency Virus found in chimpanzees
SIV_{sm}	Simian Immunodeficiency Virus found in sooty mangabeys
SSCP	single stranded conformational polymorphisms
ssDNA	single stranded deoxyribonucleic acid
T	thymine
Taq	*Thermus aquaticus*
Tat	Trans-Activator of Transcription
TBE	Tris Borate ethylenediaminetetraacetic acid
T-cell	thymus-derived cell
TEMED	tetramethylethylenediamine
TIBO	tetrahydroimidazobenzodiazepine
$tRNA^{Lys3}$	transfer ribonucleic acid lysine 3
U3	unique three
U5	unique three
UK	United Kingdom
URFs	unique recombinant forms
UV	ultraviolet
V106M	valine at position 106 changes to methionine
V3 loop	variable three loop
V3-V5	variable three to variable five loop region
Vpr	Viral Protein R
Vpu	Viral Protein U
WBC	white blood cells
WHO	World Health Organisation
Y181C	tyrosine at position 181 changes to cysteine
Y181I	tyrosine at position 181 changes to isoleucine

LIST OF TABLES

LIST OF FIGURES

ETHICS

This study was given full ethics approval by University of KwaZulu-Natal, Biomedical Research Ethics Committee, Ref.: H026/05.

ABSTRACT

KwaZulu-Natal has the highest prevalence of HIV in South Africa. The prevalence of dual infection in a normal-risk population in this region is unknown. Dual HIV infection has important implications for diagnosis, treatment response and vaccine development. This cross-sectional study aimed to establish and optimize methods for subtyping and detection of dual infection in KZN. Samples were from chronically-infected patients on ARV treatment within the ARV Rollout Programme, from sites throughout KZN. Subtyping of the samples was performed using HMA. Four samples had indeterminate results by HMA and were then cloned and sequenced. Phylogenetic analysis showed that one of the four samples was a dual infection. This study showed 1/46(2%) samples to be dually infected which suggests that the prevalence of dual infection is low in the sample population. The low prevalence of dual infection reported could be due to the low-risk profile of the sample population. It was concluded that the low prevalence of dual infection is unlikely to have a considerable impact on HIV management.

CHAPTER ONE – Introduction and Background

1.1 Introduction

More than 38.6 million people worldwide are infected with HIV, of which 24.5 million live in Sub-Saharan Africa (UNAIDS, 2006). South Africa has one of the fastest growing HIV-1 epidemics, with an estimated 5.5 million people infected (UNAIDS, 2006). The prevalence of HIV-1 varies from region to region, with the highest prevalence of 39.1% found in KwaZulu-Natal (DOH, 2004). The Anti-Retroviral Rollout Programme in South Africa is a response to this pandemic. Highly Active Anti-Retroviral treatment (HAART) in the form of triple therapy is the basis of the ARV programme.

HIV has extraordinary genetic variability. The recombinogenic activity (Pulsinelli and Temin, 1994) and the lack of a proof-reading mechanism by the viral enzyme, reverse transcriptase (Battula and Loeb, 1976) introduce mutations into the HIV genome. Dual infection i.e. infection by two subtypes of HIV-1, may lead to recombinogenic events between subtypes, and the emergence of Circulating Recombinant Forms (CRFs). This contributes to the generation of HIV diversity.

Gottlieb *et al* (2004) found 5 cases of dual infection in a cohort of 64 patients. A study using South African patients reported 6 cases of dual infection in a cohort of 31 patients (Grobler *et al*, 2004). However, there are no published data on the prevalence of dual infection in the cohort of patients being targeted for the ARV rollout programme.

The aim of this study is to establish and optimize subtyping methods and to determine the molecular epidemiology of HIV and the prevalence of dual infection in the KwaZulu-Natal ARV Rollout programme.

1.2 Discovery of HIV-1 and HIV-2

HIV-1 and HIV-2 are members of the genus *Lentivirinae* of the family *Retroviridae* (Cavallo, 1986). Retroviruses are well known as causative agents of cancers such as leukaemias and lymphomas (Bierwolf, 1983). HIV-1, initially known as LAV or HTLV III, was isolated by Robert Gallo and Luc Montagnier, respectively, in1983 (Gallo *et al*, 1983; Barre-Sinoussi *et al*, 1983). In 1984, HIV-1 was identified as the causative agent of Acquired Immune Deficiency Syndrome (AIDS). Two years later, another AIDS-causing virus, HIV-2 was discovered (Clavel *et al*, 1986).

1.2.1 Strategies for HIV genetic diversity

HIV is one of the fastest evolving organisms due to its rapid replication rate, high mutation rate, frequent recombination and natural selection (Rambaut *et al*, 2004). HIV has a viral generation time of 2.5 days with a production of ~ 10^{10} – 10^{12} new virions each day (Perelson *et al*, 1996). The high mutation rate is due to the unique properties of the RT enzyme which is largely responsible for the sequence diversity of the retrovirus family of viruses (Roberts *et al*, 1988). The characteristic features of this enzyme are its template switching ability during reverse transcription (Pulsinelli and Temin, 1994) and the lack of an exonucleolytic proof-reading mechanism (Battula and Loeb, 1976).

1.2.1.1 Proofreading mechanism

Eukaryotic cells contain the highly efficient enzyme, DNA polymerase. One of its functions is to ensure correct base pairing during DNA synthesis. Due to tautomerisation, bases may be incorrectly paired i.e. a pyrimidine and pyrimidine, purine and purine, or unmatched pyrimidine with purine. The *Klenow fragment* of DNA polymerase has the ability to undo this mismatching via its 3'-5' exonuclease activity. Removal of the mismatched nucleotide occurs via the hydrolysis of the phosphodiester linkage between non-base paired regions of the deoxyribonucleosides (Kornberg and Baker, 1992).

Unlike DNA polymerase, RT does not have a proofreading mechanism and is able to polymerise at mispaired termini (Roberts *et al*, 1988; Preston *et al*, 1988). Sequence analysis of mutations generated by RT showed that base substitutions, insertions and deletions are produced (Roberts *et al*, 1988; Preston *et al*, 1998). This lack of a proofreading mechanism allows the enzyme to produce mutations of $3x10^{-5}$ per cycle of replication (Mansky and Temin, 1995). This implies that each new genome differs from its parent by one mutation.

RT is not responsible for all the genetic errors in the HIV genome (Preston and Dougherty, 1996; Zhang, 2004). RNA polymerase II, a cellular enzyme required for the process of transcription during HIV replication contributes to the mis-incorporation of nucleotides into the genome since it also lacks a proofreading mechanism (Zhang, 2004). The lack of a proof-reading mechanism by cellular and viral enzymes is the basis for HIV diversity but the template switching ability of reverse transcriptase introduces further variability.

1.2.1.2 Template switching

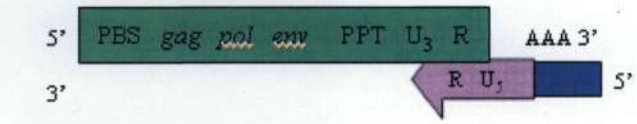
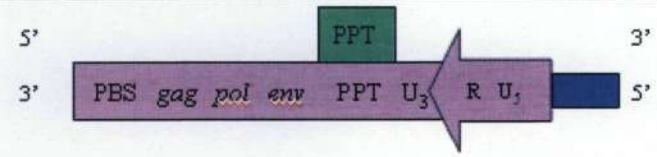
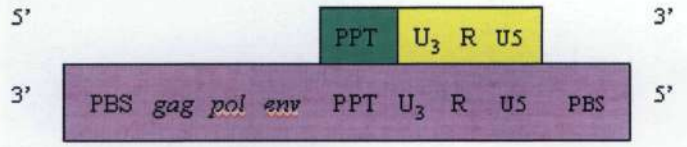
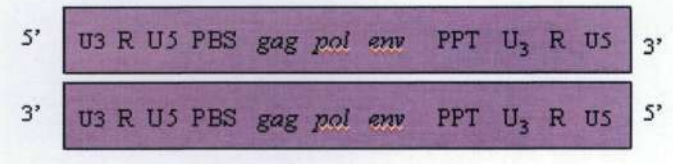
The HIV genome is known as diploid because it consists of two homologous copies of single stranded positive sense RNA. However, only one strand of RNA is converted into dsDNA and therefore the genome is sometimes referred to as pseudodiploid (Temin, 1993). Strand switching (template switching) occurs during reverse transcription of the RNA genome and results in a singular dsDNA per virion (Jetzt *et al*, 2000). Template switching occurs when reverse transcriptase switches between alternative genomic strands (Freed and Martin, 2001) and is largely responsible for the production of recombinant genetic forms (Hu and Temin, 1990).

1.2.1.2.1 Reverse transcription

cDNA synthesis is initiated with the binding of the cellular primer $tRNA^{Lys3}$ to the primer binding site (PBS) at the 5' end of the single stranded RNA molecule. This is followed by two obligatory template jumps that result in the production of long terminal repeats (LTRs) at both ends of the genome (Table 1). LTRs have a number of functions essential for HIV replication:

1) they form an intermolecular bridge of DNA-RNA hybrid linking the newly synthesised minus strand DNA with the 3' end of the same or a different genomic RNA template;
2) during integration, a step after reverse transcription of the replicative cycle, the LTR elements located at extreme ends of the linear viral DNA molecules mediate the insertion of the HIV genome into the chromosomal DNA of the host cells;
3) LTRs contribute to the packaging of progeny HIV-1 RNA genomes
(Freed and Martin, 2001).

Table 1: Reverse transcription of the HIV genome (Wagner and Hewlett, 2004)

1. $tRNA^{Lys3}$ binds to PBS on the positive RNA strand	
2. Synthesis of LTR occurs from 5' to 3' direction to copy the U5 and R region. This is the minus strand strong stop.	
3. RNAse H activity is mediated by RT on the RNA template, U5 and R regions.	
4. First strand switch: the minus strand strong stop plus the primer jumps to the 3' end of the RNA template	
5. DNA synthesis of the negative sense cDNA strand is completed. RNase H activity then removes the RNA template, except the Polypurine tract (PPT).	
6. Positive sense strand synthesis: PPT serves as the primer and the newly formed negative sense cDNA is the template. This proceeds to the PB site.	
7. Second strand switch: the partial cDNA then anneals to its own tail to complete positive strand synthesis. Binding occurs specifically at the DNA PB site.	

Reverse transcription of the viral RNA genome is a highly complex process with an efficient evolutionary strategy. It ensures the "survival of the species" by accelerating the spread of

beneficial mutations among viral quasispecies (Temin, 1991), increasing the genetic variation of HIV and mediating the repair of defective retroviral genomes (Boulerice *et al*, 1991).

1.2.1.2.2 Recombination

Packaging of RNA from different HIV-1 subtypes, intrasubtypes or intergroups into the same viral particle, in conjunction with the strand-switching ability of reverse transcriptase produces recombinant HIV in co-infected individuals (Hu and Temin, 1990; Goodrich and Duesberg, 1990; Stuhlmann and Berg, 1992). During HIV proviral DNA synthesis two necessary template 'jumps' occur for the production of LTRs at both ends of the genome (Temin, 1993).

The first jump, involves the translocation of the minus strong-stop DNA from the 5'end to the 3' end of the genome. The jump is made possible by the presence of repeat sequences at both ends of the genome. This can occur within the same strand (intramolecular) or between two different strands within the genome (intermolecular). The second jump occurs to ensure the transfer of the plus-strand ssDNA, initiated at the 3' polypurine tract, from the 3' to the 5' end of the genome. This jump is almost always intramolecular. The intermolecular jumps are responsible for recombination. (Goff, 2001)

Besides the obligatory jumps, strand switches may occur along internal genome sequences (Hu and Temin, 1990). Two models have been proposed for this activity that contributes to the development of recombinants: forced copy-choice (Coffin, 1979) and strand displacement-

assimilation (Junghans *et al*, 1982). Forced copy-choice occurs during minus-strand synthesis when RT encounters a break in the template RNA. The RT jumps to the alternative co-packaged RNA and resumes DNA synthesis (Coffin, 1979). However, this model has been refined with the suggestion that jumps during minus-strand synthesis can occur in the absence of strand breaks. This is suggestive, since the low processivity of RT would result in pauses along RNA-dependant DNA polymerisation, which would promote template switching without the need for breaks (Buiser *et al*, 1993).

The strand displacement-assimilation model suggests that the recombination process is initiated by the displacement of one fragment of internally initiated plus-strand DNA by another plus-strand transcript. The displaced fragment then binds to the minus-strand DNA derived from reverse transcription, which eventually produces a recombinant double-stranded DNA copy of the viral genome. (Goff, 2001)

These mechanisms of generating HIV diversity result in tremendous intra-patient variability of the virus, which is known as quasispecies. In addition, HIV at the population level is classified into various strains, groups, subtypes and CRFs, due to the nature of its origins and subsequent evolution.

1.3 Origin of HIV-1 and HIV-2 strains

HIV-1 and HIV-2 are related to the SIVs, the chimpanzee subspecies (Pan troglodytes troglodytes) and sooty mangabey monkeys (Cercocebus atys) respectively (Gao *et al*, 1999; Gao *et al*, 1994). The close phylogenetic relationship and geographic co-incidence of HIV-2 and SIV_{sm} support the hypothesis that HIV-2 originated from cross-species transmission of SIV_{sm} from sooty mangabeys into humans (Gao *et al*, 1994). By analogy to HIV-1 group M, HIV-2 isolates have been classified into subtypes, designated A-H. HIV-2 subtypes in terms of genetic distance are closer to the Groups of HIV-1 than to the subtypes of HIV-1 (Peeters *et al*, 2003). The innate genetic diversity of HIV-1 is displayed by three separate introductions of the virus from chimpanzees to humans. This has resulted in the three separate groups of HIV-1: M, N and O.

1.3.1 Classification of HIV

Since 1992, sequence analyses of the *env* and *gag* regions of the HIV genome have been used to classify various types of viral sequences in the AIDS pandemic (Peeters and Sharp, 2000). HIV has been divided into species (HIV-1 and HIV-2), groups, subtypes (clades), sub-subtypes (sub-clades), circulating recombinant forms (CRFs) and unique recombinant forms (URFs). Three HIV-1 groups (M, N and O) exist. Between Groups M and O the *env* protein may differ as much as 30-50% with Group N being phylogenetically equidistant from the two. The inter-subtype difference of the *env* protein is 20-30%, while intraclade variations of this protein are between 10-15% (Wainberg, 2004). The variation between the different HIV genetic forms arises as a result of the replication strategies of the virus.

1.3.1.1 Group M subtypes

By definition, viruses belonging to a subtype should resemble each other more closely than other subtypes across the entire genome (Peeters and Sharp, 2000). There are nine subtypes of HIV-1 Group M: A, B, C, D, F, G, H, J and K (Myers *et al*, 1992). Subtypes E and I were initially described as subtypes but were later determined to be recombinants (Gao *et al*, 1996; Gao *et al*, 1998). The average nucleotide percentage distance between subtypes A-D, F-H, J and K is 27% (range, 21-31%) (Moore *et al*, 1996). Certain subtypes can be subdivided into sub-subtypes as is the case for subtypes A and F. Within the Subtypes A and F, sub-clades are designated A1 and A2 and F1 and F2 respectively (Lal *et al*, 2005). The extent of divergence between F1 and F2 is similar to that seen between subtypes B and D. For this reason subtypes B and D should be considered as sub-clades of a single subtype (Peeters and Sharp, 2000). Due to historical reasons these designations cannot be altered.

1.3.1.2 Groups O and N subtypes

Group O isolates are highly divergent from Group M, demonstrating only about 50% amino acid sequence similarity to the *env* region of Group M (Vanden Haesevelde *et al*, 1994). Initially, analysis of data suggested that Group O sequences did not form clusters the same way as the Group M subtypes (Leitner, 1996). However, distinctive clades within Group O were distinguished in 1999 (Korber *et al*, 2001) but phylogenetic differentiation was not clear (Kandathil *et al*, 2005). Hence, subtypes within Group O are not yet defined.

Group N represents a small number of strains genetically distinct from Groups M and O. Infections are confined to Cameroon (Los Alamos, 2006) Not many strains have been sequenced therefore subtypes have not been determined. Phylogenetic analysis suggests that Group N is the result of a recombination event between a SIV_{cpz}- like virus and an HIV-1-like virus (Gao *et al*, 1999).

1.3.1.3 Recombinant HIV viruses

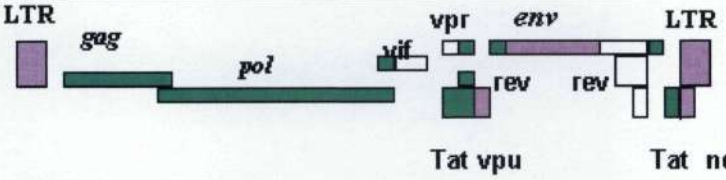
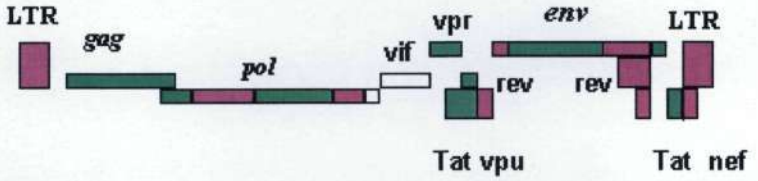
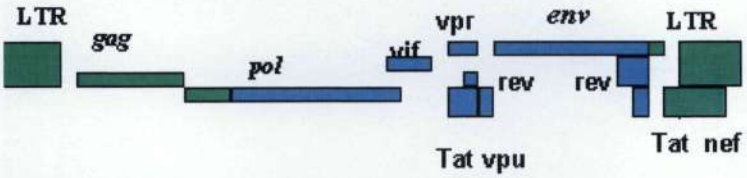
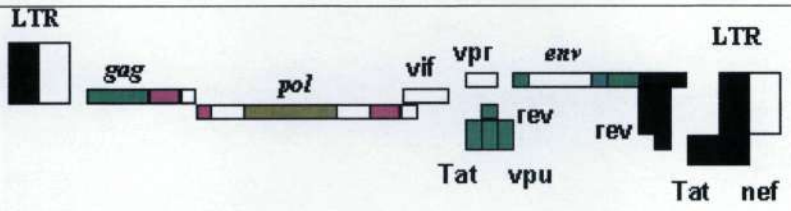
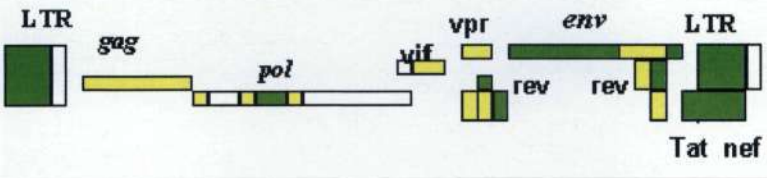
The recombinogenic rate of HIV is one of the highest of all organisms, with an estimated three recombinogenic events occurring per genome per replication cycle (Jetzt *et al*, 2000). This exceeds the mutation rate of ~0.2 errors per genome per replication cycle (Preston *et al*, 1988). These strategies strengthen the genetic diversity of HIV by contributing to the pool of genetic variants. The genetic forms arising after HIV replication must conform to the following definition to be considered a CRF: they should resemble each other over the entire genome, with similar breakpoints reflecting common ancestry from the same recombination event(s) (Robertson *et al*, 2000). The criteria for this definition are: three epidemiologically unlinked viruses with co-incidental mosaic structures and consistent phylogenetic clusters must be characterised in near full-length sequences (> 8 kb) (Robertson *et al*, 2000).

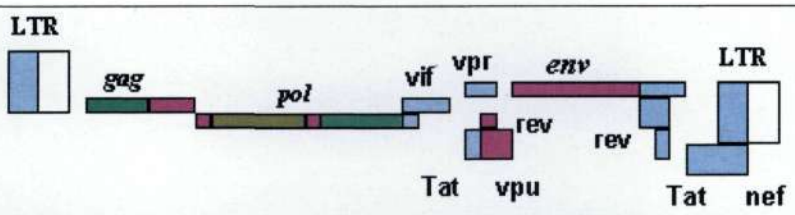
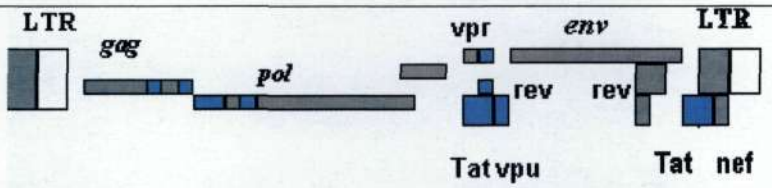
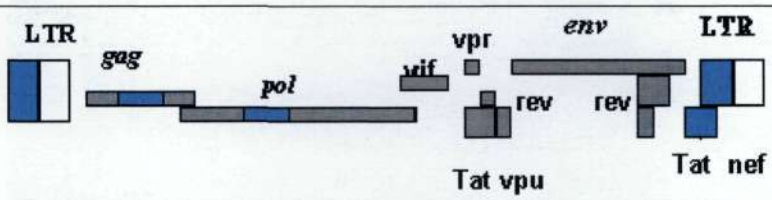
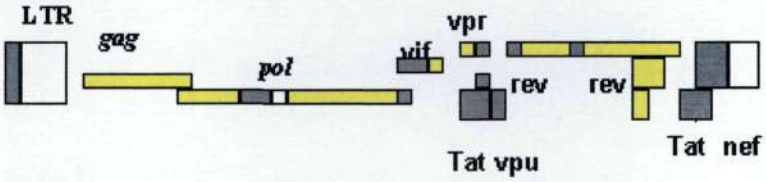
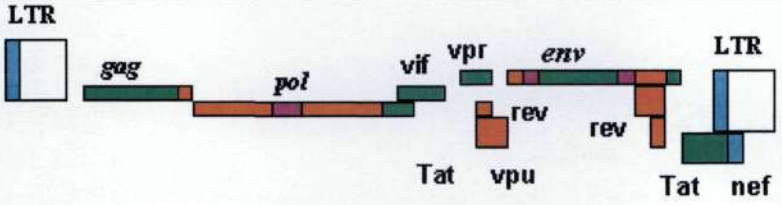
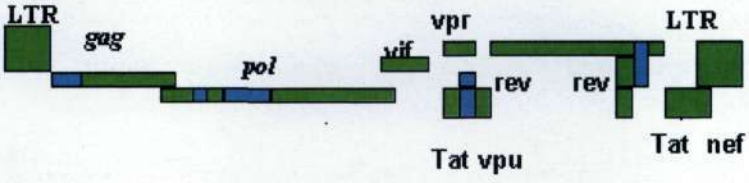
There at least 16 different CRFs identified thus far (http://www.hiv.lanl.gov/content/hiv-db/CRFs/ CRFs.html) (Table 2). Most of these CRFs have specific geographical locations where the relevant subtypes co-circulate (Leitner, 2005, Thomson *et al*, 2002c).

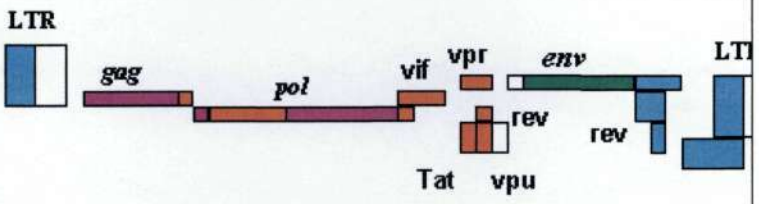
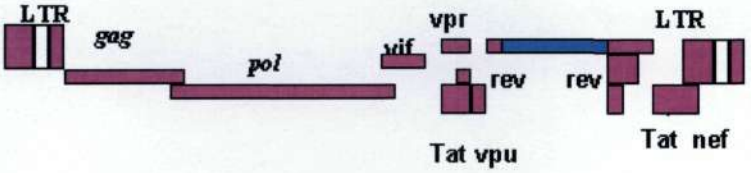
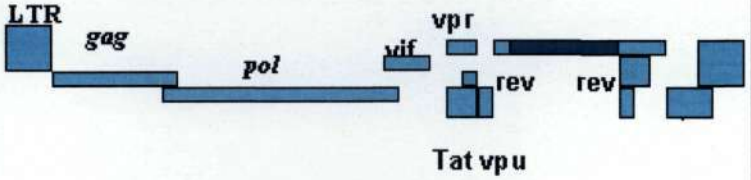
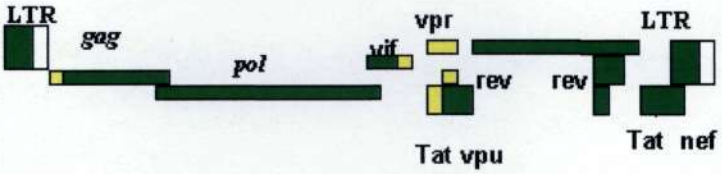
Table 2: Full genome sequences of CRFs

KEY:

A B C D E F G A2 B H K J CRF01_AE

CRF	Subtypes	STRUCTURE	GEOGRAPHICAL ORIGIN
CRF_01AE	A,E		Mekong region and the Malay Peninsula, Central Africa, USA
CRF_02AG	A,G		West and Central Africa
CRF_03AB	A, B		Russia and Ukraine
CRF_04 Cpx	A, G, H, K, unclassified		Cyprus and Greece
CRF_05 DF	D, F		Belgium, DR Congo

CRF	Subtypes	Structure	Geographical Region
CRF06_cpx	A, G, J and K		Burkina Faso, Mali
CRF07_BC	B, C		China
CRF08_BC	B, C		Southern China
CRF10_CD	C, D		Tanzania
CRF11_cpx	A, CRF01_AE, G, J		Cameroon, Gabon, DR Congo, Central African republic
CRF12_BF	B, F		Argentina, Uruguay

CRF	Subtypes	Structure	Geographical Region
CRF13_cpx	A, CRF01_AE, G, J, unclassified		Cameroon
CRF14_BG	B, G		Western Europe
CRF15_01B	CRF01_AE, B		Thailand
CRF16_A2, D	A2, D		Kenya, South Korea and Argentina

CRF09- has not been published yet. Adapted and modified from HIV Sequence Database.

1.3.1.4 Unique Recombinant forms (URFs)

Recombinant forms of HIV found in a single individual or in a single epidemiologically-linked cluster are known as URFs. These mosaic viruses like CRFs have discrete breakpoints between their genomes with differing phylogenetic affinities (Peeters and Sharp, 2000). URFs are thought to arise due to secondary recombination of a CRF (Thomson *et al*, 2002). URFs have been reported throughout the world (Vidal *et al*, 2000; Thomson *et al*, 2002a, Thomson *et al*, 2002c). High prevalences of URFs have been found in Kenya, Argentina, China and

Nigeria (Dowling *et al*, 2002, Carr *et al*, 2001, Yang *et al*, 2002, (McCutchan *et al*, 1999). The frequent finding of URFs reflects the occurrence of dual infections.

1.3.1.5 Dual infection

Dual HIV infection is the presence of more than one HIV genetic sequence i.e. Groups, subtypes, intra-subtypes and/or strain within an individual. It occurs either as a co-infection when two viral variants cause infection at or near the same time before seroconversion, or a superinfection in which primary infection is followed by infection with another viral variant after seroconversion.

Dual infections have been found in areas where multiple HIV variants circulate. For instance, co-infections of viruses belonging to distinct HIV-1 subtypes have been found in Tanzania, Kenya and Brazil (Herbinger *et al*, 2006; Steain *et al*, 2005, Ramos *et al*, 1999). Intra-subtype dual infections are also of major concern since subtype B intra-subtypes and subtype C intra-subtypes infections have been reported (Gottlieb *et al*, 2004). A patient was reported to have a multiple infection, where the original infection was with subtype B followed by superinfection with CRF01_AE and subsequently a second subtype B infection (van der Kuyl *et al*, 2005).

The possibility of dual infection is of growing concern since it is a pre-requisite for genetic recombination and the development of CRFs. Of greater importance is the possibility that dual infections could result in strains of HIV strains that are more virulent, have single or multiple drug resistant mutations or have altered tropism.

1.4 Geographical distribution of HIV genetic variants

HIV-1 subtypes are unevenly distributed throughout the world (Korber *et al*, 1997) (Figure 1). Subtype B is synonymous with North America and Europe. Although Group M subtypes and Group O viruses have been reported in the United States, less than 1 per cent of infections are due to these subtypes (McCutchan, 2000; Brodine *et al*, 1995; Rayfield *et al*, 1996). In South America, Subtypes F and C are found but subtype B predominates. B/F recombinant variants have been reported in Brazil and Argentina (Marquina *et al*, 1996; Janini *et al*, 1998).

Subtype B is most frequent in Europe but other subtypes are increasingly documented. For example, subtypes A, C, D, F, G, H, CRF01_AE and Group O infections have all been reported (McCutchan, 2000). The epidemic in Asia consists of all types of HIV-1. In India, subtype C predominates but subtypes A and B co-circulate with A/C recombinants frequently being reported (Jameel *et al*, 1995; Lole *et al*, 1999). In Thailand, CRF01_AE is transmitted heterosexually and established an explosive epidemic in the 90s (McCutchan *et al*, 1995; Weniger *et al*, 1994; Weniger and Brown, 1996a). HIV transmission in Thailand and certain areas of Bangkok is via intravenous drug use (IVDU) (Chen *et al*, 1998). Initially this resulted in the spread of subtype B viruses but data suggests that CRF01_AE is increasing in frequency through IVDU (Peeters and Sharp, 2000)

The greatest genetic diversity of HIV-1 has been found in Africa, which is consistent with it being at the epicentre of the epidemic. HIV-2, which does not show any global geographical distribution, is restricted to West Africa (Kandathil *et al*, 2005).

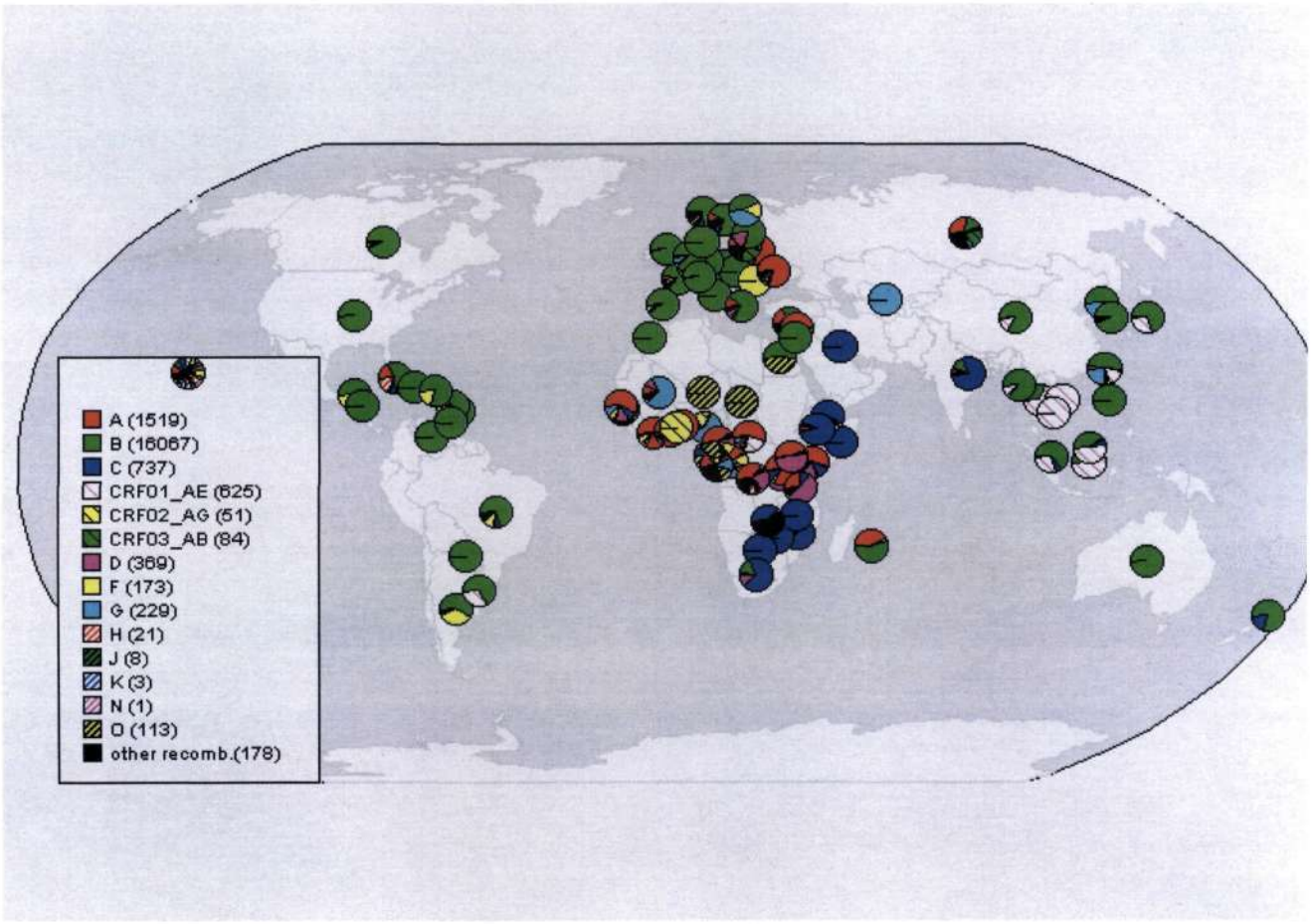


Figure 1: Geographical representation of HIV-1 subtype distribution throughout the world. Data downloaded from HIV Sequence Database, Los Alamos, NM

Group M subtypes A and C cause the most frequent infections followed by A/G recombinant variants (Wainberg, 2004). Subtype C predominates in South and East Africa (Peeters *et al*, 2003). Globally, subtype C is the greatest contributor to the epidemic, causing 50 per cent of

infections (Hemelaar *et al*, 2006). Subtype A and A/G recombinant variants are localized in West and Central Africa (Piot and Bartos, 2002). CRF02_ AG, a complex mosaic of subtypes A and G was first identified in Ibadan, Nigeria and is alternately known as IbNg (Howard and Rasheed, 1996; Carr *et al*, 1998). Five to forty per cent of infections in East and Central Africa are caused by subtype D (Peeters *et al*, 1998, Rayfield *et al*, 1998). Subtype B and other recombinant strains are rare (McCutchan, 2000) while other subtypes like H are exclusive to central Africa (Quinones-Mateu and Arts, 1999; Bikandou *et al*, 2000; Triques *et al*, 1999). The diversity of HIV is striking in Africa as new subtypes continue to be discovered. Subtype K, the most recent HIV-1 subtype to be discovered was identified in the Democratic Republic of Congo and Cameroon (Triques *et al*, 2000).

Group O is most common in West Central Africa with Group N only being reported in this region as well (Tatt *et al*, 2001; Simon *et al*, 1998, Salminen *et al*, 1997).

HIV-2 consists of groups A-H. Groups A and B are most prevalent (Kandathil *et al*, 2005). Group A infections specifically in West Africa have been reported more than Group B. Clade B infections have been limited to France, Ivory Coast, Portugal and Abidjan (Damond *et al*, 2001; Soriano *et al*, 2000; Brennan *et al*, 1997, Pieniazek *et al*, 1999). Groups C and D were identified in Liberia and Groups E and F were reported in Sierra Leone (Gao *et al*, 1994; Chen *et al*, 1997). Group G is represented by a strain found in the Ivory Coast (Yamaguchi *et al*, 2000) where Group H was also recently discovered (Damond *et al*, 2004).

1.4.1 Subtype distribution in South Africa

Two separate HIV epidemics caused the circulation of subtypes in South Africa. In the early 1980s the homosexual population presented with infections by subtypes B and D. This was followed by subtype C infections in the late 1980s in the heterosexual black population (Williamson *et al*, 1995). Continued surveillance reveals that subtype C has dominated the epidemic in South Africa (Los Alamos, 2006). This may be due to biological or socioeconomic factors. Subtypes B and D are still prevalent in the population but cause a very low percentage of infections (Treurnicht *et al*, 2002). Subtype A and a few recombinants have also been discovered (van Harmelen *et al*, 1999). (Figure 2)

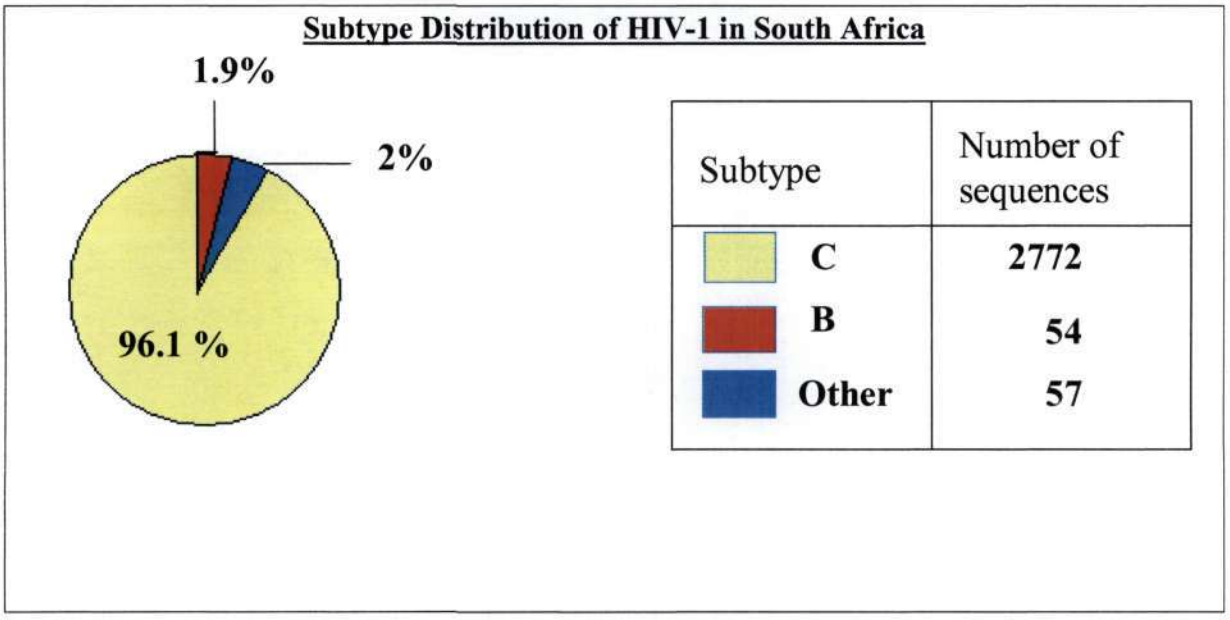


Figure 2: Distribution of subtypes in South Africa. Data downloaded from HIV Sequence Database, Los Alamos, NM

1.4.2 Prevalence of HIV in South Africa

It is estimated that approximately 5.2-6.7 million people are HIV positive in South Africa (UNAIDS, 2006). Nearly 6.4 per cent (or 6 million persons) adults are HIV-infected while infant infections account for approximately 0.3 per cent of the population. According to the National HIV and Syphilis Antenatal Seroprevalence Survey in South Africa, HIV/AIDS prevalence is on the increase. In 2005, the survey demonstrated an estimated HIV prevalence rate of 29.5% in comparison to the 2003 estimate of 27.9%. Even though the prevalence of HIV/AIDS in KZN for 2005 has declined, KZN still remains the province with the highest prevalence nationally.

1.5 Methods used to identify HIV genetic subtypes

HIV genomic variation requires extensive analysis because of its effect on disease progression, viral pathogenesis, vaccine development and drug resistance. Many subtyping techniques have been developed (de Oliveira *et al*, 2005; Hoelscher *et al*, 2002; Plantier *et al*, 2002) to examine these relationships. There are three main methods of subtyping: serotyping (the antibody response of the host using synthetic peptide corresponding to the V3 loop of gp120), heteroduplex mobility assay and sequencing.

1.5.1 Serological techniques

Serological subtyping is successful only in geographical locations where few subtypes are circulating (Chanbancherd *et al*, 1999, Kaleebu *et al*, 2000). In geographic locations where many different subtypes circulate e.g. West Africa, the results are unreliable due to cross-

reactivity between subtypes (Lal *et al*, 2005). This is because antigens currently used are based on a short region of the V3 crown domain of glycoprotein gp120 and differ by only a few amino acids amongst subtypes. For these reasons, serological assays are not the preferred method for subtyping HIV.

1.5.2 PCR-based techniques

PCR, which was invented in the 1980s, allows the characterisation and comparison of viruses at the genomic level (Kwok *et al*, 1987). It has broadened the scope of molecular virology by allowing the detection of viruses that were once difficult to detect (CMV) and amplified required regions of the genome as opposed to the entire genome (Arens, 1999). For this reason, subtyping techniques are primarily based on PCR.

1.5.2.1 Sequencing

Genomic analysis using sequencing followed by phylogenetic analysis is the most informative method of subtyping and is considered the 'gold standard.' Since 1985, when the complete genomic sequence of HIV was published, sequencing has been used to demonstrate the genetic variability of the virus (Ratner *et al*, 1985). Subtyping using sequencing has been streamlined to just the *gag* and *env* regions of the HIV genome. Recently Eshleman and colleagues (2004) used the *pol* region (RT and PR) of the HIV genome for subtype analysis. The entire genome is sequenced only in cases where highly divergent and recombinant variants exist (Gao *et al*, 1996). Although sequencing is considered tedious and expensive, it

is often used for comparative and confirmatory purposes with more recent subtyping techniques (Gottlieb *et al*, 2004; Saunders *et al*, 2005).

1.5.2.2 Heteroduplex Mobility Assay (HMA)

HMA is based on nested PCR products co-amplifying with divergent templates that randomly reanneal to form heteroduplexes that migrate with reduced mobility on a neutral polyacrylamide gel. Delwart *et al* (1993) described HMA as a simple and rapid quantitative assay for the detection and estimation of genetic divergence between HIV strains. The HIV-1 *env* region V3-V5 is a commonly used target for subtype determination using HMA (Delwart *et al*, 1994). A *gag* HMA was developed due to the *gag* gene containing less inter-subtype sequence variation than *env* (Tatt *et al*, 2000). It has been useful as an alternative to, and in conjunction with, *env* HMA in detecting recombinant strains (Bredell *et al*, 2000). However, using the *env* region of the genome in HMA is still a major cause of concern. Certain geographic locations demonstrate unamplifiable C2V3 regions, which has led to the development of an assay that targets the highly conserved gp41 region (Agwale *et al*, 2001). All variations of HMA have proven useful and successful for subtype determination.

1.5.2.3 Restriction Fragment Length Polymorphism (RFLP)

The sequence of bases in DNA can be cleaved at specific sites by restriction endonucleases. Restriction endonucleases are bacterial enzymes that cut a DNA sequence at palindromic sites. The resulting fragments are separated on an agarose gel and displayed as a banding pattern. The presence of mutations in the restriction sites results in different banding patterns of

fragments. This is the basis of using RFLP for subtyping HIV. If subtype B has more restriction sites for a specific restriction endonuclease than subtype A, then the resulting number of bands will be greater for subtype B than subtype A. Many studies have used RFLP as a subtyping method, either as a screening tool or as a confirmatory test. It has also been modified to identify specific intra-subtypes within a region (Gadelha *et al*, 2003).

1.5.2.4 Single Stranded Conformation Polymorphism (SSCP)

Conformational changes in double or single stranded DNA are used to detect mutations. These changes can be detected as alterations in the electrophoretic mobility of the nucleic acid in non-denaturing polyacrylamide gels. The separation in an electrical field is not only dependent on the shape of the molecules but also on the temperature, concentration and composition of the medium in which the molecules are moving. In SSCP the technique depends on the altered mobility of single stranded DNA molecules (Orita *et al*, 1989). Unlike double stranded DNA, single stranded DNA is flexible and will adopt a conformation determined by intramolecular interactions and base stacking that is uniquely dependent on sequence composition (Orita *et al*, 1989). SSCP is able to determine the genetic makeup of several hundred base pairs of DNA, whereas RFLP analysis can determine only a few bases (the restriction site).

1.5.2.5 Techniques used to identify dual infections

Several techniques are used to determine infection with more than one variant of HIV. Cloning followed by sequencing is the most reliable method (Steain *et al*, 2005; Brenner *et al*, 2004). However, this method is tedious and expensive so other methods, such as HMA and

RFLP, are used as screening techniques. Indeterminate HMA results occur when there are distinct slow migrating heteroduplexes or smears (Gottlieb *et al*, 2004; Grobler *et al*, 2004) or when two hetroduplexes migrate to the furthest equivalent position on a polyacrylamide gel. In these cases, HMA is confirmed by cloning and sequencing. Potential dual infections caused by distinct HIV-1 subtypes are identified by RFLP on the basis of the simultaneous presence of more than one endonuclease digestion pattern (Masciotra *et al*, 2000). A multi-region hybridisation assay (MHA) is a more recent technique which is based on the hybridisation of various primers to different regions of the HIV-1 genome (Herbinger *et al*, 2006; Hoelscher *et al*, 2002).

1.6 Impact of HIV Genetic Diversity

The genetic diversity of HIV has an impact on serological and molecular detection, treatment response and vaccine efficacy since all of these are dependent on the sequence of the target protein or HIV RNA/DNA.

1.6.1 Laboratory diagnosis

HIV diagnosis depends on the detection of one or more of the components of a viral particle or the immune response of the host to the virus (Gurtler, 1996). Constituents of a viral particle that make for easy detection include: Proteins p24 and RT, RNA or DNA. Detection of HIV-specific antibodies is the main diagnostic tool in adults. In infants, detection of p24 antigen and HIV RNA or DNA is required to confirm infection.

1.6.1.1 Serological diagnosis

Serological diagnosis is based on detection of HIV antibodies and/or antigens. The detection of antibodies to HIV depends on their highly specific binding to antigens which may either be viral lysate, recombinant proteins or synthetic peptides. Methods of detecting antibodies include the enzyme immunoassay (EIA, also commonly referred to as the enzyme-linked immunosorbent assay or ELISA), Western blot (or immunoblot), the immunofluorescence assay (IFA), immunochromatographic or agglutination rapid tests and the detuned assay (Iweala, 2004). First generation tests are based on purified viral lysate or viral antigen preparations from cultured T cells infected with HIV (culture supernatant or infected cells) (Cooper *et al*, 1987; Tegtmeier, 1988). These tests produce a substantial rate of false-positive results due to the presence of residual cellular antigens which are incorporated into the viral particle during maturation (Cooper *et al*, 1987; Tegtmeier, 1988). Second and third generation assays are based on recombinant viral antigens or synthetic peptides of selected immunodominant antigenic epitopes (Gurtler, 1996). For example structural and non-structural proteins like gp41 transmembrane envelope protein and gp120 and p24 of HIV are used. Fourth generation assays simultaneously detects anti-HIV antibodies and the HIV p24 antigen (Ly *et al*, 2001; Thomson *et al*, 2002c).

HIV diversity has an enormous impact on diagnostics. HIV antigens used in serological tests should bind antibodies against all HIV subtypes. Initially diagnostic tests were developed for subtype B viruses and were used in developing countries where majority of infections are caused by non-B type viruses (Kandathil *et al*, 2005). Although most serological tests were able to diagnose chronic HIV-1 Group M infections, difficulties arose when recent infections

with non-B subtypes had to be diagnosed. This occurred because antigens used in the assay were based on subtype B virus only (Koblavi-Deme *et al*, 2001). Studies also showed that persons infected with highly divergent HIV-1 group O failed to be diagnosed (Hackett *et al*, 1997). This problem was resolved by addition of a Group O lysate/peptide in the standard antigen mixture. Current assays are able to detect Group M and O as well as HIV-2 (Oelrichs, 2004). Due to the constant diversification of HIV, continued surveillance of commercial assays are necessary to ensure that newly emerging variants do not go undetected.

1.6.1.2 Molecular analysis

Nucleic acid testing for HIV uses specific regions in the HIV genome as its target. HIV is a complex retrovirus i.e. its genome contains accessory genes eg. *vpu*, *vpr* etc along with the necessary *gag*, *pol* and *env* regions. Target selection of the genome depends on the purpose of testing. Nucleic acid testing in the form of DNA PCR is used for diagnosis in infants whose mothers are HIV positive. Quantification of nucleic acids, expressed as viral load, is used in initiation and monitoring of ARVs.

For purposes of nucleic acid testing, the most conserved region within the genome is often chosen (Kwok and Sninsky, 1993). A conserved region is a sequence of nucleotides that remains essentially unchanged during the evolution of the viral genome. This is particularly important in nucleic acid testing because HIV is known to demonstrate heterogeneity. Furthermore, the conserved region that is used must be common in all variants of the virus.

The nucleotide sequence of primers used in nucleic acid testing of HIV has been optimized to yield equivalent amplification of Group M subtypes and not of HIV-2 or other groups of HIV-1. Furthermore, the appearance of new variants may not be identified by these primers.

1.6.2 The ARV Rollout Programme

The national ARV Rollout Programme, initiated in 2004, is essential and challenging, particularly in KZN where the prevalence of HIV is the highest in the country. Accredited ARV sites are situated throughout KwaZulu-Natal. Attendees of these clinics have their blood samples taken which is sent to a reference laboratory for CD4 T-cell testing. CD4 T-cell counts are measured using flow cytometry. Patients with CD4 counts <200 cells/μl, or WHO stage VI irrespective of CD4 count receive intensive adherence counseling in preparation for HAART. Once the patient is ready for therapy, a baseline HIV viral load is performed using PCR or NASBA based technology

Patients are then initiated on one of two ARV regimens (table 3). CD4 T-cell counts and viral loads are performed 6-monthly in order to monitor therapy. Patients who fail Regimen 1a or 1b are assessed and investigated further. If treatment failure is confirmed, the patient is switched to Regimen 2 (Table 3). Patients who fail both regimens are referred to ARV specialists for individual evaluation which may include drug resistance testing (DOH, ARV 2005).

Table 3: Recommended ARV regimens

Regimen	Drugs
1a	d4T/ 3TC/ efavirenz
1b	d4T/ 3TC/ Nevirapine*
2	Zidovudine/ ddI/ Lopinavir/ Ritonavir (Kaletra™)

* Nevirapine is substituted for efavirenz when female patients cannot guarantee reliable contraception.

1.6.2.1 Treatment monitoring

Measurement of HIV RNA levels in plasma is crucial to the management of HIV infection. Current guidelines for HIV management recommend the use of HIV RNA quantification or viral load testing for the initiation and monitoring of Highly Active Anti-Retroviral Treatment (DHHS guidelines, 2004; DOH ARV guidelines, 2004). The four most utilised commercial viral load assays are the Amplicor HIV-1 Monitor v1·5 (Roche Molecular Diagnostics), nucleic-acid-sequence-based amplification (NASBA, Organon Teknika), HIV-1 Quantiplex branched DNA (bDNA, Chiron Diagnostics), and the LCx HIV-RNA quantitative assay (LCx, Abbott) (Lal *et al*, 2005). Regular evaluations are conducted to test the performance characteristics of these assays (Swanson *et al*, 2005).

Genetic variability of HIV can result in the under-quantification or lack of sensitivity of nucleic acid quantification assays. This is a result of the mismatching of primer and probe sites that reduces the efficiency of hybridization. All nucleic acid amplification assays and/or signal amplification rely on HIV-1 sequence specific primers and/or probes. Natural polymorphisms occurring in these target regions can result in reduced or abolished hybridization of these primers which causes under-quantification or inability to detect viral

RNA. The influence of genetic heterogeneity on performance is evident and constant surveillance of this relationship is required.

1.6.3 The impact of HIV variability on vaccine development

HIV vaccine development is one of the greatest challenges in HIV research. Despite intensive research, there is as yet no effective vaccine against HIV. The genetic diversity of HIV is a major obstacle to the development of a successful vaccine. Envelope glycoproteins, in particular, may not necessarily be immunologically cross-reactive due to the variability of the genetic sequences of HIV-1 and HIV-2. Subtypes may therefore have dramatic differences at the antigenic level. Since subtypes are clustered epidemiologically in distinct geographical locations, it has been suggested that different geographical regions of the world may require different vaccines. Vaccines were initially produced against subtype B virus, which is responsible for infections in North and South America, Europe and Australia but accounts for only 12% of infections worldwide (Peeters and Sharp, 2000). More attention has now been given to developing a subtype C vaccine due to the predominance of this subtype globally (Nkolola and Essex, 2006; Xin *et al*, 2007).

HIV variants in the form of CRFs and URFs continue to emerge, and create an ongoing challenge to vaccine development. An HIV vaccine should ideally produce a broadly cross-reactive immune response that protects against multiple HIV subtypes, CRFs, URFs and quasispecies. It may be possible to limit the effects of HIV diversity on vaccine efficacy by developing vaccines based on the conserved regions of the genome, where cross-clade reactivity has been confirmed.

Genetic diversity is continuously generated during the course of infection within an infected individual. Infected persons harbour a swarm of closely related but unique viruses (also known as quasispecies) which can differ by as much as 10% after 10 years of infection, and can be 30% different between individuals (Korber *et al*, 2001). Quasispecies of the virus exist that are a genetically heterogeneous population of virions produced as a result of the inaccuracy of viral replication. This also complicates vaccine development because it demonstrates that an antibody that targets one isolate in an individual may fail to neutralise another from the same individual.

Dual infections, particularly superinfection, demonstrate that the immune system cannot reliably prevent reinfection even after responding to an initial infection. It shows that infection with one variant cannot protect an individual against further infection (immunisation). Researchers first thought that dual infections occur only as co-infections. However, superinfections were soon discovered (Jost *et al*, 2002) which clearly showed that an 'immunisation effect' was not conferred by initial infections. For vaccine researchers the task is made more difficult by the finding that the immune system cannot reliably prevent reinfection even when responding vigorously to the initial infection (Allen and Atfield, 2003).

1.6.4 Impact of genetic diversity on antiretroviral drug sensitivity and drug resistance

The management of HIV disease with highly active antiretroviral therapy (HAART) using nucleoside and non- nucleoside reverse transcriptase inhibitors (NRTI and NNRTI), protease inhibitors (PI) and more recently fusion inhibitors (Table 4) makes HIV a more manageable

disease (Re *et al*, 2005). HAART decreases the viral load to undetectable levels thereby allowing immune reconstitution (DHHS Guidelines, 2004; DoH South Africa, 2004). Furthermore, it can decrease HIV transmission, and reduce morbidity and mortality but also creates the long-term spectre of drug resistance. Variations within the *pol* region of the genome are of particular relevance to drug susceptibility and the development of resistance since antiretroviral therapy has been carefully targeted at the conserved *pol* gene products, reverse transcriptase and protease.

Table 4: FDA approved Antiretrovirals used in the treatment of HIV Infection

CLASS	DRUGS	MECHANISM OF ACTION
Nucleoside Reverse Transcriptase Inhibitors (NRTIs)	Zidovudine, lamivudine, abacavir, zalcitabine, emtricitabine, dideoxycytidine (ddC), didanosine, tenofovir disoproxil fumarate, stavudine (d4t)	NRTIs in the triphosphorylated form closely resemble normal nucleotides. They compete with the normal substrates (nucleotides) for binding to RT which then incorporates them into the growing DNA. Once added NRTIs prevent additional nucleotides from being added. This results in DNA chain termination. NRTIs are competitive inhibitors.
Non-nucleoside Reverse Transcriptase Inhibitors (NNRTIs)	Delavirdine, efavirenz, nevirapine	Non-competitive inhibitors that bind to a region of the RT enzyme outside the active site. They prevent RT activity by allosterically distorting the orientation of the main catalytic aspartic acid residues in the active site of the enzyme.
Protease Inhibitors (PIs)	Amprenavir, tipranavir, indinavir, saquinavir, saquinavir mesylate, lopinavir, ritonavir, Fosamprenavir Calcium, atazanavir sulfate, nelfinavir mesylate	Prevent the *gag-pol* polyproteins from successfully entering the active site of the protease enzyme and being cleaved. This in turn prevents the assembly of new, fully functional HIV virions.

1.6.4.1 Mechanisms of drug resistance

Drug resistance is defined as the measurable loss of drug sensitivity due to a genetic change that alters the target of an antiviral agent. Within an HIV infected individual a diverse population of viruses exist that are constantly shaped by selective pressure like viral and host factors. Antiviral therapy is an additional source of selective pressure on the virus which

selects out those variants which are best able to survive drug pressure ('natural selection'). Very simply, the antiviral agent causes suppression of replication on most of the quasispecies, but some variants will possess mutations that enable replication to occur despite the presence of the drug. These mutations change enzyme structure and/or function in ways that reduce the effectiveness of the drugs.

NRTI resistance occurs by two mechanisms:

1) mutations arise that enable the RT enzyme to discriminate against NRTIs during polymerisation. This prevents the addition of NRTIs to the growing DNA chain relative to the natural dNTP substrates (Larder, 1999; Huang *et al*, 1998).

2) pyrophosphorolysis, nucleotide excision or nucleotide-dependent primer unblocking (Meyer *et al*, 1999). This mechanism is mediated by mutations that cause the hydrolytic removal of the chain terminating NRTIs from the growing chain. This mechanism requires a pyrophosphate donor such as ATP.

NNRTIs bind to the hydrophobic pocket in the RT. This pocket is adjacent to, but not contiguous with, the active site of the enzyme. High-level resistance to NNRTIs may arise from even a single amino acid change in the NNRTI-binding pocket. Some of the NNRTI resistance mutations may also compromise viral replication capacity due to changes in the conformation of the dNTP binding pocket (Kleim *et al*, 1994) or changes in RNaseH activity (Gerondelis *et al*, 1999; Archer *et al*, 2000).

Protease inhibitor resistance mutations prevent the inhibitor binding to the PR enzymes. This occurs by structural alteration of the substrate cleft of the enzymes where only natural substrates are allowed to bind. Mutations that alter the substrate cleft directly confer resistance to one or more protease inhibitor. Mutations in the flap region contribute to reduced inhibitor binding but alone do not provide considerable resistance.

The impact of the genetic diversity of HIV on drug resistance has been studied. Many studies have compared the response of therapy in B-type and non-B type infections. Even though antiretroviral agents were developed based on subtype B infections, ARVs are now used in non-B type infections. Nucleotide divergence within the *pol* region has emerged between HIV subtypes and intra-subtypes.

The Y181C and Y181I mutations, respectively, causes resistance of group O and all strains of HIV-2 to all drugs within the NNRTI class (Quinones-Mateu *et al*, 1998; Descamps *et al*, 1995). On a smaller scale subtype F shows resistance to a non-commercialised NNRTI compound, TIBO compound while remaining sensitive to other NNRTIs, as well as NRTIs and PIs (Apetrei *et al*, 1998). Mutations at NNRTI resistance associated positions such as V106M and A98S are substantially more common in subtype C infected patients than those with subtype B infections (Grossman *et al*, 2004a). A single dose Nevirapine PMTCT study in Uganda found resistance to NVP to occur more frequently in subtype D infected mothers than those infected with subtype A viruses (Eshleman *et al*, 2001). In Brazil, protease sequence analysis showed that the majority of mutations to be greater in B than non-B subtypes (Barreto

et al, 2006). Results from the PENTA 5 paediatric study have also shown rates of resistance to be greater in non-B than B subtypes (Pillay *et al*, 2002).

ARVs were designed for use against subtype B viruses, the predominating subtype in North America and Western Europe. The vast majority of data on the genetic mechanisms of HIV-1 drug resistance has also been generated by observations on subtype B virus. However, subtype B accounts for only about 12% of the global HIV pandemic (Hemelaar *et al*, 2006). Therapy has been introduced to developing countries which implies that the number of persons with non-B virus on therapy have increased dramatically. Studies suggest that currently available protease and RT inhibitors are as active against non-B subtype viruses as they are against B-subtype viruses (Easterbrook *et al*, 2003; Bocket *et al*, 2005; Frater *et al*, 2002). However, some studies suggest that their may be differences in the mechanisms of resistance in non-B viruses (Barreto *et al*, 2006; Zarandia *et al*, 2006; Brenner *et al*, 2003; Loemba *et al*, 2002).

Preliminary data indicates that drug resistance mutations that have been described for subtype B virus have been observed in at least one non-B subtype virus as well (Kantor *et al*, 2005). However, it is not known whether drug resistance mutations arising in non-B subtypes will go unrecognized in B viruses. Mutational patterns of the virus may sometimes differ in patients who are failing therapy depending on whether they are infected by non-B or B subtype virus. For instance, reports suggest that the NNRTI resistance mutation V106M occurs more commonly in subtype C than subtype B viruses (Grossman *et al*, 2004a). Protease inhibitor mutations D30N and L90M arising during nelfinavir treatment are subtype dependent.

Mutation D30N occurs more commonly in subtype B viruses, whereas L90M is selected more frequently in subtypes C, G, and CRF01_AE (Grossman *et al*, 2004b).

The evidence of the impact of genetic diversity on drug resistance is inconclusive. Some studies point in the direction of drug resistance mutations being more common in non-B subtypes, other studies contradict this statement. Further studies are required to establish a definite conclusion on the impact of HIV diversity on treatment response.

1.7 Dual infection

Dual infection or the infection with more than one variant of the virus affects disease progression and complicates the management of HIV infection.

Progression to AIDS is faster in those dually infected than those with only one strain of HIV-1 (Gottlieb *et al*, 2004; Grobler *et al*, 2004). These individuals may have higher viral loads and lower CD4 counts, which causes more rapid disease progression (Jost *et al*, 2002; Gottlieb *et al*, 2005; Grobler *et al*, 2004). It is not known if dual infection or infection with multiple divergent strains of HIV enables rapid adaptation and immune escape or if people who are genetically predisposed to experience rapid disease progression are unable to control viral diversity, enabling the establishment of dual infections (Grobler *et al*, 2004).

Dual infection also affects treatment response (Smith *et al*, 2005). An individual initially infected with a drug sensitive variant and superinfected with a drug resistant strain of HIV

(resistant to NRTIs and protease inhibitors) failed therapy (Smith *et al*, 2005). Superinfection with multi-drug resistant virus has also been reported (Brenner *et al*, 2004). Five other subtype B superinfections have been described, as well as three intersubtype A/E and B superinfections (Allen and Altfield, 2003; Altfield et al, 2002; Brenner et al, 2004; Jost et al, 2002; Koelsch et al, 2003; Ramos et al, 2002). In all these reports drug resistance is acquired by superinfection/reinfection with a drug resistant variant. It is still not clear how dual infection with wild-type viruses from different variants would affect the selection of drug resistant viruses.

The greatest fear in terms of dual infection and drug resistance is the possibility of generating a new infectious recombinant virus resistant to more than one antiretroviral agent (Figure 3). This could be made possible by the strand-switching ability of HIV (See 1.3.1.2.) If this recombinant infects a large percentage of the population, treating HIV with antiretrovirals may be very difficult. Even if drug-sensitive viruses are responsible for dual infection, the resulting recombinant may have a greater chance of developing drug resistance than the two viruses on their own. This could be due to the selection of mutations between the two viruses that enables the optimal viral fitness.

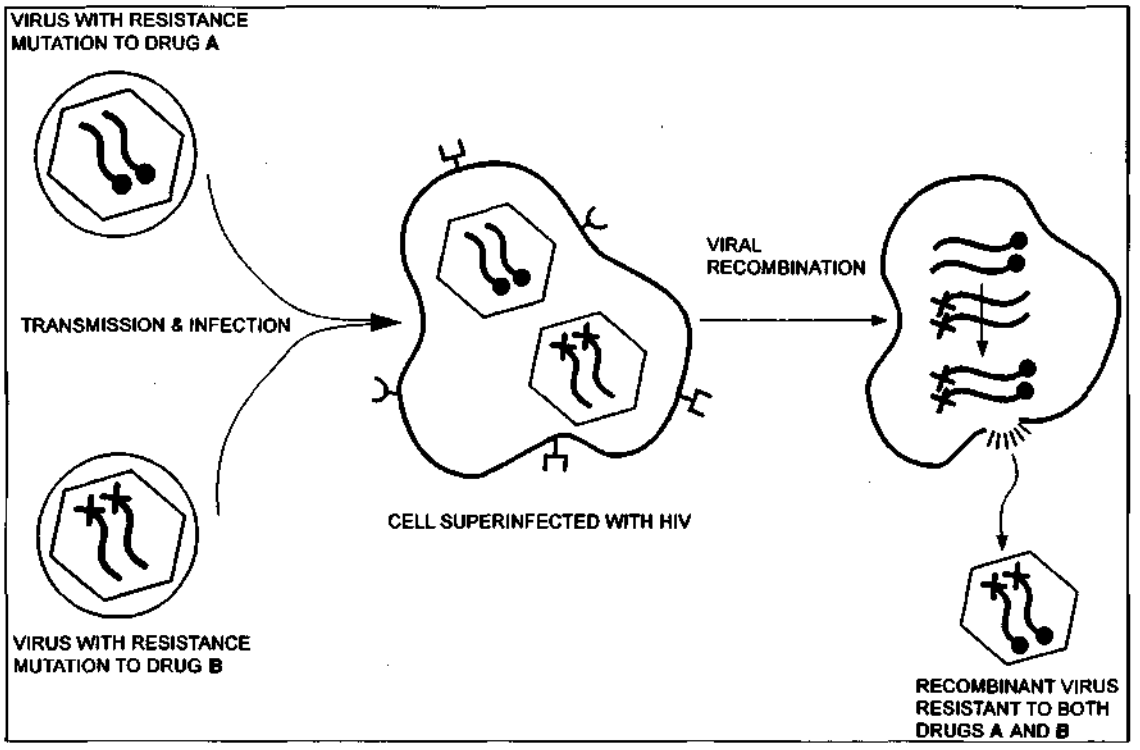


Figure 3: Dual infection of the same cell with two drug resistant variants forming a recombinant virus (Blackard *et al*, 2002)

Regular surveillance of dual infections is essential especially in areas where multiple subtypes circulate and in populations with a high percentage of individuals who are resistant to ARVs. Monitoring the frequency of dual infections is important since it is an indicator of the appearance of new recombinant viruses (CRFs and URFs). Within the last ten years dual and multiple infections have been reported in most parts of the world.

The incidence of dual infection in South Africa has not been reported at the population level. Studies have focused on high-risk HIV cohorts that have a large percentage of dual infections (Grobler *et al*, 2004). The present study aims to determine the prevalence of dual infection in

the ARV Roll-out programme in KZN which is a more general sample population than previous studies.

Although reports of dual infection have been published, undertaking this project is relevant to the pandemic since:

- The prevalence of dual infection has not been determined in KwaZulu-Natal
- Continuous surveillance of the genetic diversity of HIV is required

This knowledge could improve the management of HIV infected individuals.

1.8 Aims and objectives

Aims:

To document the subtype distribution of HIV in the KwaZulu-Natal ARV Rollout Programme

To determine the prevalence of dual HIV infection in the KwaZulu-Natal ARV Rollout Programme

Objective:

To establish and optimize methods for subtyping and detection of dual infection

Chapter two: Materials and Methods

Overview

Ethics approval was obtained from the University of KwaZulu-Natal Biomedical Ethics Committee. The protocol was approved for the degree of Master of Medical Science by the Post-graduate Education Committee of the Nelson R. Mandela School of Medicine.

Samples submitted for routine CD4 testing from the KwaZulu-Natal ARV Rollout programme were identified for use in this study. Routine testing was performed and residual whole blood, which would otherwise have been discarded, was used anonymously for this study. A minimum of 500µl whole blood was required for the study. Samples with less than the minimum volume were excluded from the study. The following information was obtained from the laboratory database: viral load and CD4 results, the patient's age and gender and the name of the ARV site attended. The data was recorded anonymously and no patient identifiers were used.

The Heteroduplex Mobility assay was performed on these samples to determine the HIV subtype and to screen for possible dual infection. Samples were screened for inter-subtype and sub-subtype dual infection using *gag* and *env* HMA respectively. Samples that were indeterminate by HMA were cloned and sequenced. Phylogenetic analysis was performed to determine whether the patient had possible dual infection.

The flow diagram (Figure 4) summarizes the study procedure.

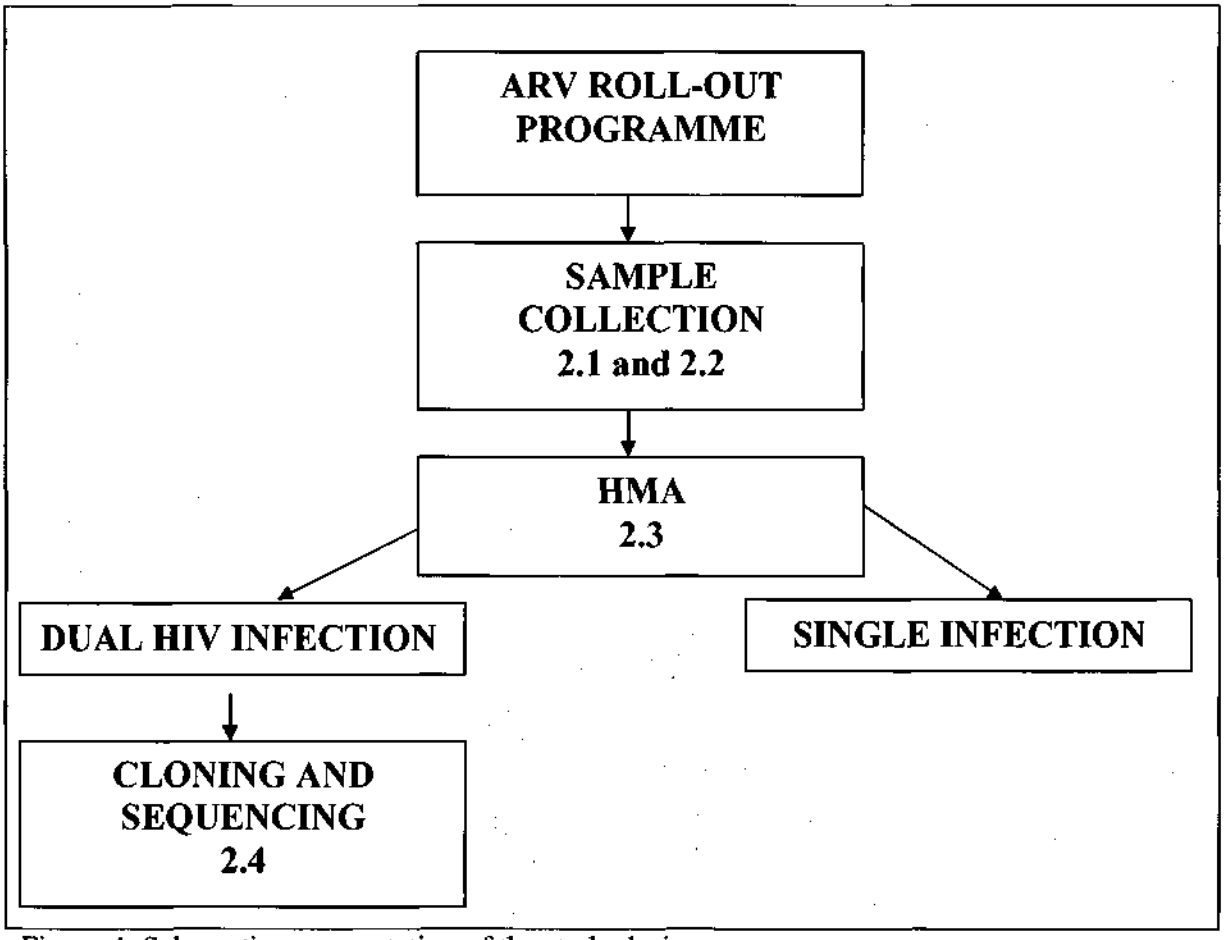


Figure 4: Schematic representation of the study design

2.1 Sample Size Calculation

This study utilized samples from a larger ongoing "treatment outcome" study investigating the impact of HIV sequence diversity and dual infection on the success of antiretroviral therapy in the KwaZulu-Natal ARV Rollout Programme. The rate of dual infection and the proportion of patients failing therapy were required to perform sample size calculations in the larger "treatment outcome" study.

Dual infection at the population level is unknown. Gottlieb *et al* (2004) established that dual HIV-1 infection is associated with rapid disease progression using 3 different cohorts, the combined size of which was 64, of which 5 individuals were dually infected (7.8%). Unpublished data from our laboratory was used to establish a crude estimate of the proportion of patients failing antiretroviral therapy (12%). CD-4 and viral load results of patients on ARV were used to establish patient success or failure, according to criteria defined by the Department of Health (DOH, ARV 2005). Data from 450 patients were analysed.

The minimum sample size required for the "treatment outcome study" was 83. This was calculated by a professional statistician using STATA™ 9 (StataCorp, College Station, Texas, USA). The prevalence of dual infection and the molecular epidemiology of HIV were described in these 83 patients for purposes of this dissertation. The patients were randomly selected using random numbers provided by the statistician, and included 53 patients who were treatment "failures" and 30 who were treatment "successes" as defined by Department of Health criteria (DOH, ARV 2005).

2.2 Sample Collection

Whole blood was routinely collected in heparin-free EDTA tubes (to prevent agglutination). These samples were sent from various ARV sites in KwaZulu-Natal to the Department of Virology, Inkosi Albert Luthuli Central Hospital, Durban for routine 6 monthly CD4 testing. The residual whole blood was used after routine testing was completed. White blood cells were extracted and stored at -70^0C within 48 hours. PCR was performed and the products were subsequently used for HMA, cloning and sequencing.

2.3 Subtyping using env and gag heteroduplex mobility assays

All samples in this study were first subtyped using *gag* HMA utilizing reference strains A, AE, AG, B, C, D, F, G, H and J. Determination of the sub-subtype was performed on *env* HMA using the reference strains of the subtype identified in the *gag* HMA . Those samples that were indeterminate by HMA were analysed further by clonal-sequencing to determine whether the patient had possible dual infection.

HMA kits were obtained from National Institutes of Health (Bethesda, USA). The HMA methods used in this study were adapted from the accompanying package inserts, and are described in the following sections.

2.3.1 Background on HMA

Base-pairing between double stranded DNA can be exactly complementary, with only A-T to G-C base interactions. The degree of base pairing influences the migration of dsDNA on a non-denaturing polyacrylamide gel. A completely complementary dsDNA will migrate faster than an identically sized PCR product (amplicon) with mismatches. According to Barlow *et al* (2000), mismatches cause 'bubbles' in the 'rod-like' amplicon that hinders its movement through the polymerised gel. Furthermore, an insertion or deletion of a nucleotide in one of the amplicon strands forms a 'bulge' that causes an even greater impediment in the electrophoretic mobility of DNA.

Heteroduplexes are formed when two non-identical but closely related single-stranded DNA fragments anneal. These molecules will have mismatches and unpaired bases resulting in the

structural distortions of 'bubbles' and 'bulges'. The retarded movement of a heteroduplex in comparison to a homoduplex on an acrylamide gel has shown to be proportional to the degree of divergence between the two sequences. Delwart *et al* (1993) were the first to realise the relevance of heteroduplex formation to viral genome diversity. Their observations showed that a PCR amplicon of the variable *env* region of HIV-1 migrated as a single band on a low resolving agarose gel, but was separated into several bands on a high-resolution neutral polyacrylamide gel (Delwart *et al*, 1993). This gave rise to the use of the Heteroduplex Mobility Assay in virology which was initially used to subtype HIV-1 (Delwart *et al*, 1994).

The principle of HMA is the mixing of an untyped sample with equal amounts of reference strains which are heated to 94-96^{0}C for two minutes to denature the dsDNA into single strands. These strands are immediately cooled on ice for annealing to occur. This generates the original homoduplexes and heteroduplexes composed of complementary single strands from the two different starting amplicons. The heteroduplex formed between the unknown sample and the most closely related reference strain will exhibit the fastest migration.

HMA has been successful in subtyping the following viruses: HCV, GBV, Norwalk-like viruses, influenza, measles, poliovirus and HIV-1. Different genes may be selected for HIV subtype determination using HMA. Delwart, Mullins and colleagues showed that heteroduplex formation using nested PCR on the *env* region could assign a subtype. Other more conserved regions of the genome like *gag* and *tat* have also been used for heteroduplex formation.

2.3.2 Methods

2.3.2.1 Sample preparation

White blood cells were isolated from whole blood using the HIV Amplicor test version 1.5 Wash buffer (Roche, Basel Switzerland) (Appendix I). Sodium phosphate, an integral component in the wash solution, isolates the white blood cells (leucocytes) from whole blood. White blood cells were stored at -70^{0}C until ready for testing.

2.3.2.2 Extraction of DNA

DNA was extracted using the NucliSens Magnetic extraction procedure (BioMerieux, Marcy l'Etoile, France) (Appendix I). This method is based on Boom chemistry using magnetic silica particles (Boom *et al*, 1989). Under high salt conditions provided by buffers, nucleic acid binds to the silica particles. Non-nucleic acid components are removed by several washing steps performed in the NucliSens miniMag (BioMerieux, Marcy l'Etoile, France). Finally, nucleic acids are eluted from the silica particles and were used in the PCR.

2.3.3 Polymerase chain reaction

PCR using *gag* primers was performed on all 83 samples. PCR using *env* primers was performed on all samples positive by *gag* PCR.

2.3.3.1 Principle of PCR

The polymerase chain reaction is the most-utilised technique for replicating DNA sequences *in vitro*. Two oligonucleotide primers, 15-30 base pairs in length, that are complementary to the target nucleic acid sequence, flank the DNA to be amplified. Primers are included in a reaction mixture that consists of the target nucleic acid, deoxynucleotides triphosphates (dNTPs), heat-stable DNA polymerase and a defined solution of salts (buffer, water, etc). (Kwok *et al*, 1987)

The mixture is exposed to repeated cycles of defined temperature changes. These thermal changes causes the denaturation (94 to97^{0}C) of the template DNA into single strands, the annealing (55-72^{0}C) of the primers to the target DNA, and the extension (72^{0}C) of the primers so that the target DNA is replicated. Subsequent heating cycles utilises the original and newly synthesised DNA strands as templates for another round of replication. Therefore, the number of target DNA strands doubles with each thermal cycle. The number of thermal cycles for infectious agents usually consists of 20 to 40 cycles. This results in a 10^5-10^6 fold increase in target nucleic acid concentrations within 3 to 4 hours.

For the purposes of this study nested PCR was performed on both *gag* and *env* genes of the HIV-1 genome. The 460bp product yielded by the nested *gag* PCR with outer primers (P202

and G1777) and inner primers (g17 and 1584) was used in the *gag* HMA. The 700bp product produced by nested *env* PCR with outer primers (ED5 and ED12) and inner primers (ES7 and ES8) was used in the *env* HMA. Reference plasmids consisting of clones from the different subtypes of HIV-1 for both the *gag* and *env* genes were also amplified. However, plasmids required amplification using only the second round PCR reaction.

2.3.3.1.1 Master Mix preparation

Specific concentrations of 10 X PCR buffer (without $MgCl_2$), $MgCl_2$, PCR grade water (distilled, autoclaved and filtered), deoxynucleotides (dNTPs), primers, and the enzyme, Taq Polymerase (Roche, Basel Switzerland) were optimised for a nested HIV DNA PCR. *Env* and *gag* HMAs required separate master mix reactions (Appendix I). The primers for each reaction were different since different regions of the genome were targeted (Table 5 and 6).

Table 5: Primers used in nested *gag* PCR

Primer	Sequence (5'-3')	Position in genome[1]
H1G777 (sense)	TCACCTAGAACTTTGAATGCATGGG	777-801
H1P202 (antisense)	CTAATACTGTATCATCTGCTCCTGT	1874-1898
H1*Gag*1584 (sense)	AAAGATGGATAATCCTGGG	1123-1141
g17 (antisense)	TCCACATTTCCAACAGCCCTTTTT	1566-1589

[1] Position relative to genome of HIV-1 group M strain ELI (Genbank accession number K03454)

Table 6: Primers used in nested *env* PCR

Primer	Sequence (5'-3')	Position in genome[2]
ED5 (sense)	ATGGGATCAAAGCCTAAAGCCATGTG	6556–6581
ED12(antisense)	AGTGCTTCCTGCTGCTCCCAAGAACCCAAG	7822–7792
ES7 (sense)	tgtaaaacgacggccagtCTGTTAAATGGCAGTATAGC	7001–7020
ES8 (antisense)	caggaaacagctatgaccCACTTCTCCAATTGTCCCTCA	7667–7647

[2] Position relative to HIV-1HXB2 genome (Genbank accession number K03455)

2.3.3.1.2 Amplification conditions

Amplification was conducted in a Biometra thermocycler (Whatman Biometra, Goettingen). Amplification conditions differed for *env* and *gag* HMA (Appendix I).

2.3.3.1.3 Detection using agarose gel electrophoresis

Agarose gel electrophoresis was conducted to detect the presence of the target HIV DNA sequence. The PCR product was visualized under ultraviolet transillumination (Whatman Biometra, Goettingen) on a two percent agarose gel stained with ethidium bromide (Sigma, Steinheim Germany), an intercalating agent of DNA (Appendix I). It was important to view a single band on the agarose gel because multiple bands interfere with distance measurements due to the formation of heteroduplexes between DNA strands of substantially different size.

2.3.3.1.4 Considerations taken to avoid contamination

Good laboratory practice guidelines were followed. Each aspect of PCR i.e. extraction, master mix preparation, amplification and detection were each allocated specific rooms or areas within the laboratory. A unidirectional flow was conducted (Figure 5) to prevent amplicon contamination (ie. pre-amplification to post-amplification).

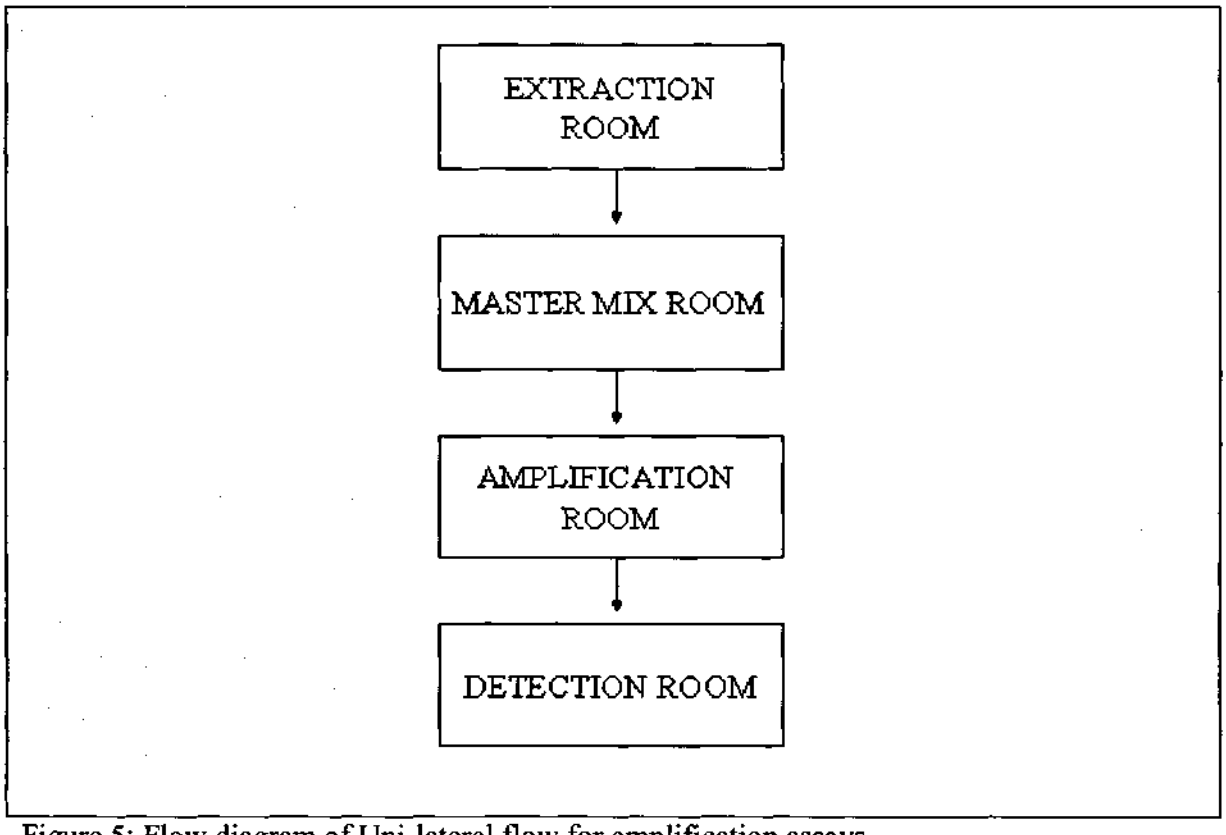


Figure 5: Flow diagram of Uni-lateral flow for amplification assays

2.3.4 Polyacrylamide gels

2.3.4.1 Background

The preparation of a polyacrylamide gel requires a monomer, acylamide and a bis-acrylamide crosslinker. A free radical initiator, tetramethylethylenediamine, and Ammonium Persulphate are added to a solution of monomer and cross-linker. Ammonium Persulphate reacts with tetramethylethylenediamine to form a tetramethylethylenediamine derivative and a free radical. The tetramethylethylenediamine radical reacts with an acrylamide molecule forming a tetramethylethylenediamine-acrylamide radical. Polymer chains are formed by repetitive addition of acrylamide monomers with a free radical terminus. Chain branching occurs when a bis-acrylamide molecule is added to the end of the chain.

2.3.4.2 Methods

A five percent polyacrylamide gel was prepared with acrylamide (Biorad, California USA), buffer and water. Tetramethylethylenediamine (Promega, Madison, WI) and Ammonium Persulphate (Molecular Sigma Biology, St. Louis USA) were added for polymerisation. For *gag* HMA, urea (Merck Chemicals, Gauteng SA) was added to allow better distinction of heteroduplexes (Appendix I). Extreme caution was used with the preparation of the polyacrylamide gel since it is a neurotoxin (Appendix III).

The mixture was loaded between assembled Hoefer gel plates (Hoefer, San Francisco USA) (Appendix I) and left to set for approximately 45 to 60 minutes. The gel containing urea required a pre-run of 15 minutes before loading of samples.

A heteroduplex annealing buffer was prepared (Appendix II) and added to the PCR product and reference strain. Each sample reaction was prepared in a 0.2ml PCR tube with 1.1 µl annealing buffer, 5 µl reference strain and 5 µl sample.

The tubes were heated at 94^0C for two minutes and immediately cooled on ice: rapid cooling facilitates the formation of stable heteroduplexes between divergent strains. The heteroduplex reaction was mixed with 3ul loading dye and carefully loaded onto the 5% non-denaturing polyacrylamide gel while the sandwich was on the caster. A molecular weight marker (Roche, Basel, Switzerland) was loaded simultaneously to serve as a reference marker.

The gel plates were carefully loaded onto the electrophoresis apparatus, placed in the gel tank and covered with 1 X Tris Borate EDTA (Appendix II). The gel was electrophoresed at a constant voltage of 250V for 3 hours for the *env* products and 250V for 2.5 hours for the *gag* products. The gel tank was placed in a water bath with ice packs to avoid the tank from heating and thus melting the gel.

The gel was removed from the apparatus, placed in an ethidium bromide (Merck, Germany) solution for 10 minutes and destained with distilled water. It was then viewed under ultraviolet light using the Biometra GelDocAnalyse version 1.0 (Whatman Biometra, Goettingen).

2.3.4.3 Electrophoretic conditions

The electrophoretic conditions are extremely important for heteroduplex stability. Therefore optimisation of the gel units, plates, acrylamide concentration, buffer and standard voltage/current conditions were carefully adjusted for this experiment.

2.4 Confirmation of dual HIV infection using cloning and direct sequencing

Cloning followed by sequencing of the clones is the most reliable method to confirm indeterminate infections (Brenner *et al*, 2004; Gottlieb *et al*, 2004). Therefore all samples analysed by HMA which showed more than one heteroduplex migrating to the furthest position on the gel, or had distinct slow migrating heteroduplexes or smears were cloned and then sequenced. *Taq* polymerase-amplified PCR products were used with a TOPO TA Cloning® Kit (Invitrogen, Carlsbad, California) to produce clones that were ready for sequencing.

2.4.1 PCR Clean-Up for sequencing and cloning

PCR products were purified using the QIAquick® PCR Purification Kit (Qiagen, Hilden). This procedure is based on spin-column technology and the selective binding properties of a uniquely designed silica-gel membrane. In effect three steps are involved ie. binding, washing and eluting of the DNA (Figure 6). DNA binds to the silica membrane in the presence of high salt concentrations while contaminants, unwanted primers and impurities such as enzymes, unincorporated nucleotides etc are washed out through the column. Binding buffers provided

in the kit maintain the high salt concentration and suitable pH (< 7.5) for DNA adsorption onto the silica membrane. Elution of the DNA occurs in the presence of low salt concentrations which are provided by the elution buffer or water.

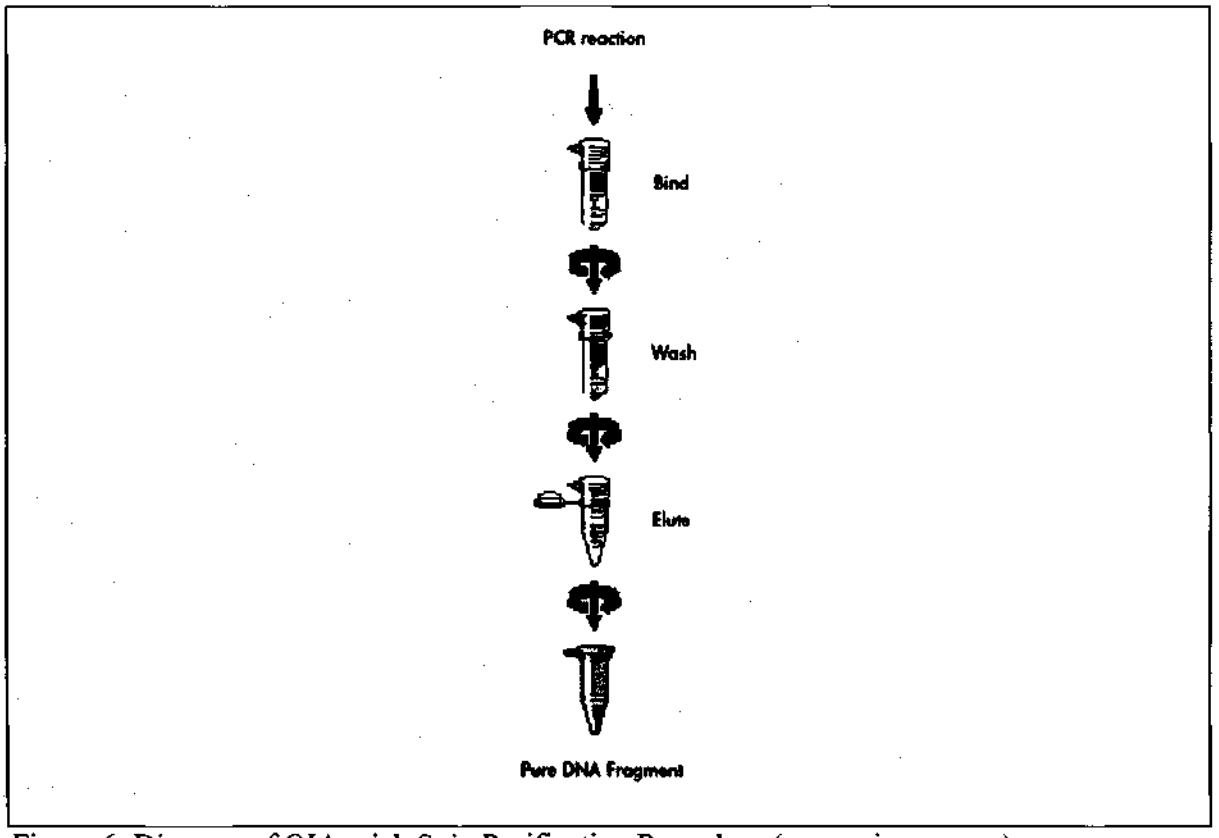


Figure 6: Diagram of QIAquick Spin Purification Procedure (www.qiagen.com)

2.4.2 Cloning

2.4.2.1 Principle

Several steps are performed to generate clones. Firstly the region of interest or the fragment of DNA containing the gene to be cloned is inserted into a circularized DNA molecule called a vector. The vector and the gene together form a recombinant DNA molecule. The vector acts as a vehicle that transports the gene into a host cell, where the recombinant DNA molecule

multiplies. The host cell also divides, and copies of the recombinant DNA molecule are passed to the progeny and further replication of the vector and gene takes place. Numerous cell divisions occur after which a colony, or clone of identical host cells are produced. Each cell in the colony contains one or more copies of the recombinant DNA molecule, the gene of interest carried by the vector is now said to be cloned.

The central component to a cloning experiment is the vector. Vectors must be capable of entering a host cell and once inside must replicate to produce multiple copies of itself. Two naturally occurring vectors are used: plasmids and virus chromosomes. For the purposes of this study a TOPO TA Cloning® Kit for Sequencing (Invitrogen, Carlsbad California) was used which contained a plasmid vector, pCR®4-TOPO® (Figure 7). The enzyme, topoisomerase was bound to the vector which changes the conformation of the plasmid by introducing or removing supercoils.

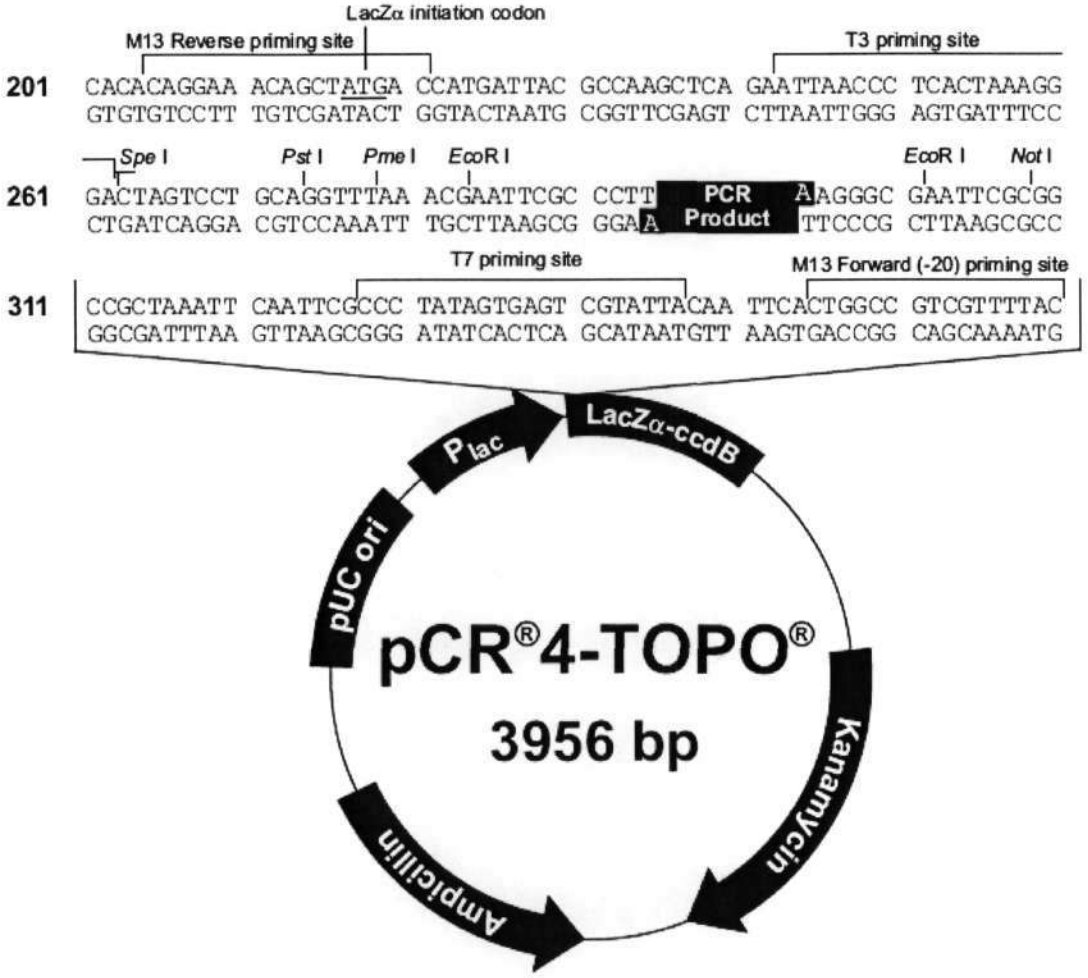


Figure 7: Diagram of vector pCR®4-TOPO®Map Invitrogen Life technologies

The PCR product to be sequenced is inserted into the cloning site situated between bases 294 and 295. This cloning site is contained within the N-terminal fragment of the *lacZ* gene of *E. coli*. This allows for blue/white screening to confirm that the insert has been included in the vector. A white colony represents the clone with the insert.

2.4.2.2 Method

A ligation reaction between the vector and the PCR product was performed in the presence of a salt solution. This resulted in the formation of a recombinant DNA molecule. In order to form multiple copies of this molecule, a transformation step was performed. This involved mixing the ligation reaction with DH5 a competent cells, and culturing the cells on LB agar plates. Clones were analysed by PCR for either the *env* or *gag* region of the HIV genome. PCR positive clones were sequenced (Appendix I).

2.4.3 Direct Sequencing

2.4.3.1 Introduction

Sequencing determines the order of the nucleotides of a gene or PCR fragment. In this study, sequencing was performed using ABI Prism BigDye Terminator v3.1 ready reaction Cycle Sequencing Kit and an ABI cycle sequencer 3100 (ABI, United Kingdom). A sequencing reaction i.e. denaturation, annealing and extension occurs that is similar to PCR but only one primer is used per reaction. Deoxynucleotides tri-phosphates (dNTPs) and dideoxynucleotides tri-phosphates (ddNTPs) are added which bind to complementary target sequences. When a ddNTP is incorporated, the extension reaction is terminated because a ddNTP contains an H-atom on the 3rd carbon atom whereas dNTP's contain an OH-atom on that position. The ddNTP's are fluorescently labeled. Therefore, it is possible to detect the color of the last base of this fragment on an automated sequencer. The sequencer separates sequencing ladders by using capillary arrays, which are automatically filled with polymer solution. In this study sequencing was performed to confirm dual infections determined by HMA.

2.4.3.2 Pre-sequencing sample preparation

PCR products were purified using the QIAquick® PCR Purification Kit (See 3.4.1) (Qiagen, Hilden). The concentrations of the PCR products were determined by comparison of the PCR products to a mass ladder (Invitrogen, Carlsbad, California).

2.4.3.3 Master mix preparation for sequencing reaction

A separate master mix was prepared for each primer (Appendix I). Only the second round primers were used ie ES7/ES8 and 1584/g17 for the *env* region and *gag* region respectively. 15µl of master mix was aliquoted into the appropriate wells and 5µl of diluted product into the appropriate well. The product and the master mix were mixed by pipetting and then placed in the thermocycler where the sequencing reaction took place under the conditions indicated in Table 7:

Table 7: Amplification conditions for sequencing reaction

Temperature (^{0}C)	Duration (seconds)	
96	10	} 25 cycles
50	5	
60	240	
4	indefinitely	

2.4.3.4 Clean-up of Sequencing reaction

An in-house clean-up method was used. This was based on the use of isopropanol (Merck, Germany) and ethanol (Merck, Germany) in conjunction with centrifugation (Appendix I). Samples were sequenced immediately or stored at -20^{0}C for no longer than two weeks.

2.4.3.5 Post-Sequencing Reaction

After the sequencing reaction and clean-up were performed the samples were reconstituted in Hi-Di™ Formamide (ABI, United Kingdom) and denatured at 95^{0}C for 3 minutes. Samples were sequenced on an ABI 310 cycle sequencer (ABI, United Kingdom) at the South African Sugar Research Institute.

2.5 Phylogenetic Analysis

Sequences of the 460bp *gag* regions of the HIV genome obtained from the ABI Sequencing system were edited using SeqMan (DNA Star) and BioEdit. Sequences were aligned using Clustal X. Phylogenetic Trees were constructed using PAUP Version 4. Phylogenetic trees were visualised in TreeView. Recombination and diversity on sequences was determined by Simplot® and MEGA®, respectively. Procedures described in the Los Alamos website were used to rule out contamination.

Chapter three – Results

3.1 Sample collection

Eighty three samples were analysed. All the patients were on ARV therapy. The age, gender, ARV site, CD4, viral load and subtyping results are summarized in table 8.

Table 8: Patient details

STUDY NUMBER	HIV-1 VIRAL LOAD	CD4 COUNT	AGE	GENDER	*Subtyping by *gag* HMA	*Subtyping by *env* HMA	ARV SITE
21	2200	28	36	F	C	unamplifiable	Addington
54	2150	152	35	M	C	C	KEH
56	410	507	6	F	C	unamplifiable	Stanger
61	60100	685	31	M	C	C	M.G.M.H.
76	590000	148	34	-	C	C	M.G.M.H.
77	21000	14	29	F	C	C	M.G.M.H.
84	290000	138	44	F	C	unamplifiable	M.G.M.H.
85	4400	318	31	F	C	unamplifiable	M.G.M.H.
86	170000	155	20	F	C	C	M.G.M.H.
89	<400	272	33	F	C	C	C.T.K.
90	<400	137	33	M	C	C	C.T.K.
91	130000	120		F	C	C	C.T.K.
92	<400	253	31	F	C	Indeterminate by *env* HMA	C.T.K.
94	54	250	45	F	C	C	C.T.K.
95	<25	81	45	F	C	unamplifiable	C.T.K.
96	<400	393	26	F	C	unamplifiable	C.T.K.
97	<25	323	44	F	C	unamplifiable	C.T.K.
98	<400	253	35	F	C	unamplifiable	K.E.H.
113	3200	325	32	M	Untypeable by HMA	Untypeable by HMA	Stanger
114	303000	112	29	M	C	C	-
125	450	257	37	M	Untypeable by HMA	Untypeable by HMA	Stanger
127	735	22	34	M	C	C	C.T.K.
128	1900	256	35	F	C	C	Stanger
130	6400	91	30	M	C	C	Stanger
137	1800	52	38	M	C	unamplifiable	Addington
144	140000	123	32	F	C	C	Stanger
148	169000	122	39	F	C	C	K.E.H.

STUDY NUMBER	HIV-1 VIRAL LOAD	CD4 COUNT	AGE	GENDER	*Subtyping by *gag* HMA	*Subtyping by *env* HMA	ARV SITE
160	2670	441	6	F	Untypeable by HMA	Untypeable by HMA	K.E.H.
164	4700	135	22	F	Untypeable by HMA	Untypeable by HMA	M.G.M.H.
182	210000	252	33	F	C	C	K.E.H.
201	18000	161	11	M	C	unamplifiable	P.M.M.H.
203	2500	316	26	F	C	unamplifiable	R.K.K.
205	44000	21	35	F	C	Indeterminate by *env* HMA	M.G.M.H.
207	14000	350	29	F	C	unamplifiable	K.E.H.
210	69000	154	36	M	Untypeable by HMA	Untypeable by HMA	Stanger
213	610	368	47	F	C	unamplifiable	R.K.K.
215	24600	81	37	F	Untypeable by HMA	Untypeable by HMA	R.K.K
224	87000	167	39	F	C	C	G.J.C
233	5600	171	33	F	C	C	R.K.K.
256	<25	450	29	F	C	unamplifiable	K.E.H.
258	<25	323	30	F	C	C	M.G.M.H.
261	<25	135	24	F	C	unamplifiable	M.G.M.H.
262	<25	190	36	F	A	unamplifiable	M.G.M.H.
263	2100	147	36	M	C	Indeterminate by *env* HMA	M.G.M.H.
266	16000	132	28	F	C	C	K.E.H.
268	3000	184	29	F	C	C	K.E.H.
276	<400	290	31	F	C	unamplifiable	St. Apollinaris
277	<25	117	40	M	C	C	K.E.H.
280	412	303	49	F	C	unamplifiable	Stanger
303	2600	265	10	F	C	unamplifiable	K.E.H.
339	<400	210	35	F	C	unamplifiable	Stanger
340	<400	146	32	F	C	C	Stanger
341	<400	127	34	F	C	C	R.K.K.
344	<400	254	40	F	C	C	Stanger
345	<25	482	7	F	C	unamplifiable	K.E.H.
346	<25	273	7	M	C	C	C.T.K.
347	1000	239	31	F	C, Dual infection confirmed by clonal-sequencing	Indeterminate by *env* HMA.	C.T.K.

STUDY NUMBER	HIV-1 VIRAL LOAD	CD4 COUNT	AGE	GENDER	*Subtyping by *gag* HMA	*Subtyping by *env* HMA	ARV SITE
350	<25	216	35	F	C	unamplifiable	M.G.M.H.
351	<25	255	29	M	C	C	K.E.H.
354	<25	168	31	F	C	C	G.J.C.
355	200000	284	26	F	C	C	St. Apollonaris
357	<25	328	26	F	C	C	R.K.K.
358	4500	218	36	M	C	C	R.K.K.
359	<25	210	36	M	C	unamplifiable	M.G.M.H.
362	65000	110	28	-	C	C	G.J.C.
378	14000	232	31	F	Untypeable by HMA	Untypeable by HMA	K.E.H.
385	660	204	33	M	C	C	R.K.K.
400	14000	98	53	F	Untypeable by HMA	Untypeable by HMA	Stanger
428	33000	218	42	M	Untypeable by HMA	Untypeable by HMA	Stanger
429	470000	507	7	F	Untypeable by HMA	Untypeable by HMA	Stanger
436	22000	365	28	F	C	C	K.E.H.
439	120000	148	9	F	C	unamplifiable	Stanger
441	52000	360	56	M	C	C	R.K.K.
442	2200	185	33	F	C	C	K.E.H.
443	3500	241	29	F	C	unamplifiable	K.E.H.
445	590	1257	7	M	C	unamplifiable	K.E.H.
450	1100	347	-	F	C	C	K.E.H.
453	12000	421	56	F	C	unamplifiable	K.E.H.
454	<25	258	-	F	C	C	R.K.K.
456	<25	149	36	F	C	unamplifiable	R.K.K.
457	<25	428	23	F	C	unamplifiable	R.K.K.
459	<25	79	42	M	C	C	R.K.K.
460	<25	327	32	F	C	unamplifiable	R.K.K.

* The subtyping results are discussed further in Section 3.3, page 66.

Note: CTK: Christ the King, GJC: G.J. Crookes KEH: King Edward Hospital, MGMH: Mahatma Ghandi Memorial Hospital, PMMH: Prince Mshiyeni Memorial Hospital, RKK: R. K. Khan

3.2 Polymerase Chain Reaction

3.2.1 *gag* products

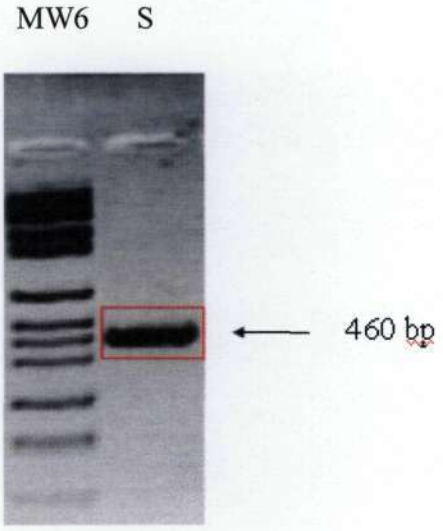


Figure 8: Detection of a 460 bp *gag* PCR product on a 2% agarose gel (boxed area, indicated by arrow). Molecular weight marker 6 is used as a reference for the PCR product.

Subtyping using HMA requires the identically sized segment of the equivalent position of the genome to be amplified in all samples and reference strains. The *gag* primers generate a 460 bp *gag* gene fragment (Figure 8) corresponding to the region coding for amino acid 132 of p24 up to amino acid 40 of p7 in the genome. The *gag* PCR was positive in 87.9% (73/83) of samples. The median viral load and CD4 count of unamplifiable samples was 19300 copies/ml and 225 cells/µl, respectively. However, those samples that were successfully amplified by *gag* PCR had a median viral load of 6000 copies/ml and a median CD4 count of 215.25 cells/µl.

3.2.2 env products of 700bp

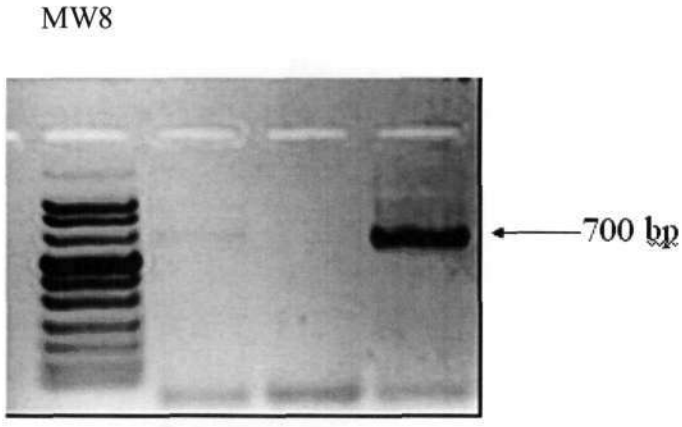


Figure 9: Detection of an *env* PCR product on a 2% agarose gel. The product is 700bp using molecular weight 8 as a reference (shown by arrow).

The *env* region is the most variable region of the HIV genome and is useful in subtype determination and determining sequence diversity. HMA is dependent on the mismatches between the sample and reference strains; therefore both the samples and the reference strains had to be amplified to the same sizes ie 700bp for *env* HMA. It was extremely important that the PCR product was a single band of the correct sized fragment. Figure 9 is a general representation of the *env* PCR product.

Env PCR was performed on the 73 samples positive by *gag* PCR. Of these 73 samples, 47 (64%) were *env* PCR positive. The remaining 26 samples were unamplifiable by *env* primers ES7 and ES8. The median viral load of the unamplifiable samples was 3950 copies/ml and the median CD4 count was 277.5 cells/µl. The amplifiable samples had a median viral load and CD4 count of 11200 copies/ml and 184 cells/µl, respectively.

3.3 Heteroduplex Mobility Assay Results - Polyacrylamide gels

3.3.1 *gag* Heteroduplex mobility Assay

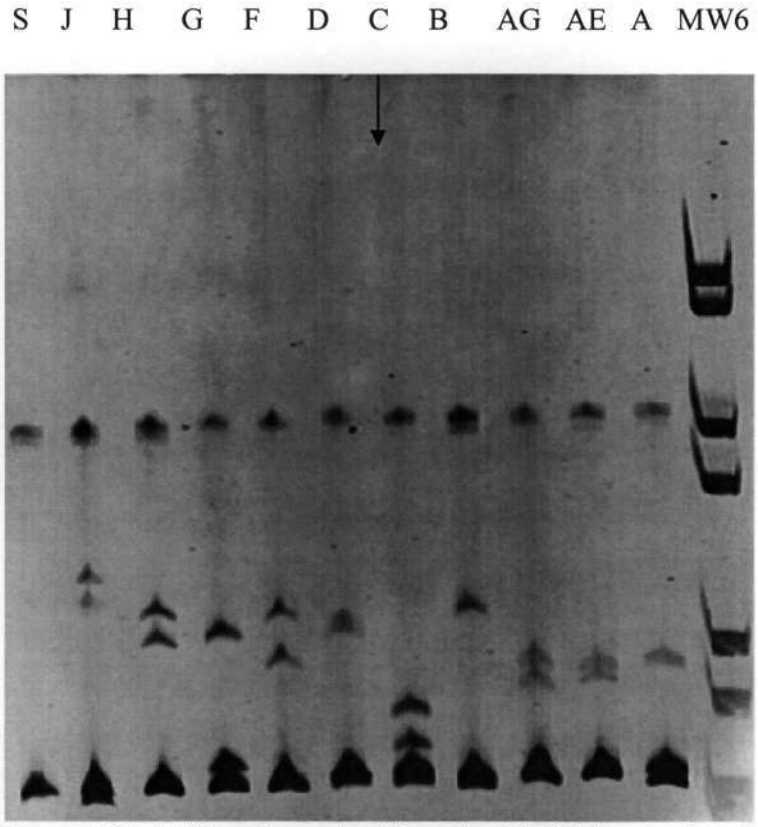


Figure 10: A 5% polyacrylamide gel with 20% urea showing subtype C as the resulting subtype (indicated by arrow). Reference strains of HIV-1 A, AE, AG, B, C, D, F, G, H and J of the *gag* region are used in subtyping. S represents the sample migrating without a reference strain.

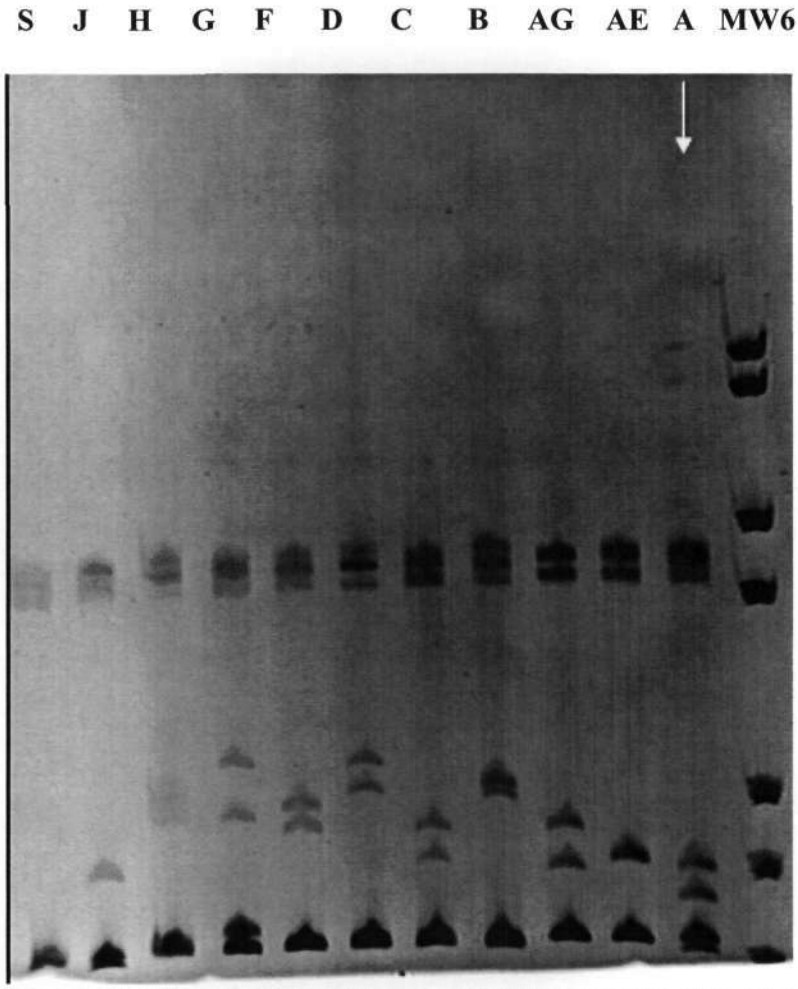


Figure 11: A 5% polyacrylamide gel with 20% urea showing subtype A as the resulting subtype (indicated by arrow). Reference strains of HIV-1 A, AE, AG, B, C, D, F, G, H and J of the *gag* region are used in subtyping. S represents the sample control i.e. the sample migrating without a reference strain.

All reference HIV-1 strains (A, AE, AG, B, C, D, F, G, H and J) from the *gag* HMA kit were mixed with each sample to form heteroduplexes. The resulting subtype was determined by the heteroduplex that migrated to the furthest point on the polyacrylamide gel. In Figure 10, the sample is a subtype C while in Figure 11 the sample is a subtype A.

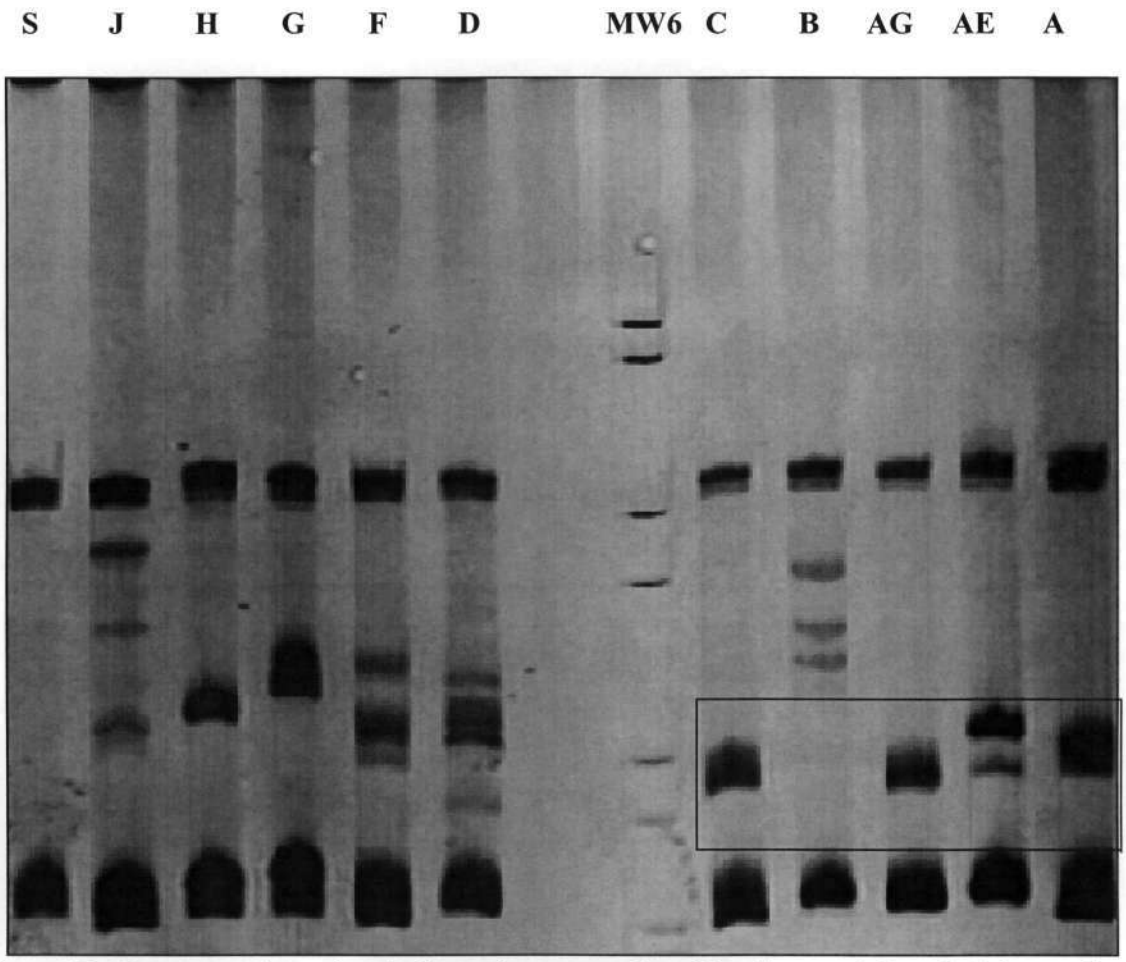


Figure 12: Representation of an indeterminate sample with subtypes AG and C detected on a 5% polyacrylmide gel using 20% urea (indicated by boxed area). A, AE, AG, B, C, D, F, G, H and J are HIV-1 subtypes used to from heteroduplexes with the sample.

In some instances, a 20% urea gel resulted in more than one subtype migrating to the furthest position on the gel (Figure 12). It was suggested that in cases where the gel presents multiple bands at equivalent positions, a 30% urea gel should be used to confirm this result (Heyndrickx *et al*, 2000). A 30% urea gel allowed for more separation between subtypes with better resolution of heteroduplexes. As Figure 13 shows, this sample is a subtype C.

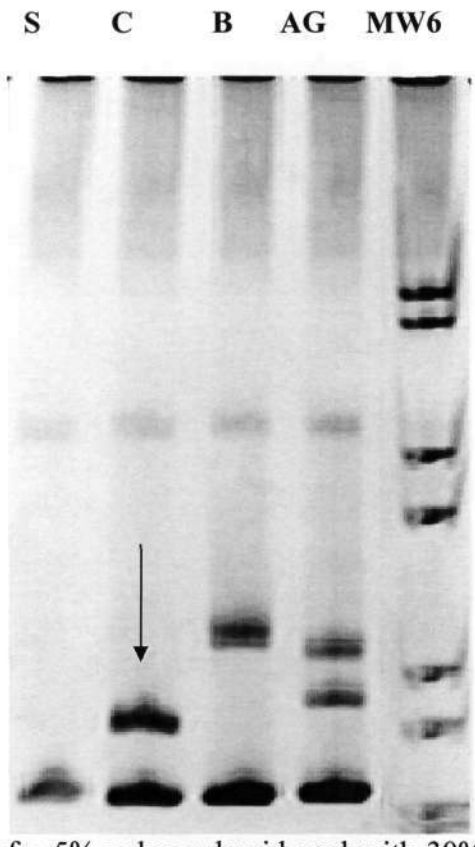


Figure 13: Representation of a 5% polyacrylamide gel with 30% urea. This gel was used to confirm heteroduplexes migrating to the equivalent furthest position on the 20% urea gel. AG, B and C are HIV-1 subtypes used to form heteroduplexes with the sample. S is the sample with water used as a control for diversity within the sample. The arrow points to the resulting subtype for this sample which is subtype C.

All samples analysed by *gag* HMA were distinctly of a single subtype. None of the samples were indeterminate. Subtype C accounted for 98.6% of infections (n=73) while 1 sample was a subtype A infection (Table 8).

3.3.2 env Heteroduplex Mobility Assay

3.3.2.1 Single HIV infections

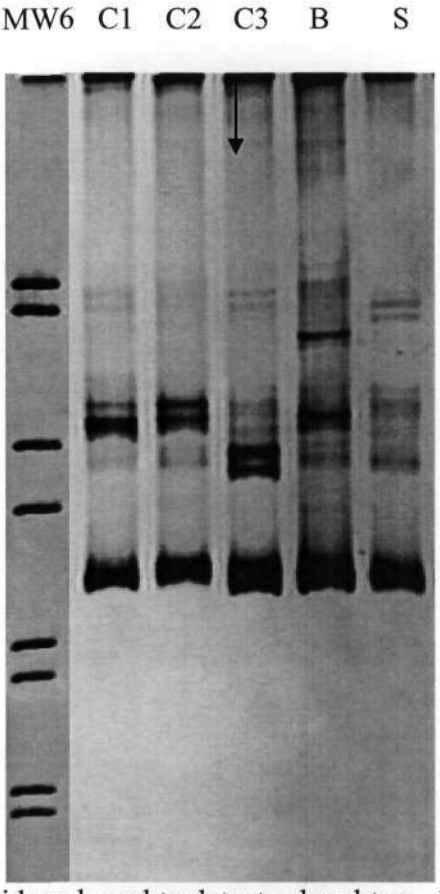


Figure 14: A 5% polyacrylamide gel used to detect sub-subtype dual infections using *env* reference strains C1, C2, C3 and B. The arrow points to the resulting sub-subtype which is C3. Lane S is the control lane of the sample and water.

Env HMA was used to determine sub-subtypes. The sub-subtypes C1, C2 and C3 were used to form heteroduplexes with the amplified PCR product. Like *gag* HMA, the heteroduplex migrating to the furthest point on the gel is the resulting subtype or intra-subtype for that sample. A representation of this is shown in Figure 14.

3.3.2.2 Indeterminate Result

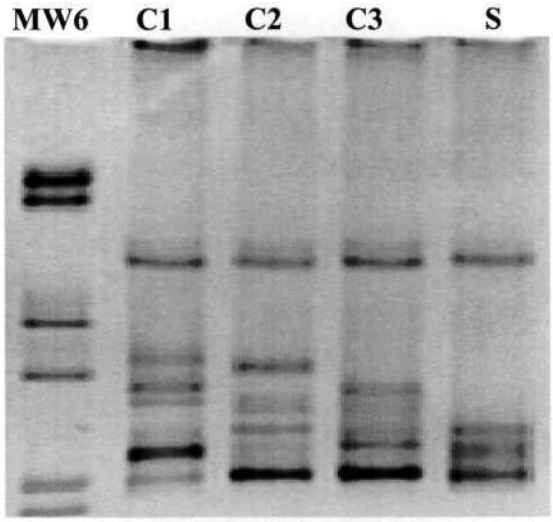


Figure 15: Representation of indeterminate *env* HMA on a 5% Polyacrylamide gel. Sub-subtypes C1, C2 and C3 are used with the sample. S represents the sample electrophoresed without reference strains.

Subtypes are determined on HMA by the heteroduplex that migrates to the furthest position in the gel. Samples in which a single subtype cannot be unequivocally determined require clonal sequencing to determine the subtype or the presence of dual infection or recombination. Such indeterminate results occur when there are distinct slow migrating heteroduplexes or smears (Gottlieb *et al*, 2004; Grobler *et al*, 2004) or when two hetroduplexes migrate to the furthest equivalent position on a polyacrylamide gel. In this study, 4/47 (8.5%) of the samples showed bands on HMA which could not be subtyped as a single infection and therefore required clonal sequencing as confirmation. An example of such a gel is shown in Figure 15.

3.4 Cloning Results

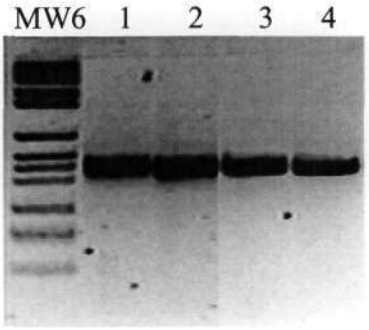


Figure 16: Agarose gel detection of *gag* PCR products amplified from clones

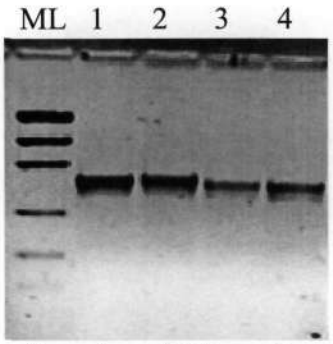


Figure 17: Agarose gel detection of *env* PCR products amplified from clones

Selected clones analysed by PCR that contain the inserts of interest i.e. *gag* (460 bp) and *env* (700bp) are shown (Figures 16 and 17, respectively).

3.5 Sequencing results

The clones of the 4 indeterminate samples on HMA were sequenced. The phylogenetic analysis of these clones are shown below. The electropherogram from clone 1 of sample 92 is shown in Appendix II. The results of the other clones were similar and have therefore not been shown.

3.6 Phylogenetic Analysis

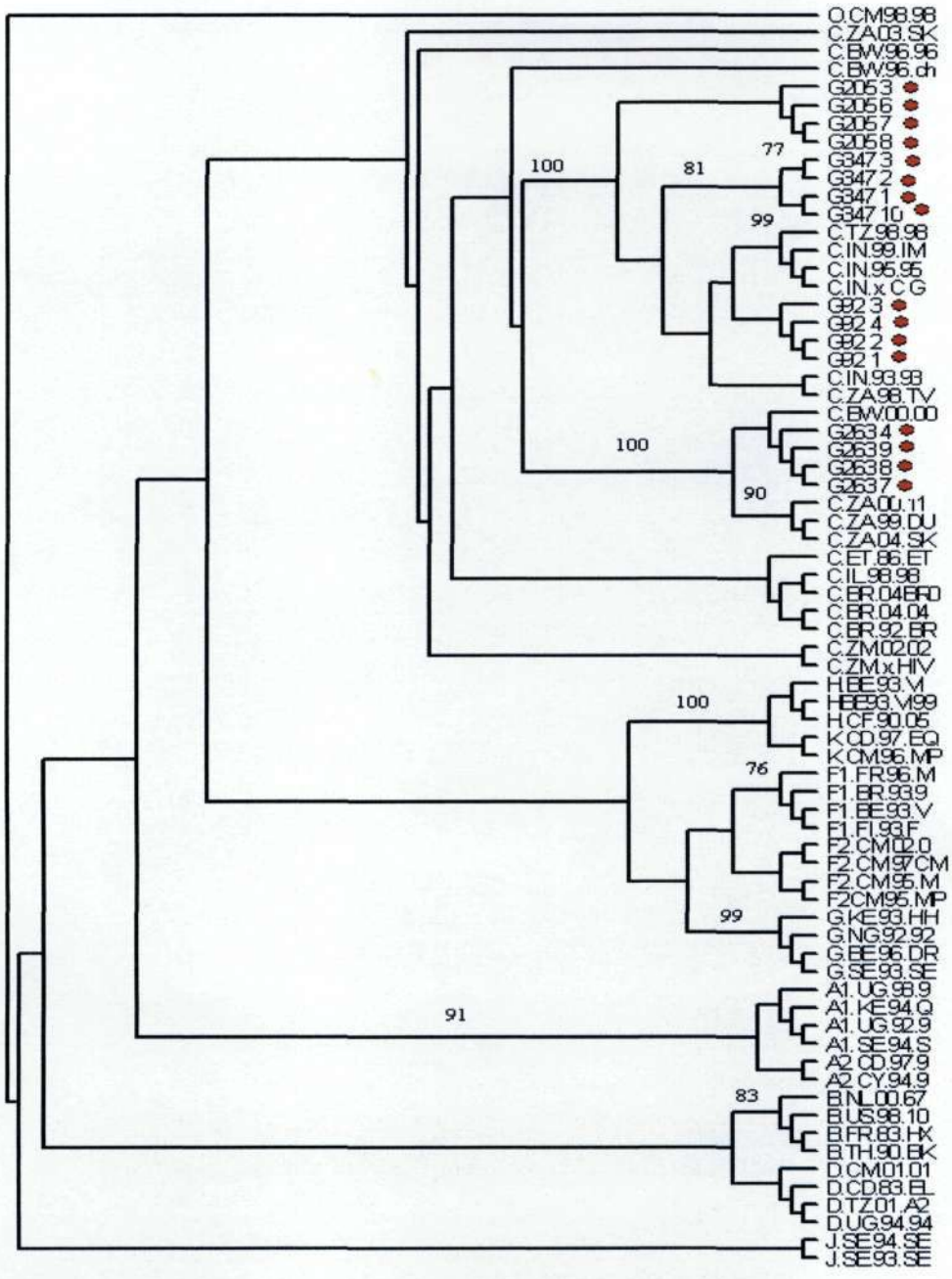


Figure 18: Phylogenetic tree showing HIV-1 subtypes. HIV-1 subtypes using *gag* sequences (460 bp) were used to construct this tree with the neighbour-joining method. Subtype C *gag* clones of samples 92, 205, 263 and 347 are shown with in red dots. Reference sequences are from the Los Alamos database (available at: http://www.hiv.lanl.gov). Bootstrap values (500 replicates) >75% are shown.

The results of the phylogenetic analysis of the 4 clones using the *gag* region (460bp) of HIV-1 subtypes A, B, C, D, F, G, H, J and K are shown in Figure 18. These show the monophyletic clustering of the clones of samples 92, 205, 263 and 347. Dual infection is defined as the presence of two infecting viruses within an individual that are no more closely related to one another than to at least one unlinked strain in the HIV database (Gottlieb *et al*, 2004). Based on this definition, these samples cannot be classified as inter-subtype dual infection.

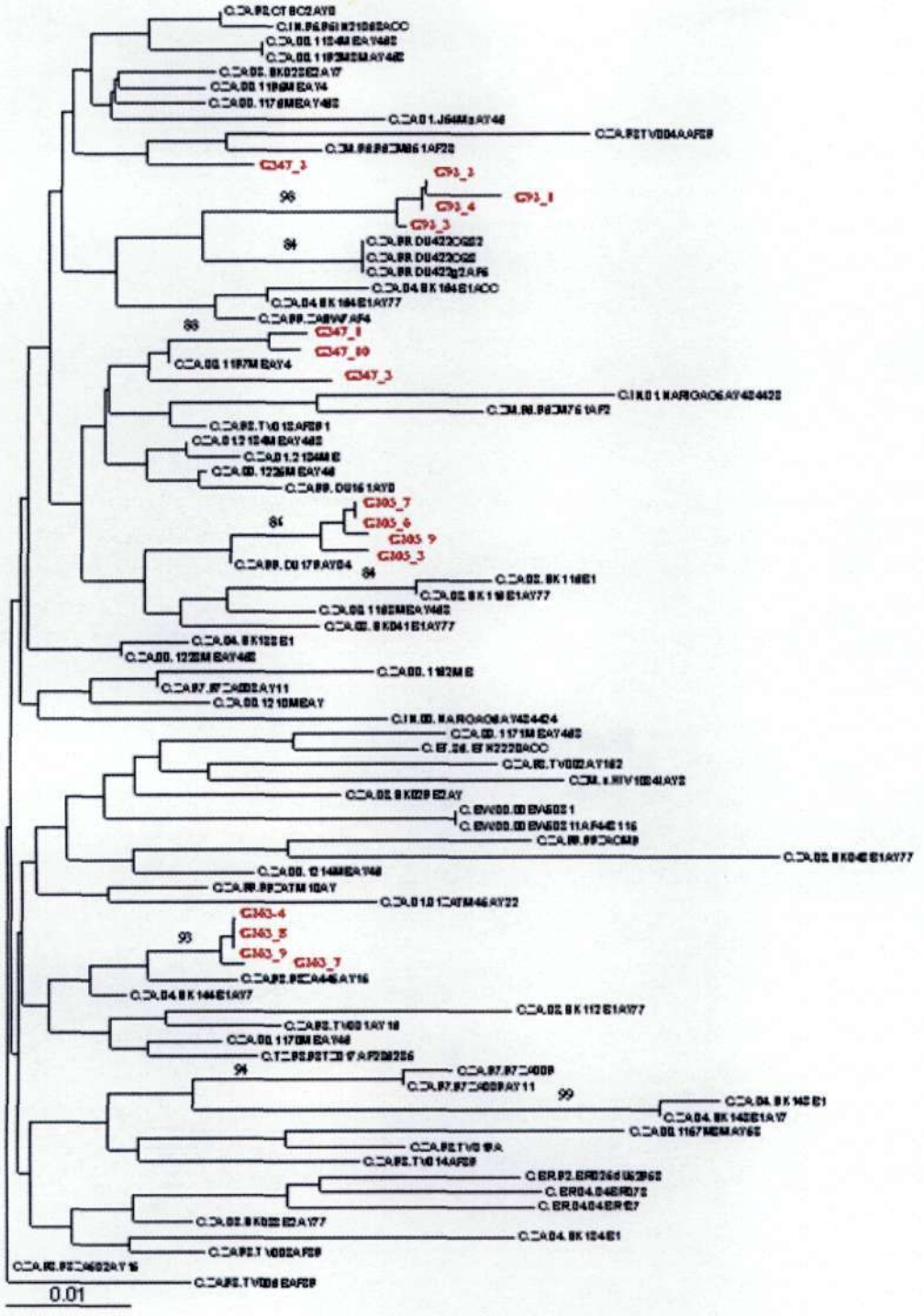


Figure 19: Phylogenetic tree of *gag* Subtype C sequences. HIV-1 *gag* sequences (460bp) were used to construct this neighbour-joining method. Subtype C clones of samples 92, 205, 263 and 347 are shown with symbols. Reference sequences are from the Los Alamos database (available at: http:/www.hiv.lanl.gov). Five hundred bootstrap replicates were analysed and those >75% are shown.

The results of the phylogenetic analysis of the clones using the *gag* region (460bp) of HIV-1 subtype C are shown in Figure 19. This shows dual infection of sample 347 which grouped separately on the phylogenetic tree. Clones of samples 92, 205 and 263 cluster monophyletically and therefore cannot be classified as inter-subtype dual infection. Simplot was used to exclude recombination. Further details of these samples appear in Table 8.

4. Chapter Four – Discussion

HIV, the causative agent of AIDS, has extraordinary genetic variability. This diversity is due to the lack of a proof-reading mechanism by the enzyme reverse transcriptase (Battula and Loeb, 1976) and its template switching ability during reverse transcription (Pulsinelli and Temin, 1994).

The genetic diversity of HIV has an impact on the management of the disease. A major concern is that all variants of HIV must be recognizable by diagnostic tests and be equally susceptible to ARVs. In addition, an HIV vaccine must be able to prime the immune system for infection with any form of the virus. The appearance of new HIV variants due to the recombinogenic activity of the virus will have an impact on diagnosis, treatment and prevention.

The diversity of HIV in the form of species, groups, subtypes, sub-subtypes, URFs and CRFs creates the possibility of a patient being infected with more than one variant i.e. dual infection. This occurs either as a co-infection when two viral variants cause infection at or near the same time before seroconversion, or as superinfection (reinfection) in which primary infection is followed by infection by another viral variant after seroconversion.

Dual HIV infection is a pre-requiste for recombination, which results in CRFs and URFs. Dual infection causes rapid disease progression (Gottlieb *et al*, 2004) and may also affect treatment response (Smith *et al*, 2005). HIV superinfection implies that the immune system

cannot prevent a second infection after responding vigorously to an initial infection. This has important consequences for vaccine development.

Continued surveillance for the emergence of dual infection is therefore necessary due to its impact on HIV management and prevention. The objective of this study was to establish and optimize methods for subtyping and detection of dual infection. The first aim was to determine the prevalence of dual infection in the target population and the second, to document the subtype distribution of HIV in KwaZulu-Natal.

Samples submitted for routine CD4 testing from the KwaZulu Natal ARV Rollout programme were identified for use in this study. Routine testing was performed and residual whole blood, which would otherwise have been discarded, was used anonymously for this study. The Heteroduplex Mobility assay was performed on these samples to determine the HIV subtype and to screen for possible dual infection. Samples were screened for inter-subtype and sub-subtype dual infection using *gag* and *env* HMA respectively. Samples that were indeterminate by HMA were cloned and sequenced. Phylogenetic analysis was performed to determine whether the patient had possible dual infection. Dual infection by clonal-sequencing and phylogenetic analysis was defined as the presence of two infecting viruses within an individual that are no more closely related to one another than to at least one unlinked strain in the HIV database. (Gottlieb *et al*, 2004). Recombination and diversity were determined by Simplot® and MEGA®.

Dual infection was found in 1/47 (2%) of the study population. The prevalence of dual infection in KwaZulu-Natal has not been previously reported for a normal-risk population. The prevalence of dual infections in other population-based studies was higher than in the present study. In Tanzania, the prevalence of dual infection was 28% (Herbinger *et al*, 2006). In a sex-worker cohort study the percentage of dual infection was 19.6% (Grobler *et al*, 2004). However, the risk groups analysed were different to the present study. The Tanzanian study used normal-risk and high-risk populations (Herbinger *et al*, 2006). Grobler *et al* (2004) also used a high-risk population in which subjects had multiple exposures (up to 437 exposures per year). The samples from the ARV Rollout programme were from a diverse group of individuals who most likely had a lower number of exposures per year.

Furthermore, Grobler's study was longitudinal in which a cohort of female sex workers were followed-up over a course of 24 months. The risk of dual infection was highest at the time of initial infection and most likely represented co-infection. The present study is cross-sectional and consists of patients on ARVs with CD4 counts below 200 cells/µl. Most of the patients in this study are therefore likely to have had chronic infection and the study sample would not have been taken during the time of initial infection.

The low prevalence of dual infection in the present study could also be attributed to several technical factors. HMA may not detect dual infections if the second virus is present in low amplifiable copy numbers (Grobler *et al*, 2004). In addition, four clones were sequenced for each sample suspected to have dual infection. Studies that have detected dual infections have used between 20-134 clones per sample to confirm dual infections (Grobler *et al*, 2004;

Gottlieb *et al*, 2004). Selecting more clones increases the chances of detecting dual infections. In future studies, more clones need to be selected in order to increase the sensitivity of the technique in identifying dual infections.

The HIV-1 subtype distribution in South Africa consists mainly of subtype C infections. CRFs and subtypes A, B and D also contribute to the epidemic but account for only 5.3% of infections in South Africa (Los Alamos, 2006). The limited circulation of HIV subtypes in South Africa could explain the low percentage of dual infections (2%) reported in this study. In areas that have multiple subtypes circulating, dual infection between subtypes is more frequent (Herbinger *et al*, 2006).

The present study reported an intra-subtype C dual infection. This reflects the predominance of subtype C in KwaZulu-Natal. Intra-subtype dual infections have also been documented in other studies. The South African sex-worker cohort reported intra-subtype C dual infections (Grobler *et al*, 2004). Gottlieb *et al* (2004) reported four intra-subtype B infections and one intra-subtype C dual infection. Other reports of dual infections between subtypes also reflect the main circulating strains in the area. For instance, HIV-1 and HIV-2 dual infections have been reported in Senegal (Curlin *et al*, 2004). Co-infections with subtype A and subtype D has been reported in Kenya (Steain *et al*, 2005).

The second aim of this study was to determine the molecular epidemiology of HIV in KwaZulu-Natal. Infections were all subtype C, except one which was subtype A (1/47). These results are in accordance with previous data which demonstrate that subtype C has dominated

the pandemic in South Africa (Los Alamos, 2007). However, non-C types have been reported. Subtype B and subtype D are responsible for a few infections (Treurnicht *et al*, 2002). Subtype A and a few recombinants are also present in the population (van Harmelen *et al*, 1999) and may have been introduced to South Africa by immigrants from neighbouring countries like Mozambique, Zimbabwe, Botswana and Namibia or other parts of Africa (Los Alamos, 2007).

The finding of an intra-subtype C dual infection in this study reflects the predominance of one subtype in KwaZulu-Natal. Other dual infection studies using South African cohorts have also reported intra-subtype C dual infections (Gottlieb *et al*, 2004; Grobler *et al*, 2004). In addition, 98.6% (n=73) of the samples were subtype C infections. Globally, subtype C is the greatest contributor to the epidemic, causing 50 per cent of infections (Hemelaar *et al*, 2006). The circulation of subtypes in South Africa is due to two separate epidemics. Initially the homosexual population presented with infections by subtypes B and D which was followed by a subtype C epidemic in the heterosexual black population (Williamson *et al*, 1995). Other subtypes have also been identified in South Africa and the present study identified a subtype A infection.

Subtype C infections have spread rapidly throughout the world. Its prevalence of 50 per cent currently dominates the pandemic (Hemelaar *et al*, 2006). Strains were first introduced in Ethopia in the early eighties (Salminen *et al*, 1996). Countries within Southern African have the highest prevalence of Subtype C infections. China, India and Brazil are areas highly affected by subtype C infections as well (Los Alamos, 2007).

Viral, host and socioeconomic factors may be related to the increase in C-type infections. Viral differences have been noted in HIV-1 Group M subtypes. Subtype C is the only variant to have an extra NF-κB binding site. Studies have shown that the triple (NF)-κB motif confers a higher promoter activity. The response to the proinflammatory cytokine tumor necrosis factor-a is increased with the triple (NF)-κB configuration. This suggests that subtype C may have a replication advantage in individuals with chronic immune activation. Other differences between subtype C and the other HIV Group M subtypes have been identified. For instance, the protease genes may have increased catalytic activity relative to other subtypes (Oliveira *et al*, 2003). The regulatory proteins, *Tat* and *Rev* are prematurely truncated and the *vpu* reading frame has a 15-bp insertion at the 5' end (Huang *et al*, 2003) in HIV subtype C. The rapid spread of subtype C may be due to the use of the co-receptor CCR-5 (Ball *et al*, 2003). Further studies are required to validate these queries.

The findings of this study need to be confirmed in large-scale epidemiological and long-term longitudinal investigations. However, there are limitations to the methods used for subtyping and detection of dual infection. Sequencing is the gold standard for subtyping but is expensive and labour-intensive. Furthermore clonal-sequencing requires a large number of clones to be selected. This procedure is expensive and labour-intensive and therefore impractical in the clinical setting.

The HMA subtyping method is also problematic. Many samples were non-amplifiable by both *gag* and *env* primers from the HMA kit used in this study. Eighty-three samples were selected for PCR. The *gag* PCR was positive on 73 samples. The remaining 10 samples were

unamplifiable by *gag* PCR. The median viral load of the unamplifiable samples (19300 copies/ml) was greater than the median viral load of the amplifiable samples (6000 copies/ml). This implies that a lower copy number, as measured by plasma viral load, is not a reason for non-amplification by *gag* PCR. 46 out of 73 samples were positive by *env* PCR. Unlike *gag* PCR, *env* PCR was more likely to be positive if the patient had a high viral load and a low $CD4^+$ cell count.

Rigorous optimization tests were conducted to improve the number of PCR positive samples. Primer and $MgCl_2$ concentrations were altered individually to improve primer annealing. Template volumes were also changed. DNA concentrations were determined by spectrophotometry to ensure that DNA extraction was optimal. However, the samples remained non-amplifiable even with optimal PCR conditions.

The inability to amplify samples by PCR could be due to primer-target mismatches. This could arise as a result of the genetic diversity of the target sequences. This has implications for subtyping HIV since primers used for subtype determination must amplify all HIV variants. The primers used for *gag* and *env* PCR should be designed to anneal to all types of HIV variants. Furthermore, HMA kits should have different PCR primers and reference strains for different geographic regions. For instance, an HMA kit for South Africa should consist of reference subtype strains from South Africa and primers designed on the basis of the subtype circulation in South Africa. Another limitation of the HMA kit is the lack of a positive control for dual infection. In addition, HMA is only a screening method for dual infection, which requires confirmation by further clonal sequencing and phylogenetic analysis.

Despite technical and methodological limitations, this study has highlighted important aspects of the molecular epidemiology of HIV in this region.

This study confirms the predominance of subtype C in this study population and therefore highlights the importance of developing a subtype C vaccine for this region. On the other hand, the finding of a single subtype A infection within this region emphasizes the importance of designing a vaccine with broad cross-clade efficacy. A more unusual finding in this study, which may be of interest to vaccine research, is the finding of a dual sub-subtype C infection. This implies that infection with one particular subtype of HIV is insufficient to prevent infection by viruses belonging to the very same subtype. This raises the possibility that should a subtype specific vaccine be used in this region, that not all infections by the same subtype would be prevented. However, a limitation of this study is that it did not determine how dual infection occurred i.e. whether it occurred as a co-infection or a superinfection. Further prospective studies are required to document the occurrence of superinfection, the associated immune responses and the potential impact this would have on subtype C vaccines.

The occurrence of dual infection has implications for other preventative measures. Safe-sex needs to be practiced by HIV-infected persons since they may not be immune to infection by other variants of HIV. HIV positive patients need to be counseled that dual infection causes rapid disease progression (Gottlieb *et al*, 2004) and may also affect treatment response (Smith *et al*, 2005). The magnitude of these effects needs to be investigated within the ARV Rollout programme in this region.

The present study is the first performed in South Africa using samples from a normal-risk population to determine dual infection. This study confirmed 1/47(2%) samples to be dually infected. The low prevalence of dual infection suggests that this phenomenon is unlikely to have a significant impact on diagnosis, treatment and vaccine development at this point in time. However, due to the dynamic nature of the epidemiology of HIV, continued surveillance is necessary to monitor the trends in the prevalence of dual infection.

Conclusion

The present study aimed to determine the prevalence of dual HIV infection in KZN. One case of intra-subtype C dual infection was identified. The rest of the sample population was single HIV-infections consisting of subtype C infections and one subtype A infection. The incidence of dual infection in KZN is low. At present, dual infection is unlikely to have an impact on diagnosis, treatment and prevention. However continued surveillance of dual infection and the molecular epidemiology of HIV is necessary due to the dynamic nature of the HIV pandemic.

Appendix I

Isolation of white blood cells (Package Insert: AMPLICOR® HIV-1 DNA Test, version 1.5)

HIV-1 DNA was extracted within a designated Extraction room. Whole blood was manually inverted 10-15 times to ensure thorough mixing of blood. 1ml of Blood Wash was added to each 2ml screw-cap tube. Removal of caps from each tube of whole blood was performed with the use of a gauze pad to avoid aerosol contamination. 500 µl of whole blood was pipetted into each tube containing Bld WS using a micropipette with an aerosol barrier tip. The tubes were capped and then mixed by inverting tubes 10 to 15 times. The tubes were then incubated for 5 minutes at room temperature after which the tubes were inverted 10 to 15 times again. The tubes were incubated for an additional 5 minutes at room temperature. Using a microcentrifuge (Jouan, France) the tubes were centrifuged at full speed (14000 rpm) to pellet the white blood cells.

The resulting supernatent was aspirated with a fine–tip transfer pipette without disturbing the WBC pellet. 1ml of Bld WS was added to each tube, which was then vortexed to resuspend the pellet. The tubes were microcentrifuged at maximum speed for 3 minutes. This step was performed for at least 3 times or until the pellet was free of red blood cells.

The final aspiration of supernatent was carefully performed to remove all traces of Bld WS. The dry pellet could be extracted immediately or stored at -70^0C until extraction.

PCR: Master Mix preparation

The *env* subtyping master mix was prepared according to concentrations from Heteroduplex mobility assay, HIV-1 *env* subtyping kit, protocol version 5 (National Institutes of Health, Bethesda, USA). This is shown in Figure 20.

First round (50 ul reaction)	
10 X PCR reaction buffer (without $MgCl_2$)	5.0 µl
$MgCl_2$ (25mM)	2.5 ul
Water	25.0 µl
dNTPs (10mM)	1.0 µl
ED5 (5pmoles/ul)	2.0 ul
ED12 (5pmoles/ul)	2.0 ul
Taq (1U/ul)	2.5 ul
Template	10.0 ul
Second round (50 ul reaction)	
10 X PCR reaction buffer (without $MgCl_2$)	5.00 ul
$MgCl_2$ (25mM)	2.50 ul
Water	34.20 ul
dNTPs(10mM)	1.00 ul
ES7 (5pmoles/ul)	2.00 ul
ES8 (5pmoles/ul)	2.00 ul
Taq (1U/ul)	1.25 ul
Template	2.0 ul

Figure 20: *env* nested PCR master mix

The master mix for *gag* gene subtyping was prepared according to the protocol HIV-1 group M *gag* HMA Subtyping Kit (Protocol Version 5, February 2002) (National Institutes of Health, Bethesda, USA). This is shown in Figure 21.

First round (50 ul reaction)	
10 X PCR reaction buffer (without $MgCl_2$)	5.0 ul
$MgCl_2$ (25mM)	3.0 ul
Water	36.0 ul
dNTPs (10mM)	0.5 ul
H1G777 (10pmoles/ul)	1.0 ul
H1P202 (10pmoles/ul)	1.0 ul
Taq (1U/ul)	2.5 ul
Template	1.0 ul
Second round (50 ul reaction)	
10 X PCR reaction buffer (without $MgCl_2$)	10 ul
$MgCl_2$ (25mM)	10 ul
Water	64 ul
dNTPs(10mM)	1 ul
g17 (10pmoles/ul)	4 ul
H1*Gag*1584 (10pmoles/ul)	4 ul
Taq (1U/ul)	5 ul
Template	2 ul

Figure 21: *gag* nested PCR master mix

PCR: Amplification conditions

***Env* nested PCR amplification conditions** {HIV-1 *env* subtyping kit, protocol version 5 (National Institutes of Health, Bethesda, USA)}

The amplification conditions used for both rounds of PCR were the following:

- 3 cycles of 94°C for 1 minute, 55°C for 1 minute and 72°C for 1 minute,
- 32 cycles of 94°C for 15 seconds, 55°C for 45 seconds, 72°C for 1 minute,
- Final extension at 72°C for 5 minutes

***Gag* nested PCR Amplification conditions {*gag* HMA Subtyping Kit (Protocol Version 5, February 2002) (National Institutes of Health, Bethesda, USA)}**

First round:

- 1 cycle of 94°C for 2 minutes
- 35 cycles of 94°C for 30 seconds, 50°C for 30 seconds, 72°C for 90 seconds,
- Final extension at 72°C for 7 minutes

Second round:

- 1 cycle of 94°C for 2 minutes
- 35 cycles of 94°C for 30 seconds, 50°C for 45 seconds, 72°C for 1 minute,
- Final extension at 72°C for 7 minutes

PCR: Agarose gel detection of PCR products

1.4 grams of agarose powder (Whitehead Scientific, Brackenfell, SA) was added to 70 mls of 0.5 X Tris Borate EDTA (TBE) buffer. The solution was microwaved for 2 minutes after which it was cooled to 50-60^{0}C. 3µl of ethidium bromide was pipetted into the agarose solution. Extreme caution was used when ethidium bromide was used. The solution was then poured into a casting tray arranged with combs and left to set in a dark area until ready for use.

An electrophoresis tank of volume 2.5 L was filled with 500 mL of 0.5 X TBE. The agarose gel was placed in the tank with the top of the gel (the end with the wells) at the anode. 10 µl of a marker with the molecular weight of 700 kb was loaded onto the gel. For each sample, 10 µl of sample mixed with 2 µl of loading dye was pipetted into a well on the gel. The electrophoresis tank was then connected to the Biorad Powerpack (Biorad, California) and ran at 100 V for 2hrs.

Agarose gels were viewed under ultraviolet illumination and the presence of DNA was observed by an orange band. This was due to ethiduim bromide intercalating between DNA. Images of agarose gels were captured by the Biometra, BioDoc Analyse 1.0.

Preparation of a 5% polyacrlamide gel {*env* HMA Subtyping Kit, Protocol Version 5, National Institutes of Health, Bethesda, USA)}

A 50 ml mixture was prepared with 8.3 ml acrylamide stock, 5 ml 10X TBE and 36.7 ml water. Polymerization was initiated by mixing in 50 mg of ammonium persulphate and 33 µl of tetramethylethylenediamine. For *gag* HMA 14g urea was added to the mixture.

Assembly of Hoefer Apparatus

Constructing the gel sandwich and insert into caster (Figure 22)

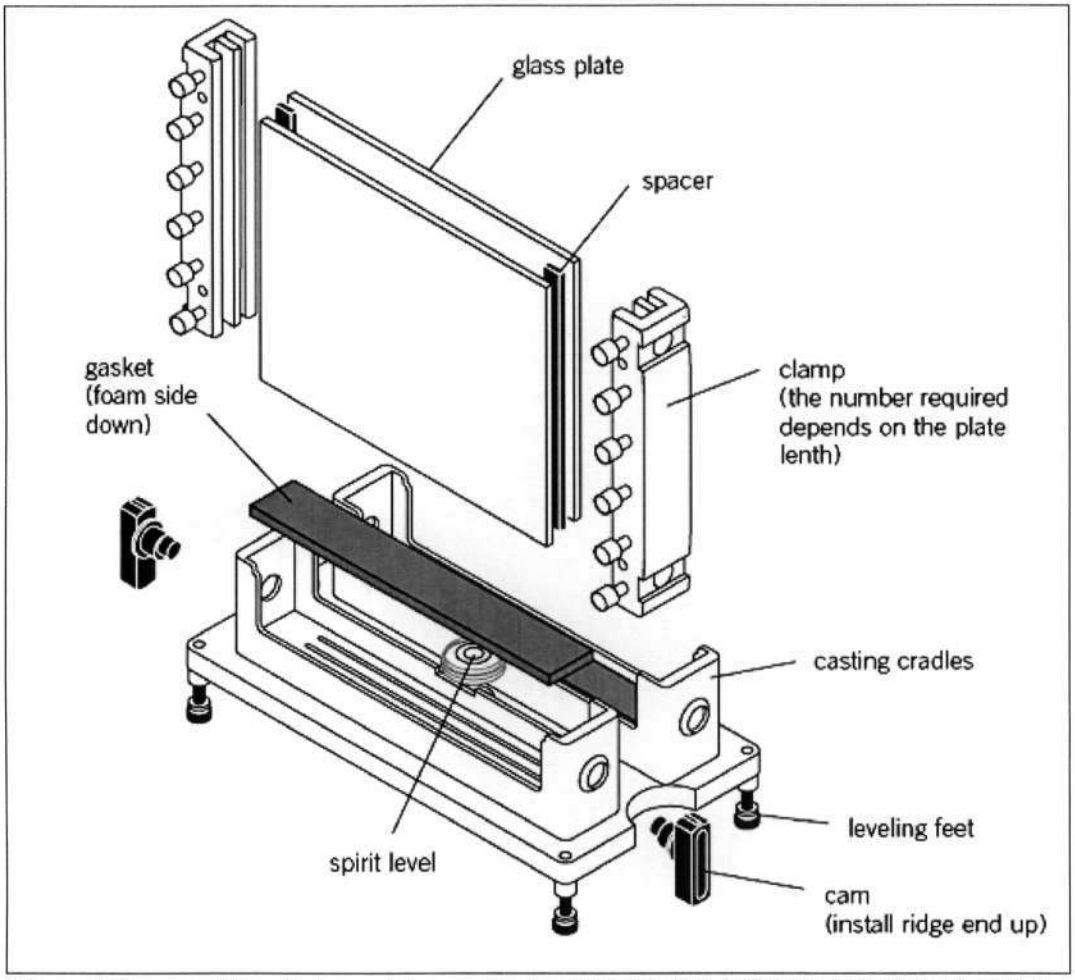


Figure 22: assembly of glass plates on caster (Hoefer SE 600/SE 660, SE600-IM/Rev. A0/06-04)

Place the spirit level into the center of the caster and adjust the leveling feet. Position gasket into grooves of the caster (make sure the foam side is down). Construct a gel sandwich as follows:

- Lay one glass plate flat on a surface.
- Place the spacers vertically opposite each other.
- Lay the other glass plate on top of the spacers.

Secure the sandwich by clamping the sides of the plates with spacers. Tighten screws when sandwich is aligned. Place assembled sandwich onto casting cradle. Insert a cam into the hole on either side of the casting tray. Seal the glass sandwich against the casting gasket by turning both cams as far as needed, usually 90^0C, 150^0C up to 180^0C. The cam action presses the plates down into the gasket which seals the bottom of the sandwich. The seal is complete once the glass edge appears darker and nearly transparent against the gasket. Do not turn the cam past this point.

Final assembly

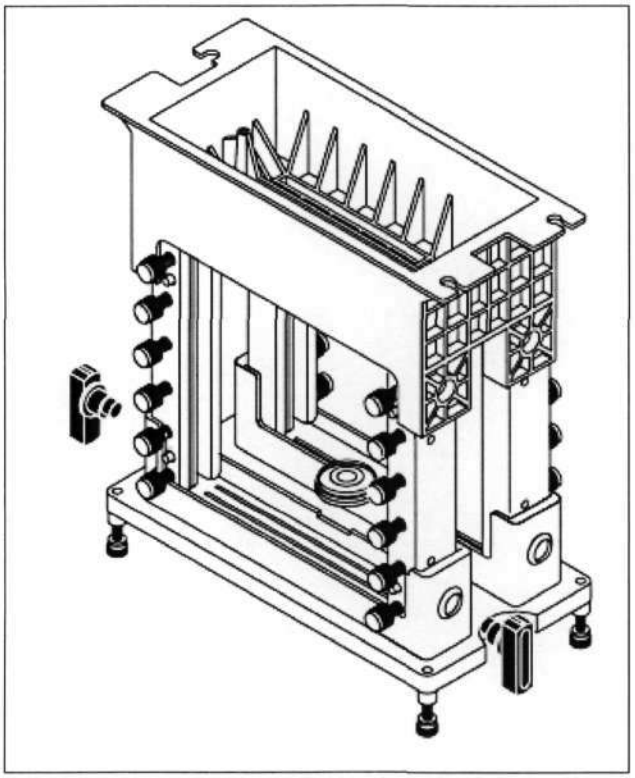


Figure 23: Positioning of upper chamber to gel sandwich (Hoefer SE 600/SE 660, SE600-IM/Rev. A0/06-04)

Turn the upper buffer chamber upside down and place gaskets into grooves provided (Figure 23). Place upper chamber on the top of the gel sandwich situated on the caster. Remove cams from caster and place in the holes on the upper buffer chamber. Turn the cams into place until the sandwich is firmly aligned with the upper buffer chamber. Pour 100ml of buffer into the chamber. Check that no buffer is leaking around the gasket.

Connect heat exchanger to cold water supply using tubing. Position heat exchanger into lower buffer chamber. The heat exchanger must always be in place for all runs since the lower electrode is integrated into the heat exchanger.

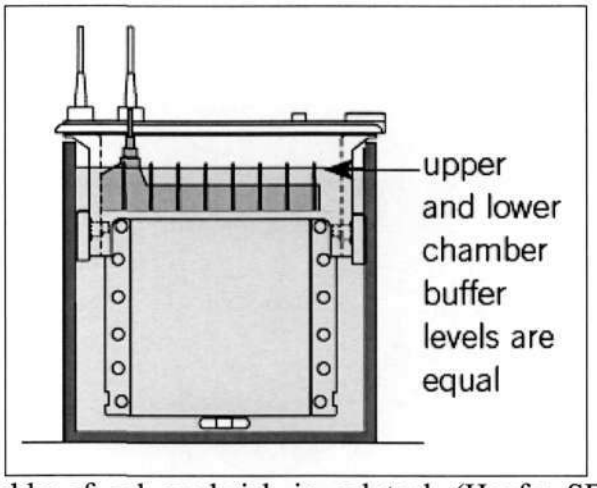


Figure 24: Final assembly of gel sandwich in gel tank (Hoefer SE 600/SE 660, SE600-IM/Rev. A0/06-04)

Fit the upper buffer chamber into the lower buffer chamber (Figure 24). Do this very carefully to avoid disturbing samples in the wells. Add 4.5-5 L of buffer to the lower chamber. The upper chamber requires 450 – 600mls of buffer – just enough to cover the upper chamber ribs

but should not contact the banana plug. Place the safely lid on the unit. Plug the leads into the jacks of a power supply.

Determining concentration of PCR products using a mass ladder

A 1.2% agarose (Whitehead Scientific, Brackenfell SA) gel was prepared to run the clean PCR product. Five microliters of each sample was mixed with 1µl loading dye and compared to a mass ladder (Invitrogen, Carlsbad USA) to determine the concentration. The concentrations per 5µl of product were calculated e.g. 100ng/5µl. For uniformity, all products were diluted to the same concentrations.

The Sequencing master mix per sample was prepared as listed in Table 9.

Table 9: Sequencing: Master Mix

Water	8.5µl
Ready reaction mix	4.0µl
5 X sequencing buffer	2.0µl
*primer (5 pmol/µl)	0.5µl
total volume	15µl

*primers were those used in the second PCR reaction for the *env* and *gag* regions ie ES7/ES8 and 1584/g17 respectively.

Clean-up of sequencing reaction

Tube clean up

Eighty microliters of 75% isopropanol was added to each PCR tube. After vortexing and incubating for 15 minutes, tubes were spun at full speed (14000Xg) for 20 minutes. The supernatant was aspirated. Two hundred microliters of freshly prepared ethanol (70%) was added to each tube (wash step). The tubes were vortexed, then centrifuged at full speed for 5 minutes. The supernatant was aspirated and the wash step was repeated. A quick spin was

performed and droplets of ethanol were aspirated. The sequencing products were dried at 50^0C for 5 minutes. Samples could be sequenced immediately or stored at -20^0C for no longer than two weeks.

Plate clean up

Prepare a mix of 3M sodium acetate and absolute ethanol in a reagent trough using the following volumes per reaction: 2µl sodium acetate and 50µl absolute ethanol. Prepare 10 reactions extra when using a multi-channel pipette. Pipette 55µl of mix into each well. Cover plate with a sticky plate cover. Vortex for 5 seconds. Centrifuge for 20 minutes at 3000 x g. Prepare a 70% ethanol solution: 150µl 70% ethanol is used per reaction. Prepare 10 extra reactions and prepare directly into the reagent trough and keep in freezer until ready for use. Fold 2 pieces of tissue into rectangles approximately the same size as the plates. After centrifugation, turn plate onto the rectangular piece of tissue paper. Place the other piece onto the base of the centrifuge bucket. Spin for 5 minutes at 150 x g. Immediately add 150µl of 70% ethanol to each well (do not let the plate dry as this will cause the presence of dye-blobs in the sequence). Cover plate with the sticky plate cover. Centrifuge for 5 minutes at 3000 x g. Fold tissue into rectangular pieces approximately the size of the plate. Turn plate onto the prepared tissue paper. Place the other piece onto the base of the centrifuge bucket. Spin for 1 minute at 150 x g. Air dry plate in the dark for approximately 5 minutes. Samples can be stored at -20^0C.

Cloning (Package Insert: TOPO TA Cloning® Kit for Sequencing, Version N, July 2004)

Ligation step – the cloning reaction

The following were used to prepare a cloning reaction (per sample):

a. Mix the following reagents together with the respective volumes:

Fresh PCR product	2µl
Salt solution	1µl
Sterile water	2µl
TOPO® vector	1µl

b. mix reaction gently (do not pipette up and down) by swirling pipette in reaction

c. incubate for 5 minutes at room temperature (22 to 23^{0}C)

d. place the reaction on ice or store overnight at -20^{0}C

Transformation using DH5a - T1 Chemically Competent Cells

a. Add 2 µl of the TOPO Cloning reaction from the Ligation step into a vial of DH5a Chemically competent *E.coli* and mix gently. Do not mix by pipetting up and down.
b. Incubate on ice for 5 minutes
c. Check temperature of water bath using a thermometer – ensure that it is at 42^{0}C. Heat-shock the cells for 30 seconds at 42^{0}C without shaking.
d. Immediately transfer the tubes to ice.
e. Add 250 µl S.O.C. medium (room temperature) to tubes containing cells
f. Cap the tube tightly and shake the tube horizontally at 37^{0}C for 1 hour.

g. Spread 10 to 50 µl from each transformation on a prewarmed selective plate and incubate overnight at 37^0C. Recommendation: two volumes should be plated ie. 25 and 50 µl to ensure well-spaced colonies. Use Aseptic technique when spreading the culture.

a. draw a grid on the back of an LB agar plate and number each block from 1 to16. This is shown in Figure 25. Work in the region of the flame.

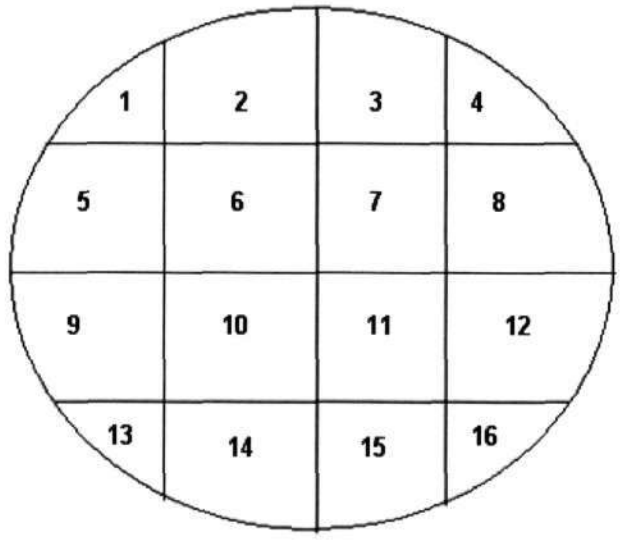


Figure 25: representation of grid on agar plate

b. select white colonies – especially those in the region of blue colonies and those without satalites.
c. pick out a colony at a time using a picker and touch the designated position on the LB agar master plate eg. The first colony picked will touch the centre of block 1.
d. place the end of the picker in a PCR tube with 50 µl water. Eg. The first colony picker will be placed in a PCR tube labelled 1. Discard picker.
e. Once plating with colonies is completed leave plate in incubator set at 37^0C.
f. Vortex PCR tubes with water and plasmids.
g. Boil for 10 minutes at 95^0C. (use thermocycler)
h. Centrifuge at maximum speed for 2 minutes.
i. Add 10 µl sample into master mix.

j. Amplify using programme *gag* PCR on Perkin Elmer 3100.
k. Detect products using agarose gel (1%). A 460bp product should be viewed.
l. Once cultures have grown on master plates store plates in fridge.

Preparation of LB (Luria-Bertani) Medium and plates

Materials

1.0% Tryptone or Casein peptone (Merck, Germany)
0.5% Yeast Extract
1.0% NaCl (Merck, Germany)
Deionised water
Agar

Methods

1. to make up 1L, dissolve 10g tryptone , 5g yeast extract, and 10g NaCl in 950 ml deionised water.
2. stir using magnetic stirrer and flee
3. adjust the pH of the solution to 7.0 with NaOH and bring volume up to 1L.
4. Autoclave on liquid cycle for 20 minutes at 15 psi or 121 kph. Allow solution to cool to 55^{0}C and add antibiotic (Ampicillin, Kanamycin etc).
5. Store at room temperature or at 4^{0}C.

Preparation of LB agar plates

1. prepare LB medium as above, but add 15g/L agar before autoclaving
2. autoclave on liquid cycle for 20 minutes at 15 psi.
3. after autoclaving, cool to 55^{0}C, add antibiotic (50-100µg/ml of ampicillin or 50 µg/ml kanamycin), and pour into 10 cm plates.
4. leave at room temperature to harden.
5. Optional: Spread Xgal (100 µg/ml) over hardened LB plate.
6. store at 4^{0}C in the dark.

Appendix II

The following were taken from: HIV-1 *env* subtyping kit, protocol version 5 (National Institutes of Health, Bethesda, USA)

EDTA concentrations

Stock EDTA = 500g with molecular weight of 373.7

0.5M EDTA

Dissolve 186.1g of EDTA in 1 L distilled water.

10 X heteroduplex annealing buffer

For a total volume of 100 mls add:

20 mls of a 5M NaCl solution,

1.575g of Tris HCl (500g) (Sigma, Steinheim Germany),

4 ml of 0.5M EDTA (Merck, germany)

Fill up to 100mls with deionised H_2O.

10 X Tris Borate EDTA buffer

For 1 L 10 X stock dissolve in deionised water:

106.5 g of Tris base

55g of Boric acid (Merck chemicals, Gauteng)

40 ml of 0.5 M EDTA (20mM EDTA)

10mM dNTP solution

a. For a total volume of 50µl, add 5 µl of each nucleotide (100mM) (Amersham, UK) into an eppendorf.

b. Add 30 µl of PCR grade water to the solution.

5 X Tris Borate EDTA buffer

Weigh out the following for a 1L solution:

Trizma Base - 60.7g

Boric Acid - 25.66g

EDTA - 1.86g

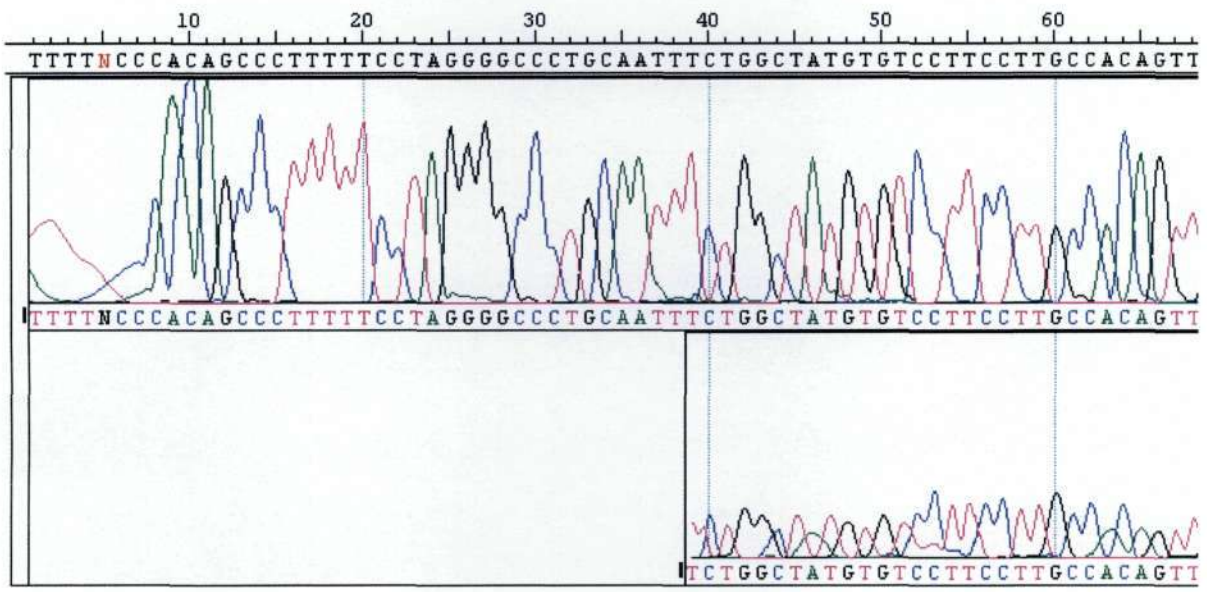
Make up to 1L volume with distilled water.

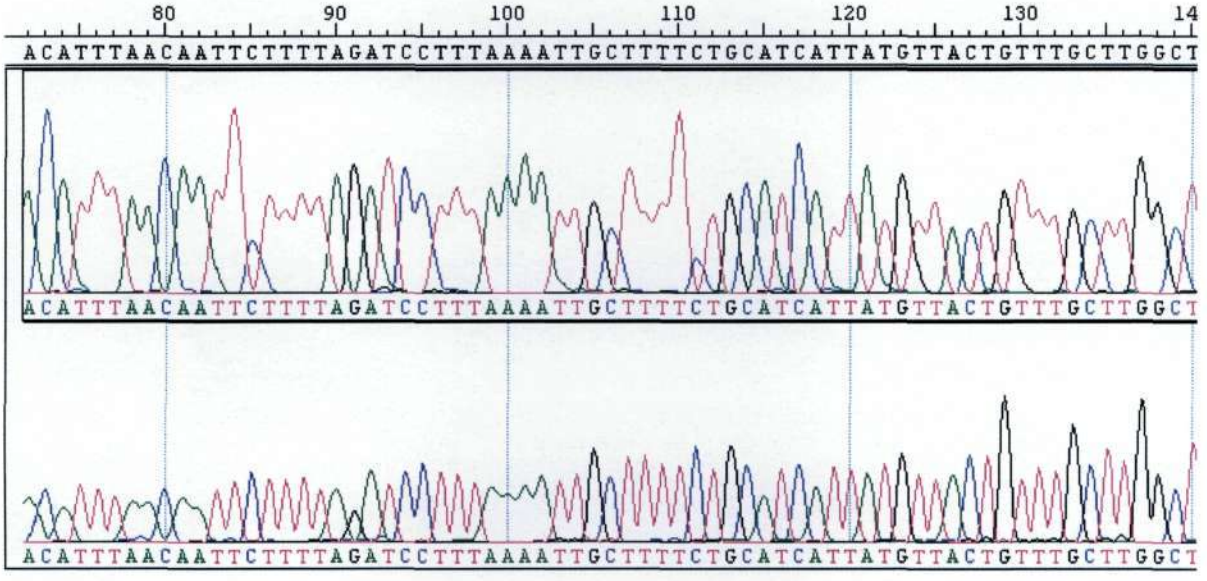
0.5 X Tris Borate EDTA buffer

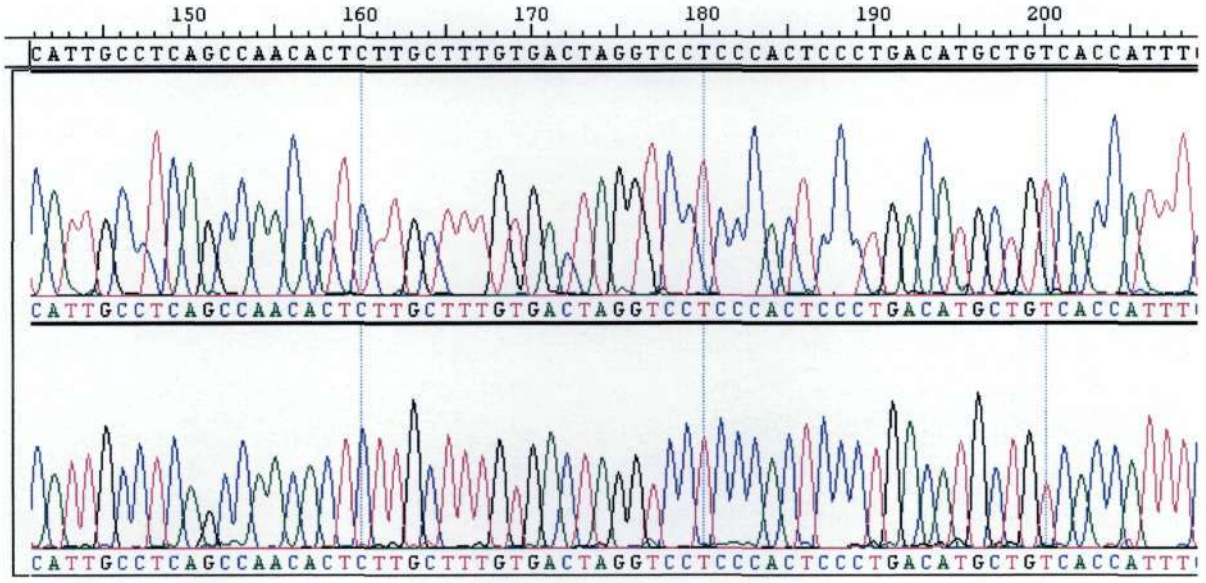
Dilute 100mls of 5 X TBE with 900mls of dH_2O.

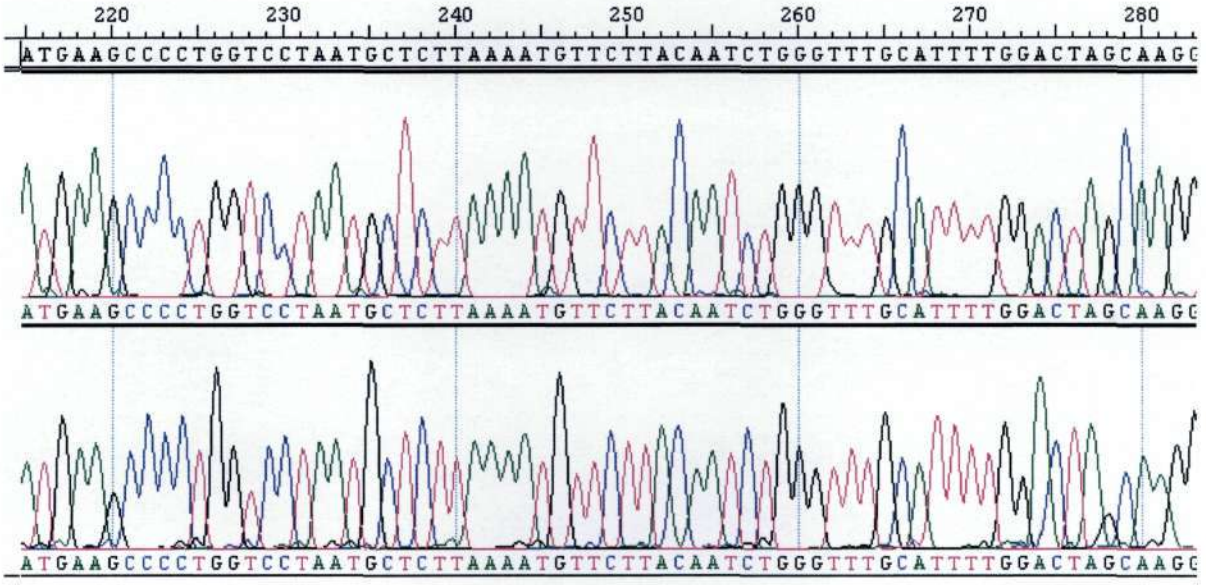
Electropherogram

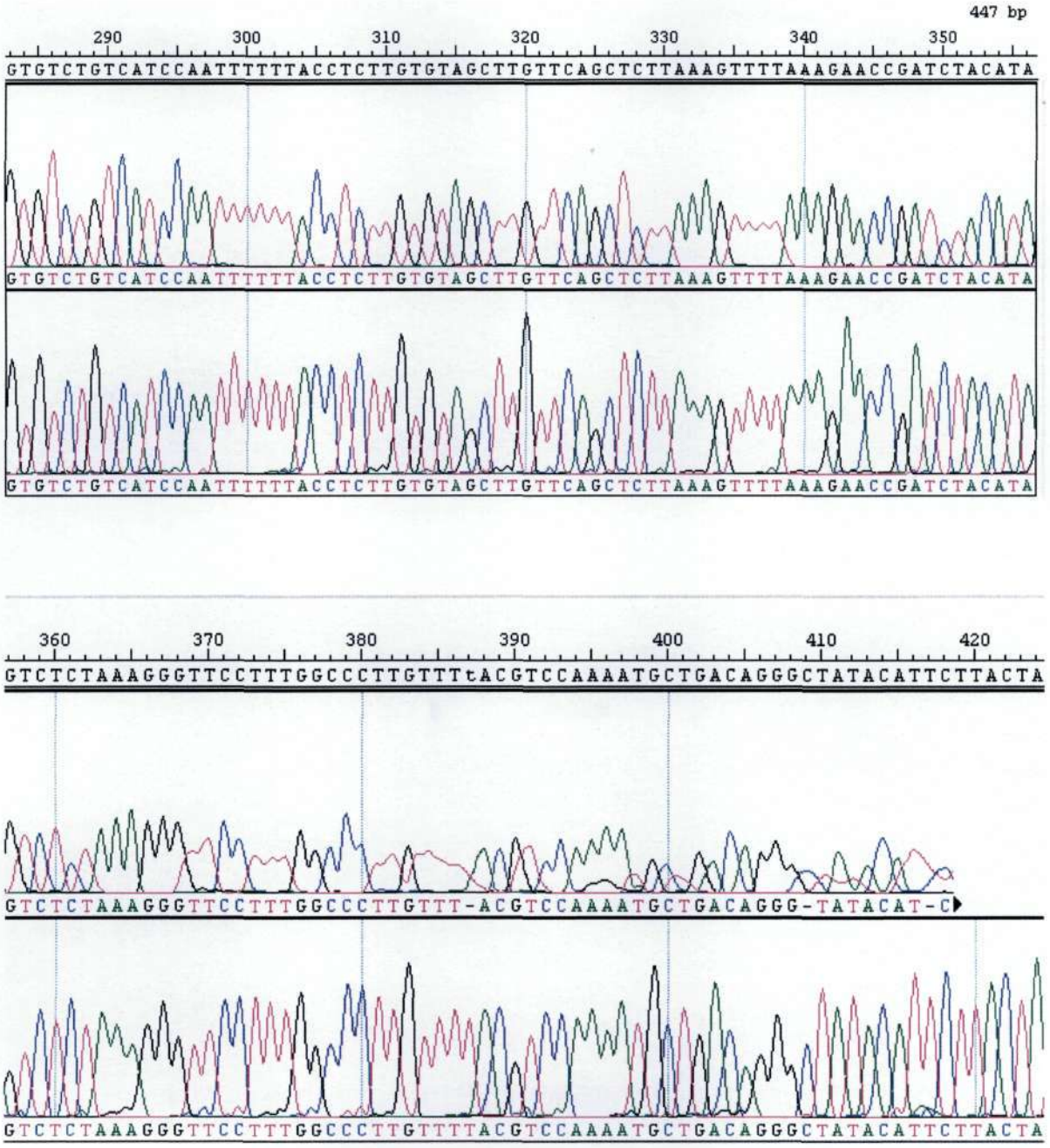
Clone92_1






150
160
170
180
190
200
CATTGCCTCAGCCAACACTCTTGCTTTGTGACTAGGTCCTCCCACTCCCTGACATGCTGTCACCATTT
CATTGCCTCAGCCAACACTCTTGCTTTGTGACTAGGTCCTCCCACTCCCTGACATGCTGTCACCATTT
CATTGCCTCAGCCAACACTCTTGCTTTGTGACTAGGTCCTCCCACTCCCTGACATGCTGTCACCATTT


220
230
240
250
260
270
280
ATGAAGCCCCTGGTCCTAATGCTCTTAAAATGTTCTTACAATCTGGGTTTGCATTTTGGACTAGCAAGC
ATGAAGCCCCTGGTCCTAATGCTCTTAAAATGTTCTTACAATCTGGGTTTGCATTTTGGACTAGCAAGC
ATGAAGCCCCTGGTCCTAATGCTCTTAAAATGTTCTTACAATCTGGGTTTGCATTTTGGACTAGCAAGC


447 bp
290
300
310
320
330
340
350
GTGTCTGTCATCCAATTTTTTACCTCTTGTGTAGCTTGTTCAGCTCTTAAAGTTTTAAAGAACCGATCTACATA
GTGTCTGTCATCCAATTTTTTACCTCTTGTGTAGCTTGTTCAGCTCTTAAAGTTTTAAAGAACCGATCTACATA
GTGTCTGTCATCCAATTTTTTACCTCTTGTGTAGCTTGTTCAGCTCTTAAAGTTTTAAAGAACCGATCTACATA
360
370
380
390
400
410
420
GTCTCTAAAGGGTTCCTTTGGCCCTTGTTTtACGTCCAAAATGCTGACAGGGCTATACATTCTTACTA
GTCTCTAAAGGGTTCCTTTGGCCCTTGTTT-ACGTCCAAAATGCTGACAGGG-TATACAT-C
GTCTCTAAAGGGTTCCTTTGGCCCTTGTTTTACGTCCAAAATGCTGACAGGGCTATACATTCTTACTA

Appendix III

Laboratory safety

Ethidium bromide

Ethidium bromide is a highly mutagenic agent and must be handled as carefully as possible. It can be absorbed through the skin so direct contact must be avoided by the use of chemically resistant gloves in conjunction with a laboratory coat, and closed toe-shoes. Ethidium bromide is an irritant to the skin, eyes, mouth and upper respiratory tract. It must be stored away from oxidising agents in a cool, dry place. Traces of this mutagen are found in agarose and polyacylamide gels, TBE buffers, laboratory gloves, equipment and debris.

Acrylamide

Acrylamide is a toxic substance and probable mutagen. Its most common routes of exposure include: the skin, inhalation and ingestion. Acrylamide used in the laboratory can be in two forms, a powder or liquid. The powder form is extremely dangerous as the acrylamide dust can become airborne and enter the respiratory tract. Pre-mix acrylamide or ready-made gels are much safer for use. When using the acrylamide powder work in a fume hood, and wear gloves, a laboratory coat, goggles and a face shield. Even when working with the polymerised form of acrylamide and electrophoresis equipment exposed to acrylamide protective laboratory gear must be worn. Acrylamide must be stored in tightly closed containers to prevent acrylamide entering the atmosphere. It does not form stable aerosols. Acrylamide should also be stored in a cool place away from oxidising agents, reducing materials, bases, metals, UV light and other contaminants.

Waste bins

Polyacylamide and agarose gels are discarded in medical waste boxes for incineration.

Charcoal filtration

TBE exposed to ethidium bromide is mixed with activated charcoal eg 1g activated charcoal per 1L TBE and left to stand for 24 hours. The solution is then filtered with a Whatman and then poured into a medical waste drum for incineration.

Gloves, equipment and debris

Gloves and paper towel exposed to ethidium bromide are placed in medical waste boxes for incineration. Casting trays, combs etc. are washed with bleach, rinsed with water and dried with 70% ethanol for reuse.

References

Agwale SM, Robbins KE, Odma L, Saekhou A, Zeh C, Edubio A, Njoku OM, Sani-Gwarzo N, Gboun MF, Gao F, Reitz M, Hone D, Folks TM, Pieniazek D, Wambebe C, Kalish ML. (2001) Development of an *env* gp41-Based Heteroduplex Mobility Assay for Rapid Human Immunodeficiency Virus Type 1 Subtyping. *Journal of Clinical Microbiology*, June: 2110–2114

Allen T, Altfeld M. HIV-1 superinfection. (2003) *Journal of Allergy and Clinical Immunology*, 112: 829-835

Altfeld M, Allen TM, Yu XG, Johnston MN, Agrawal D, Korber BT, Montefiori DC, O'Connor DH, Davis BT, Lee PK, Maier EL, Harlow J, Goulder PJR, Brander C, Rosenburg ES, Walker BD. (2002) HIV-1 superinfection despite broad CD8_ T-cell responses containing replication of the primary virus. *Nature*, 420(6914): 434–439

Apetrei C, Descamps D, Collin G, Loussert-Ajaka I, Damond F, Duca M, Simon F, Brun-Vezinet F. (1998) Human immunodeficiency virus type 1 subtype F reverse transcriptase sequence and drug susceptibility. *Journal of Virology*, 72: 3534-3538

Archer RH, Dykes C, Gerondelis P, Lloyd A, Fay P, Reichman RC, Bambara RA, Demeter LM. (2000) Mutants of human immunodeficiency virus type 1 (HIV-1) reverse transcriptase resistant to nonnucleoside reverse transcriptase inhibitors demonstrate altered rates of RNase H cleavage that correlate with HIV-1 replication fitness in cell culture. *Journal of Virology*, 74:8390-401

Arens M. (1999) Methods for Subtyping and Molecular Comparison of Human Viral Genomes. *Clinical Microbiology Reviews*, October: 612–626

Ball, S. C., A. Abraha, K. R. Collins, A. J. Marozsan, H. Baird, M. E.Quinones-Mateu, A. Penn-Nicholson, M. Murray, N. Richard, M. Lobritz, P. A. Zimmerman, T. Kawamura, A. Blauvelt, E. J. Arts. (2003) Comparing the *ex vivo* fitness of CCR5-tropic human immunodeficiency virus type 1 isolates of subtypes B and C. *Journal of Virology*, 77:1021–1038

Barre-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vezinet-Brun F, Rouzioux C, Rozenbaum, Montagnier L. (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficieny syndrome (AIDS). *Science,* 1983: 868-871

Barreto CC, Nishyia A, Arau´jo LV, Ferreira JE, Busch MP, Sabino EC. (2006) Trends in Antiretroviral Drug Resistance and Clade Distributions Among HIV-1–Infected Blood Donors in Sao Paulo, Brazil. *Journal Acquired Immune Deficiency Syndrome*, 41: 338-341

Battula N and Loeb LA. (1976) On the fidelity of DNA Replication: Lack of exodeoxyribonuclease activity and error-correcting function in avian myeloblastosis virus DNA polymerase. *The Journal of Biological Chemistry*, 251(4): 982-986

Bierwolf D. (1983) The possible etiological role of retroviruses in human tumors. *Arch Geschwulstforsch,* 53(4): 303-13

Bikandou B, Takehisa J, Mboudjeka I, Ido E, Kuwata T, Miyazaki Y, Moriyama H, Harada Y, Taniguchi Y, Ichimura H, Ikeda M, Ndolo PJ, Nzoukoudi M-Y, M'Vouenze R, M'Pandi M, Parra HJ, M'Pele P, Hayami M. (2000) Genetic subtypes of HIV type 2 in Republic of Congo. *AIDS Research and Human Retroviruses*, 16: 613-619

Blackard JT, Cohen DE, Mayer KH. (2002) Human immunodeficiency virus superinfection and recombination: Current state of knowledge and potential clinical consequences. *Clinical Infectious Diseases*, 34 (8): 1108–1114

Bocket L, Cheret A, Deuffic-Burban S, Choisy P, Gerard Y, de la Tribonniere X, Viget N, Ajana F, Goffard A, Barin F, Mouton Y, Yazdanpanah Y. (2005) Impact of human immunodeficiency virus type 1 subtype on first-line antiretroviral therapy effectiveness. *Antiviral Therapy*, 10: 247–254

Boom R, Sol CJA, Salimans MMM, Jansen CL, Wertheim-van Dillen PME, van der Noordaa J. (1990) Rapid and Simple Method for Purification of Nucleic Acids. *Journal of Clinical Microbiology*, 28 (3): 495-503

Boulerice F, Li XG, Lvovich A, Wainberg MA.(1991) Recovery of infectious human immunodeficiency virus type 1 after fusion of defectively infected clones of U-937 cells. *Journal of Virology*, 65:5589–5592

Bredell H, Hunt G, Morgan B, Tiemessen CT, Martin DJ and Morris L. (2000) Identification of HIV Type 1 Intersubtype Recombinants in South Africa Using *env* and *gag* Heteroduplex Mobility Assays. *Aids Research and Human Retroviruses,* 16(5): 493–497

Brennan CA, Yamaguchi J, Vallari AS, Hickman RK, Devare SG. (1997) Genetic variation in human immunodeficiency virus type 2: identification of a unique variant from human plasma. *AIDS Research and Human Retroviruses, 13*: 401-4

Brenner B, Routy J-P, Quan Y, Moisi D, Oliveira M, Turner D, Wainberg MA, and co-investigators of the Quebec Primary Infection Study. (2004) Persistence of multidrug-resistant HIV-1 in primary infection leading to superinfection. *AIDS*, 18:1653–1660

Brenner B, Turner D, Oliveira M, Moisi D, Detorio M, Carobene M, Marlink RG, Schapiro J, Roger M, Wainberg MA. (2003) A V106M mutation in HIV-1 clade C viruses exposed to efavirenz confers cross-resistance to non-nucleoside reverse transcriptase inhibitors. *AIDS*, 17: 1–5

Brodine SK, Mascola JR, Weiss PJ, Ito SI, Porter KR, Artenstein AW, Garland FC, McCutchan FE, Burke DS. (1995) Detection of diverse HIV-1 genetic subtypes in the USA. *Lancet*, 346:1198-1199

Buiser RG, Bambara RA, Fay PJ. (1993) Pausing by retroviral DNA polymerases promotes strand transfer from internal regions of RNA donor templates to homopolymeric acceptor templates. *Biochimica et Biophysica Acta,* October 19,1216(1): 20–30

Carr JK, Ávila M, Carrillo MG, Salomon H, Hierholzer J, Watanaveeradej V, Pando MA, Negrete M, Russell KL, Sanchez J, Birx DL, Andrade R, Vinoles J, McCutchan FE. (2001) Diverse BF recombinants have spread widely since the introduction of HIV-1 into South America. *AIDS*, 15: F41–F47

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

Cavallo G and Cavallo R. (1986) Retroviruses: current classification system. *Journal of Bacteriology and Viral Immunology,* 79(7-12): 288-94

Chanbancherd P, Brown AE, Trichavaroj R, Tienamporn P, Puthakird P, Limpairojn N, Van Cott TC, de Souza MS. (1999) Application of Dried Blood Spot Specimens for Serologic Subtyping of Human Immunodeficiency Virus Type 1 in Thailand. *Journal of Clinical Microbiology*, 37(3): 804-806

Chen YM, Lee CM, Lin RY, Chang HJ. (1998) Molecular epidemiology and trends of HIV-1 subtypes in Taiwan. *Journal of Acquired Immune Deficiency Syndrome and Human Retrovirology*, 19: 393-402

Chen Z, Luckay A, Sodora DL, Telfer P, Reed P, Gettie A, Kanu JM, Sadek RF, Yee J, Ho DD, Zhang A, Marx PA. (1997) Human immunodeficiency virus type 2 (HIV-2) seroprevalence and characterization of a distinct HIV-2 genetic subtype from the natural range of simian immunodeficiency virus infected sooty mangabeys. *Journal of Virology*, 71: 3953-60

Clavel F, Guetard D, Brun-Vezinet F, Chamaret S, Rey M-A, Santos-Ferreira MO, Laurent AG, Dauguet C, Katlama C, Rouzioux C, Klatzmann D, Champalimaud JL, Montagnier L. (1986) Isolation of a new human retrovirus from West African patients with AIDS. *Science,* 233:343-346

Coffin JM. (1979) Structure, replication, and recombination of retrovirus genomes: some unifying hypotheses. *Journal of General Virology*, 42(1): 1–26

Cooper DA Imrie AA, Penny R. (1987) Antibody response to human immunodeficiency virus after primary infection. *Journal of Infectious Diseases,* 155:1113-8

Curlin ME, Gottlieb GS, Hawes SE, Sow PS, Ndoye I, Critchlow CW, Kiviat NB, Mullins JI. (2004): No evidence for recombination between HIV type 1 and HIV type 2 within the envelope region in dually seropositive individuals from Senegal. *AIDS Research and Human Retroviruses*, 20(9): 958–963

Damond F, Apetrei C, Robertson DL, Souquiere S, Lepretre A, Matheron S, Plantier JC, Brun-Vezinet F, Simon F. (2001) Variability of human immunodeficiency virus type 2 (HIV-2) infecting patients living in France. *Virology*, *280*: 19-30

Damond F, Worobey M, Campa P, Farfara I, Colin G, Matheron S, Brun-Vezinet F, Robertson DL, Simon F. (2004) Identification of a highly divergent HIV type 2 and proposal for a change in HIV type 2 classification. *AIDS Research and Human Retroviruses*, *20*: 666-72

de Oliveira T, Deforche K, Cassol S, Salminen M, Paraskevis D, Seebregts C, Snoeck J, van Rensburg EJ, Wensing AM, van de Vijver DA, Boucher CA, Camacho R, Vandamme AM. (2005) Automated genotyping system for analysis of HIV-1 and other microbial sequences. *Bioinformatics,* 21(19): 3797-800

de Oliveira T, Engelbrecht S, van Rensburg EJ, Gordon M, Bishop K, zur Megede J, Barnett SW, Cassol1 S. (2003) Variability at Human Immunodeficiency Virus Type 1 Subtype C Protease Cleavage Sites: an Indication of Viral Fitness? *Journal of Virology*, 77 (17): 9422–9430

Delwart EL, Sheppard HW, Walker BD, Goudsmit J, Mullins JI. (1994) Human Immunodeficiency Virus Type 1 Evolution In Vivo Tracked by DNA Heteroduplex Mobility Assays. *Journal of Virology*, 68 (10): 6672-6683

Delwart EL, Shpaer EG, Louwagie J, McCutchan FE, Grez M, Rubsamen-Waigmann H, Mullins JI. (1993) Genetic relationships determined by a DNA heteroduplex mobility assay: analysis of HIV-1 *env* genes. *Science*, 262(5137): 1257–1261

Department of Health. (2006) National HIV and syphilis antenatal seroprevalence survey in South Africa 2005

Department of Health and Human Services (DHHS). Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents. October 10, 2006. Available online at aidsinfo.nih.gov/Guidelines/GuidelineDetail.

Descamps D, Collin G, Loussert-Ajaka I, Saragosti S, Simon F, Brun-Vezinet F. (1995) HIV-1 group O sensitivity to antiretroviral agents. *AIDS*, 9: 977-978

Devereux HL, Youle M, Johnson MA, Loveday C. (1999) Rapid decline in detectability of HIV-1 drug resistance mutations after stopping therapy. *AIDS*, 13(18):F123-7

Dowling WE, Kim B, Mason CJ, Wasunna KM, Alam U, Elson L, Birx DL, Robb ML, McCutchan FE, Carr JK. (2002) Forty-one near full-length HIV-1 sequences from Kenya reveal an epidemic of subtype A and A-containing recombinants. *AIDS*, 16:1809–1820

Easterbrook PJ, Smith M, Geretti AM, *et al.* (2002) Impact of HIV-1 viral subtype on disease progression. XIV International AIDS Conference. July 7-12, Barcelona

Eshleman SH, Becker-Pergola G, Deseyve M, Guay LA, Mracna M, Fleming T, Cunningham S, Musoke P, Mmiro F, Jackson JB. (2001) Impact of HIV-1 subtype on women receiving single dose nevirapine prophylaxis to prevent HIV-1 vertical transmission (HIV network for prevention trials 012 study). *Journal of Infectious Diseases*, 184: 914–917

Eshleman SH, Hackett J, Swanson P , Cunningham SP, Drews B, Brennan C, Devare SG, Zekeng L, Kaptue' L, Marlowe N. (2004) Performance of the Celera Diagnostics ViroSeq HIV-1 Genotyping System for Sequence-Based Analysis of Diverse Human Immunodeficiency Virus Type 1 Strains. *Journal of Clinical Microbiology,* June, 2711–2717

Frater AJ, Dunn DT, Beardall AJ, Ariyoshi K, Clarke JR, McClure MO, Weber JN. (2002) Comparative response of African HIV-1-infected individuals to highly active antiretroviral therapy. *AIDS*,16: 1139–1146

Freed EO and Martin MA. (2001) HIVs and their replication. In Field's Virology, Editors-in-chief: Knipe DM, Howley PM (eds.) Philadelphia, PA, USA. Lippincott Williams and Wilkins. Volume 2, 4th edition, 1971-2041

Gadelha SR, Shindo N, Cruz JNM, Morgado MG, Galvao-Castro B. (2003) Molecular Epidemiology of Human Immunodeficiency Virus-1 in the State of Ceará, Northeast, Brazil. *Memórias do Instituto Oswaldo Cruz,* Rio de Janeiro, *98*(4): 461-464

Gallo RC, Sarin PS, Gelmann EP, Robert-Guroff M, Richardson E, Kalyanaraman VS, Mann D, Sidhu GD, Stahl RE, Zolla-Pazner S, Leibowitch J, Popovic M. (1983) Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science, New Series*, 220(4599):865-7

Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, Cummins LB, Arthur LO, Peeters M, Shaw GM, Sharp PM, Hahn BH. (1999) Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature*, 397: 436-41

Gao F, Robertson DL, Carruthers CD, Li Y, Bailes E, Kostrikis LG, Salminen MO, Bibollet-Ruche F, Peeters M, Ho DD, Shaw GM, Sharp PM, Hahn BH. (1998) An isolate of human immunodeficiency virus type 1 originally classified as subtype I represents a complex mosaic comprising three different group M subtypes (A, G and I). *Journal of Virology*, 72: 10234-10241

Gao F, Robertson DL, Morrison SG, Hui H, Craig S, Decker J, Fultz PN, Girard M, Shaw GM, Hahn BH, Sharp PM. (1996) The heterosexual human immunodeficiency virus type 1 epidemic in Thailand is caused by an intersubtype (A/E) recombinant of African origin. *Journal of Virology,* 70:7013–7029

Gao F, Yue L, Robertson DL, Hill SC, Hui H, Biggar RJ, Neequaye AE, Whelan TM, Ho DD, Shaw GM, Sharp PM, Hahn BH. (1994) Genetic diversity of human immunodeficiency virus type 2: evidence for distinct sequence subtypes with differences in virus biology. *Journal of Virology, 68*: 7433-47

Gerondelis P, Archer RH, Palaniappan C, Reichman RC, Fay PJ, Bambara RA, Demeter LM. (1999) The P236L delavirdine-resistant human immunodeficiency virus type 1 mutant is replication defective and demonstrates alterations in both RNA 5,-end- and DNA 3,-end-directed RNase H activities. *Journal of Virology*, 73:5803-13

Goff SP. (2001) Retroviridae: The Retroviruses and their Replication. In Fields Virology. Knippe DM, Howley PM, Griffin DE (eds.). Philadelphia, PA, USA: Lippincott Williams & Wilkins, 1871–1939

Goodrich DW, Duesberg PH. (1990) Retroviral recombination during reverse transcription. *Proceedings of the National Academy of Sciences, USA*, 87:2050-2056

Gottlieb GS, Nickle DC, Jensen MA , Wong KG, Grobler J, Li F, Liu S-L, Rademeyer C, Learn GH, Karim SSA, Williamson C, Corey L, Margolick JB, Mullins JI. (2004) Dual HIV-1 infection associated with rapid disease progression. *Lancet*, 363(9409): 619 – 622

Grobler J, Gray CM, Rademeyer C, Seoighe C, Ranjee G, Karim SA, Morris L, Williamson C. (2004) Incidence of HIV-1 dual infection and its association with increased viral load set point in a cohort of HIV-1 subtype C-infected female sex workers. *Journal of Infectious Diseases*, 190(7):1355–1359

Grossman Z, Istomin V, Averbuch D, Lorber M, Risenberg K, Levi I, Chowers M, Burke M, Yaacov BN, Schapiro JM. (2004) Israel AIDS Multi-Center Study Group: Genetic variation at NNRTI resistance-associated positions in patients infected with HIV- subtype C. *AIDS*, 18: 909–915

Grossman Z, Paxinos EE, Averbuch D, Maayan S, Parkin NT, Engelhard D, Lorber M, Istomin V, Shakad Y, Mendelson E, Ram D, Petropoulos CJ, Schapiro JM. (2004) Mutation D30N is not preferentially selected by human immunodeficiency virus type 1 subtype C in the development of resistance to nelfinavir. *Antimicrobial Agents Chemotherapy*, 48:2159-2165

Gurtler L. (1996) Difficulties and strategies of HIV diagnosis. *Lancet,* 348: 176–79

Hacket J, Brennan C, Hickman K, Vallari A, Swanson P, Rayfield M, Schable C, Subbarao S, Pieniazek D, Sullivan P, Schochetman G, Devare S. (1997) Molecular and Serologic characteristion of a HIV-1 group O isolate identified in the United States. 4th Conference on Retroviruses and Opportunistic Infections. Washington, DC – January 22-26

Hemelaar J; Gouws E ; Ghys PD; Osmanov S. (2006) Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS*, 20(16):W13-W23

Heyndrickx L, Janssens W, Zekeng L, Musonda R, Anagonou S'V, Van der Auwera G, Coppens S, Vereecken K, de Witte K, van Rampelbergh R, Kahindo M, Morison L, McCutchan FE, Carr JK, Albert J, Essex M, Goudsmit J, Asjo B, Salminen M, Buve A, Study Group on Heterogeneity of HIV epidemics in African Cities, van der Groen G. Simplified Strategy for Detection of Recombinant Human Immunodeficiency Virus Type 1 Group M Isolates by *gag/env* Heteroduplex Mobility Assay. *Journal of Virology*, 74(1), 363–370

Herbinger K-H, Gerhardt M, Piyasirisilp S, Mloka D, Arroyo MA, Hoffmann O, Maboko L, Birx DL, Mmbando D, McCutchan FE, Hoelscher M. (2006) Frequency of HIV Type 1 Dual Infection and HIV Diversity: Analysis of Low- and High-Risk Populations in Mbeya Region, Tanzania. *Aids Research and Human Retroviruses*, 22(7), 599–606

Hirsch MS, Conway B, D'Aquila RT *et al.* Antiretroviral Drug Resistance Testing in Adults with HIV Infection: Implications for Clinical Management. *Journal of the American Medical Association.* 1998; 1984-91

Hoelscher M, Dowling WE , Sanders-Buell E, Carr JK , Harris ME , Thomschke A, Robb ML ,Birx DL, McCutchan FE. (2002) Detection of HIV-1 subtypes, recombinants,and dual infections in east Africa by a multi-region hybridization assay. *AIDS*, 16:2055-2064

Howard TM, Rasheed S. (1996) Genomic structure and nucleotide sequence analysis of a new HIV type 1 subtype A strain from Nigeria. *AIDS Research and Human Retroviruses*, 12: 1413-1425

Hu W-S, Temin HM. (1990) Genetic consequences of packaging two RNA genomes in one retroviral particle: Pseudodiploidy and high rate of genetic recombination. *Proceedings of the National Academy of Sciences, USA* 87: 1556-1560

Huang DD, Glesler TA, Bremer JW. (2003) Sequence characterization of the protease and partial reverse transcriptase proteins of the NED panel, and international HIV type 1 subtype reference and standards panel. *AIDS Research and Human Retroviruses*, 19: 321-328

Huang H, Chopra R, Verdine GL, Harrison SC. (1998) Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science,* 282:1669-1675

Iweala OI. (2004) HIV diagnostic tests: an overview. *Contraception*, 70:141–147

Jameel S, Zafrullah M, Ahmed M, Kapoor GS, Sehgal SA. (1995) Genetic analysis of HIV-1 from Punjab, India reveals the presence of multiple variants. *AIDS*, 9:685-690

Janini LM, Tanuri A, Schechter M, Peralta JM, Vicente ACP, Torre ND, Pieniazek NJ, Luo C-C, Ramos A, Soriano V, Schochetman G, Rayfield MA, Pieniazek D. (1998) Horizontal and verical transmission of human deficiency virus type 1 dual infections caused by viruses of subtypes B and C. *Journal of infectious diseases*, 177: 227-231

Jetzt AE, Yu H, Klarmann GJ, Ron Y, Preston BD, Dougherty JP. (2000) High rate of recombination throughout the human immunodefiency virus type 1 genome. *Journal of Virology*, 74:1234-1240

Jost S, Bernard MC, Kaiser L, Yerly S, Hirschel B, Samri A, Autran B, Goh L-E, Perrin L. (2002) A patient with HIV-1 superinfection. *New England of Journal Medicine*, 347(10): 731–736

Junghans RP, Boone LR, Skalka AM. (1982) Retroviral DNA H structures: displacement-assimilation model of recombination. *Cell*, 30(1):53–62

Kaleebu P, Yirrell D, French N, Lyagoba F, Rutebemberwa A, Cheingsong-popov R, Gilks C, Biryahwaho B, Weber J, Whitworth J. (2000) An Improved Algorithm for Determining HIV Type 1 Subtypes in a Primary Laboratory in Uganda. *Aids Research and Human Retroviruses*, 16(7):621–625

Kandathil AJ, Ramalingam S, Kannangai R, David S, Sridharan G. (2005) Molecular epidemiology of HIV. *Indian Journal of Medical Research,* 121: 333-344

Kantor R, Katzenstein DA, Efron B, Carvalho AP, Wynhoven B, Cane P, Clarke J, Sirivichayakul S, Soares MA, Snoeck J, Pillay C, Rudich H, Rodrigues R, Holguin A, Ariyoshi K, Bouzas MB, Cahn P, Sugiura W, Soriano V, Brigido LF, Grossman Z, Morris L, Vandamme A-M, Tanuri A, Phanuphak P, Weber JN, Pillay D, Harrigan PR, Camacho R, Schapiro JM, Shafer RW. (2005) Impact of HIV-1 Subtype and Antiretroviral Therapy on Protease and Reverse Transcriptase Genotype: Results of a Global Collaboration. *PLoS Medicine* April 2 (0325) 4 e112

Kleim JP, Bender R, Kirsch R, Meichsner C, Paessens A, Riess G. (1994) Mutational analysis of residue 190 of human immunodeficiency virus type 1 reverse transcriptase. *Virology,* 200:696-701

Koblavi-Deme S, Maurice C, Yavo D, Sibailly TS, N'guessan K, Kamelan-Tano Y, Wiktor SZ, Roels TH, Chorba T, Nkengasong JN. (2001) Sensitivity and specificity of human immunodeficiency virus rapid serologic assays and testing algorithms in an antenatal clinic in Abidjan, Ivory Coast. *Journal of Clinical Microbiology,* 39:1808-12

Koelsch KK, Smith DM, Little SJ, Ignacio CC, Macaranas TR, Brown AJL, Petropoulos CJ, Richman DD, Wong JK. (2003) Clade B HIV-1 superinfection with wild-type virus after primary infection with drug-resistant clade B virus. *AIDS,* 17(7):F11–16

Korber B, Foley B, Leitner T, McCutchan F, Hahn B, Mellors JW, Myers G, Kuiken C. (1997) Human retroviruses and AIDS: A compilation and analysis of nucleic acid and amino acid sequences. *Los Alamos National Laboratory, Los Alamos, NM*

Korber BT, Foley B, Gaschen B, Kuiken C. (2001) Epidemiological and Immunological Implications of the Global Variability of HIV-1. In: Retroviral Immunology: Immune Response and Restoration, edited by Guiseppe Pantaleo and Bruce D Walker. 2001, 1-32

Los Alamos HIV Sequence Database available at: http://www.hiv.lanl.gov/components/hiv-db. Accessed on 15 December 2006

Kornberg A, Baker J. DNA polymerase I of *E.coli*. In DNA Replication. 2nd edition, 1992, 130 -131

Kwok S and Sninsky SS. (1993) PCR detection of human immunodeficiency virus type 1 proviral DNA sequences. In: Diagnostic Molecular Biology: Principles and Applications. Eds. Persing, D.H., Smith, T.F., Smith, F.C., *et al*. ASM, Washington, DC.

Kwok S, Mack DH, Mullis KB, Poiesz B, Ehrlich G, Blair D, Friedman-Kien A, Sninsky JJ. (1987) Identification of Human Immunodeficiency Virus Sequences by Using In Vitro Enzymatic Amplification and Oligomer Cleavage Detection. *Journal of Virology*, May p. 1690-1694

Lal RB, Chakrabarti S and Yang C. (2005) Impact of genetic diversity of HIV-1on diagnosis, antiretroviral therapy & vaccine development. *Indian Journal of Medical Research*, 121: 287-314

Larder BA, Stammers DK. (1999) Closing in on HIV drug resistance. *Nature Structural Biology*, 6:103-6

Leitner T. (1996) Genetic Subtypes of HIV-1.Theoretical Biology and Biophysics, Group T-10, MS K710, *Los Alamos National Laboratory, Los Alamos NM 87545*

Leitner T, Korber B, Daniels M, Calef C, Foley B. (2005) HIV-1 Subtype and Circulating Recombinant Form (CRF) Reference Sequences, *Los Alamos National Laboratory, Los Alamos, NM*, 87545

Loemba H, Brenner B, Parniak MA, Ma'ayan S, Spira B, Moisi D, Oliveira M, Detorio M, Wainberg MA. (2002) Genetic divergence of human immunodeficiency virus type 1 Ethiopian clade C reverse transcriptase (RT) and rapid development of resistance against nonnucleoside inhibitors of RT. *Antimicrobial Agents Chemotherapy*, 46: 2087– 2094

Lole K, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, Ray SC. (1999) Full-length human immunodeficiency virus type 1 genomes from subtype C infected seroconverters in India, with evidence of intersubtype recombination. *Journal of Virology*, 73:152-160

Ly TD, Laperche S, Courouce AM. (2001) Early detection of human immunodeficiency virus infection using third and fourth generation screening assays. *European Journal of Clinical Microbiological Infectious Diseases,* 20:104–10

Mansky LM, Temin HM. (1995) Lower In Vivo Mutation Rate of Human Immunodeficiency Virus Type 1 than That Predicted from the Fidelity of Purified Reverse Transcriptase. *Journal of Virology*, Aug, 5087–5094

Marquina S, Leitner T, Rabinovich RD, Benetucci J, Libonatti O, Albert J. (1996) Coexistence of subtypes B, F, and a B/F *env* recombinant of HIV type 1 in Buenos Aires, Argentina. *AIDS Research and Human Retroviruses,* 12: 1651-1654

Masciotra S, Livellara B, Belloso W *et al.* (2000) Evidence of a High Frequency of HIV-1 Subtype F Infections in a Heterosexual Population in Buenos Aires, Argentina. *Aids Research and Human Retroviruses*, 16 (10):1007–1014

Masemola A, Mashishi T, Khoury G, Mohube P, Mokgotho P, Vardas E, *et al.* (2004) Hierarchical targeting of subtype C human immunodeficiency virus type 1 proteins by CD8+ T cells: correlation with viral load. *Journal of Virology*, 78(7):3233–43

McCutchan FE, Carr JK, Bajani M, Sanders-Buell E, Harry TO, Stoeckli TC, Robbins KE, Gashau W, Nasidi A, Janssens W, Kalish ML. (1999) Subtype G and multiple forms of A/G intersubtype recombinant human immunodeficiency virus type 1 in Nigeria. *Virology*, 254:226–234

McCutchan FE. (2000) Understanding the genetic diversity of HIV-1. *AIDS*, (suppl 3): S31-S44

Moore JP, Cao Y, Leu J, Qin L, Korber B, Ho D. (1996) Inter- and intraclade neutralization of human immunodeficiency virus type 1: genetic clades do not correspond to neutralization serotypes but partially correspond to gp120 antigenic serotypes. *Journal of Virology*, 70:427–444

Myers G, Korber B, Berzofsky JA, Smith RF, Pavlakis GN (Eds): Human Retroviruses and AIDS 1992. Los Alamos: Los Alamos National Laboratory

Novitsky V, Cao H, Rybak N, Gilbert P, McLane MF, Gaolekwe S, *et al.* (2002) Magnitude and frequency of cytotoxic T-lymphocyte responses: identification of immunodominant regions of human immunodeficiencyvirus type 1 subtype C. *Journal of Virology*, 76(20):10155–68

Oelrichs R. (2004) The subtypes of human immunodeficiency virus in Australia and Asia. *Sexual Health*, 1:1-11

Orita, M., H. Iwahana, H. Kanazawa, K. Hayashi, T. Sekiya. (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proceedings of the National Academy of Sciences, USA,* 86:2766–2770

Peeters M , Esu-Williams E, Nzilambi N, *et al.* (1998) Molecular epidemiology of HIV-1 genetic subtypes in West and Central Africa. *XII International Conference on AIDS.* Geneva, June [abstract 11163]

Peeters M, Sharp PM. (2000) Genetic diversity of HIV-1: the moving target. *AIDS*, 14 (suppl 3): 129-S140

Peeters M, Toure-Kane C, Nkengasong JN. (2003) Genetic diversity of HIV in Africa: impact on diagnosis, treatment, vaccine development and trials. *AIDS*, 17: 2547 –2560

Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. (1996) HIV-1 dynamics *in vivo*: virion clearance rate, infected cell life-span, and viral generation time. *Science*, 271, 1582–1586

Pieniazek D, Ellenberger D, Janini LM, Ramos AC, Nkengasong J, Sassan-Morokro M, Hu DJ, Coulibally I-M, Ekpini E, Bandea C, Tanuri A, Greenberg AE, Wiktor SZ, Rayfield MA. (1999) Predominance of human immunodeficiency virus type 2 Subtype B in Abidjan, Ivory Coast. *AIDS Research and Human Retroviruses*, 15: 603-8

Pillay D, Walker AS, Gibb DM, de Rossi A, Kaye S, Ait-Khaled M, Munoz-Fernandez M, Babiker A, for the Paediatric European Network for Treatment of AIDS (PENTA). (2002) Impact of human immunodefiency virus type 1 subtypes on virologic response and emergence of drug resistance among children in the Paediatric European Network for Treatment of AIDS (PENTA) 5 trial. *Journal of Infectious Diseases*, 186: 617-625

Piot P, Bartos M. The epidemiology of HIV and AIDS. In: Essex M, Mboup D, Kanki PJ, Marlink PJ, Tou SR, editors. AIDS in Africa, 2nd ed. New York: 2002. 200-217

Plantier J-C, Vergne L, Damond F, MBoup S, MPoudi-NGole E, Buzelay L, Farfana I, Brand D, Peeters M, Brun-Vezinet F, Delaporte E, Barin F. (2002) Development and Evaluation of a DNA Enzyme Immunoassay Method for *env* Genotyping of Subtypes A through G of Human Immunodeficiency Virus Type 1 Group M, with Discrimination of the Circulating Recombinant Forms CRF01_AE and CRF02_AG. *Journal of Clinical Microbiology*, March. 1010–1022

Preston BD, Dougherty J P. (1996) Mechanisms of retroviral mutation. *Trends in Microbiology*, 4: 16–21

Preston BD, Poiesz BJ, Loeb LA. (1988) Fidelity of HIV-1 reverse transcriptase. *Science, New Series*, 242(4882): 1168-1171

Pulsinelli GA and Temin HM. (1994) High rate of mismatch extension during reverse transcription in a single round of retrovirus replication. *Genetics*, 91: 9490-9494

Quinones-Mateu M, Arts E. (1999) Recombination in HIV-1: update and implications. *AIDS Reviews*, 1: 89-100

Quinones-Mateu ME, Albright JL, Mas A, Soriano V, Arts EJ. (1998) Analysis of pol gene heterogeneity, viral quasispecies, and drug resistance in individuals infected with group O strains of human immunodeficiency virus type 1. *Journal of Virology*, 72: 9002-9015

Rambaut A, Posada D, Crandall KA, Holmes EC. (2004) The causes and consequences of HIV evolution. *Nature Reviews Genetics*, 5: 52-61

Ramos A, Hu DJ, Nguyen L, Phan KO, Vanichseni S, Promadej N, Choopanya K, Callahan M, Young NL, McNicholl J, Mastro TD, Folks TM, Subbarao S (2002) . Intersubtype human immunodeficiency virus type 1 superinfection following seroconversion to primary infection in two injection drug users. *Journal of Virology*, 76(15):7444–52

Ramos A, Tanuri A, Schechter M, Rayfield MA, Hu DJ, Cabral MC, Bandea CI, Baggs J, Pieniazek D. (1999) Dual and Recombinant Infections: An Integral Part of the HIV-1 Epidemic in Brazil. *Emerging Infectious Diseases,* 5(1): 65-74

Ratner L, Haseltine W, Patarca R, Livak KJ, Starcich B, Josephs SF, Doran ER, Rafalski FA, Whitehorn EA, Baumeister K, Ivanoff L, Petteway SR, Pearson ML, Lautenberger JA, Papas TS, Ghrayeb J, Chang NT, Gallo RC, Wong-Staal F. (1985)Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature*, 313:277–284

Rayfield MA, Downing RG, Baggs J, Hu DJ, Pieniazek D, Luo C-C, Biryahwaho B, Otten RA, Sempala SDK, Dondero TJ and the HIV Variant Working Group. (1998) A molecular epidemiologic survey of HIV in Uganda. *AIDS*, 12: 521-527

Rayfield MA, Sullivan P, Bandea CI, Britvan L, Otten RA, Pau CP, Pieniazek D, Subbarao S, Simon P, Schable CA, Wright AC, Ward J, Schochetman G. (1996) HIV-1 group O virus identified for the first time in the United States. *Emerging Infectious Diseases*, 2: 209-212

Re MC, Bon I, Schiavone P, Vitone F, Gibellini D. (2005) Discordant resistance interpretations in multi-treated HIV-1 patients. *International Journal of Antimicrobial Agents*, 25: 211–215

Roberts JD, Katarzyna B, Kunkel TA. (1988) The accuracy of reverse transcriptase from HIV-1. *Science, New Series*, 242(4882): 1171-1173

Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, Gao F, Hahn BH, Kalish ML, Kuiken C, Learn GH, Leitner T, McCutchan F, Osmanov S, Peeters M, Pieniazek D, Salminen M, Sharp PM, Wolinsky S, Korber B. (2000) HIV-1 nomenclature proposal. *Science*, 288:55-56

Salminen MO, Johansson B, Sonnerborg A, Ayehunnie S, Gotte D, Leinikki P, Burke DS, McCutchan FE. (1996) Full-length sequence of an Ethiopian human immunodeficiency virus type 1 (HIV-1) isolate of genetic subtype C. *AIDS Research and Human Retroviruses*, 12: 1329-39

Salminen MO, Carr JK, Robertson DL, Hegerich P, Gotte D, Koch C, Sanders-Buell E, Gao F, Sharp PM, Hahn BH, Burke DS, McCutchan FE. (1997) Evolution and probable transmission of intersubtype recombinant human immunodeficiency virus type 1 in a Zambian couple. *Journal of Virology*, *71*: 2647-55

Saunders NA, Alexander S, Tatt I. (2005) *env* Gene typing of human immunodeficiency virus type 1 strains on electronic microarrays. *Journal of Clinical Microbiology*, 43(4):1910-6

Simon F, Mauclere P, Roques P, Loussert-Ajaka I, Muller-Trutwin MC, Saragosti S, *et al.* (1998) Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *National Medicine*, *4*:1032-7

Smith DM, Wonga JK, Hightower GK, Ignacio CC, Koelsch KK, Petropoulos CJ, Richman DD, Little SJ. (2005) HIV drug resistance acquired through superinfection. *AIDS*, 19:1251–1256

Soriano V, Gomes P, Heneine W, Holguin A, Doruana M, Antunes R, Mansinho K, Switzer WM, Araujo C, Shanmugan V, Lourenco H, Gonzalez-Lahoz J, Antunes F. (2000) Human immunodeficiency virus type 2 (HIV-2) in Portugal: Clinical Spectrum, Circulating Subtypes, Virus Isolation, and Plasma Viral Load. *Journal of Medical Virology*, 61:111-6

Steain MC, Wang B, Yang C, Shi Y-P, Nahlen B, Lal RB, Saksena NK. HIV Type 1 Sequence Diversity and Dual Infections in Kenya. (2005) *Aids Research and Human Retroviruses*, 21(10): 882-885

Stuhlmann H, Berg P. (1992) Homologous recombination of copackaged retrovirus RNAs during reverse transcription. *Journal of Virology*, 66: 2378-2388

Swanson P, de Mendoza C, Joshi Y, *et al.*(2005) Impact of Human Immunodeficiency Virus Type 1 (HIV-1) Genetic Diversity on Performance of Four Commercial Viral Load Assays: LCx HIV RNA Quantitative, AMPLICOR HIV-1 MONITOR v1.5, VERSANT HIV-1 RNA 3.0, and NucliSens HIV-1 QT. *Journal of Clinical Microbiology*, 43(8)3860–3868

Tatt IA, Barlow KL, Clewley JP. (2000) A *gag* gene heteroduplex mobility assay for subtyping HIV-1. *Journal of Virological Methods*, 87: 41–51

Tatt ID, Barlow KL, Nicoll A, Clewley JP. (2001) The public health significance of HIV-1 subtypes. *AIDS*, (Suppl 5) 15: S59-71

Tegtmeier GE. (1988) Current tests for serologic detection of HIV-infection. *Journal of Clinical Immunoassay*, 11: 112-24

Temin HM. (1991) Sex and recombination in retroviruses. *Trends in Genetics*, 7:71–74

Temin HM. (1993) Review: Retrovirus variation and reverse transcription: Abnormal strand transfers result in retrovirus genetic variation. *Proceedings of the National Academy of Sciences, USA*, 90: 6900-6903

Thomson MM, Delgado E, Herrero I, Villahermosa ML, Vázquez-de Parga E, Cuevas MT, Carmona R, Medrano L, Perez-Alvarez L, Cuevas L, Najera R. (2002) Diversity of mosaic structures and common ancestry of human immunodeficiency virus type 1 BF intersubtype recombinant viruses from Argentina revealed by analysis of near full-length genome sequences. *Journal of General Virology*, 83:107–19

Thomson MM, Nájera R. (2001) Travel and the introduction of human immunodeficiency virus type 1 non-B subtype genetic forms into Western countries. *Clinical Infectious Diseases*, 32:1732–37

Thomson MM, Pérez-Álvarez L, Nájera R. (2002) Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy. *The Lancet Infectious Diseases*, 2:461-471

Treurnicht FK, Smith T-L, Susan Engelbrecht, Claassen M, Robson BA, Zeier M, van Rensburg EJ. (2002) Genotypic and phenotypic analysis of the *Env* Gene from South African HIV-1 subtype B and C isolates. *Journal of Medical Virology*, 68(2): 141-146

Triques K, Bourgeois A, Saragosti S, Vidal N, Mpoudi-Ngole E, Nzilambi N, Apetrei C, Ekwalanga M, Delaporte E, Peeters M. (1999) High diversity of HIV-1 subtype F strains in Central Africa. *Virology*, 259:99-109

Triques K, Bourgeois A, Vidal N, Mpoudi-Ngole E, Mulanga-Kabeya C, Nzilambi N, Torimiro N, Saman E, Delaporte E, Peeters M. (2000) Near full-length genome sequencing of divergent African type 1 subtype F viruses leads to the identification of a new HIV type 1 subtype designated K. *AIDS Research and Human Retroviruses*, 16:139-151

UNAIDS. AIDS Epidemic Update 2006, <http://www.unaids.org/worldaidsday/2006/press/Epiupdate.HtmlReport> (2006).UNAIDS: Overview of the Global AIDS Epidemic. Report on the Global AIDS Epidemic

van der Kuyl AC, Kozaczynska K, van der Burg R, Zorgdrager F, Back N, Jurriaans S, Berkhout B, Reiss P, Cornelissen M. (2005)Triple HIV Infection. *New England Journal of Medicine*, 352(24):2557-2559

Van Harmelen JH, Van Der Ryst E, Loubser AS, York D, Madurai S, Lyons S, Wood R, Williamson C. (1999) A Predominantly HIV Type 1 Subtype C-Restricted Epidemic in South African Urban Populations. *AIDS Research and Human Retroviruses*, 15(4): 395-398

Vanden-Haesevelde MV, Decourt J-L, De Leys RJ, Vanderborght B, van der Groen G, van Heuverswijn H, Saman E. . (1994) Genomic cloning and complete sequence analysis of a highly divergent African human immunodeficiency virus isolate. *Journal of Virology*, 68: 1586-1596

Vella S, Mildvan D. (1998) Clinical treatment: overview. *AIDS*, 12; S147-8

Vidal N, Peeters M, Mulanga-Kabeya C, Nzilanbi N, Robertson D, Ilunga W, Sema H, Tshimanga K, Bongo B, Delaporte E.. (2000) Unprecedented degree of human immunodeficiency virus type 1 (HIV-1) group M genetic diversity in the Democratic Republic of Congo suggests that the HIV-1 pandemic originated in Central Africa. *Journal of Virology*, 74: 10498–58

Wagner EK, Hewlett MJ. Retroviruses: Converting RNA to DNA. In: Basic Virology. 2nd edition. Blackwell Publishing. 2004. pgs 362-366

Wainberg MA. (2004) HIV-1 subtype distribution and the problem of drug resistance. *AIDS*, 18 (suppl 3): S63-S68

Weniger BG, Brown T. (1996) The march of AIDS through Asia. *New England Journal of Medicine*, 335: 343-345

Weniger BG, Takebe Y, Ou C-Y, Yamazaki S. (1994) The molecular epidemiology of HIV in Asia. *AIDS*, 8(suppl 2):S13-S28

Williamson C, Engelbrecht S, Lambrick M, van Rensburg EJ, Wood R, Bredell W, Williamson A-L. (1995) HIV-1 subtypes in different risk groups in South Africa. *Lancet*, 346:782

Xin K-Q, Sekimoto Y, Takahashi T, Mizuguchi H, Ichino M, Yoshida A, Okuda K.(2006) Chimeric adenovirus 5/35 vector containing the clade C HIV *gag* gene induces a cross-reactive immune response against HIV. Vaccine, 25 (19): 3809-3815

Yamaguchi J, Devare SG, Brennan CA. (2000) Identification of a new HIV-2 subtype based on phylogenetic analysis of full-length genomic sequence. *AIDS Research and Human Retroviruses*; *16*: 925-30

Yang R, Xia X, Kusagawa S, Zhang C, Ben K, Takebe Y. (2002) On-going generation of multiple forms of HIV-1 intersubtype recombinants in the Yunnan Province of China. *AIDS*, 16:1401–1407

Zarandia M, Tsertsvadze T, Carr JK, Nadai Y, Sanchez JL, Nelson K. (2006) HIV-1 Genetic Diversity and Genotypic Drug Susceptibility in the Republic of Georgia. *Aids Research and Human Retroviruses*, 22(5):470–476

Zhang J. (2004) Host RNA polymerase II makes minimal contributions to retroviral frame-shift mutations. *Journal of General Virology*, 85: 2389–2395