Chemokine coreceptor usage and HIV-1 subtype C reservoir in AIDS patients

Te-chang (Mike) Lo

A dissertation submitted to the Faculty of Science, University of Witwatersrand, Johannesburg, in fulfillment of the requirement for the degree of Master of Science

Johannesburg, 2009

University of Witwatersrand

Department of Molecular Medicine and Haematology

This is dedicated to those I know and love.
Thank you for your faith, for putting up with me and urging me on

"It is the best of time; it is the worst of time."

- Charles Dickens

Acknowledgement

I would like to thank the following people who contributed to the work:

Dr. Pascale Willem, my supervisor.

Dr. Francois Venter, Dr. Dimitri and Dr. Thomas Kalk for their assistance in collecting patient samples.

Dr. Tonie Cilliers and Mr. Jabulani Nhlapo for HIV infected cell-line cultures.

Ms. Mia Coetzer and Dr. Maria Papathanasopoulos for their gag PCR primers and protocol.

Ms Eleanor Cave and Ms. Mia Coetzer for their assistance in sequencing and subsequent data analyses.

Financial support:

The author was supported by MRC bursary during the course of the degree.

Declaration

I hereby declare that this dissertation is my own unaided work unless specified. It is being submitted for Masters of Science (Full research) at the University of Witwatersrand, Johannesburg, South Africa. It has not been submitted for any other degrees at any other university.

_____

Te-chang (Mike) Lo

Signed _____ day of _____, 2009

Table of Content

List of Figures

Acronyms and abbreviations

AIDS	Acquired Immune Deficiency Syndrome
APC/s	Antigen Processing Cell/s
ART	AntiRetroviral Therapy
CD4/8	Cluster Designation 4/8
CNS	Central Nervous System
CCR5	CC chemokine Receptor 5
CXCR4	CXC chemokine Receptor 4
DCs	Dendritic Cells
DNA	Deoxyribose Nucleic Acid
Env	Envelope
FISH	Fluorescent *in situ* Hybridization
FDCs	Follicular Dendritic Cells
Gag	Group-specific antigens
gp120/160	glycoprotein 120/160
GIT	GastroIntestinal Tract
HAART	Highly Active AntiRetroviral Therapy
HAD	HIV Associated Dementia
HAM	HTLV-I-Associated Myelopathy
HIV	Human Immunodeficiency Virus
HTLV	Human T Lymphotropic Virus
IECs	Intestinal Epithelial Cells
ISH	*In Situ* Hybridization

LTR/s	Long Terminal Repeat/s
MDMs	Monocyte Derived Macrophages
MIP	Macrophage Inflammatory Protein
NHL	Non-Hodgkin Lymphoma
OI/s	Opportunistic Infection/s
PCP	*Pneumonocystis carinii* Pneumonia
Pol	reverse transcriptase, protease & integrase
R5X4	CCR5CXCR4-using
RNA	Ribose Nucleic Acid
RT	Reverse Transcriptase
Nef	Negative factor
RANTES	Regulated upon Activation, Normal T cells Expressed and Secreted
R/H	rapid/high
RRE	Rev Response Element
SDF-1	Stromal Derived Factor-1
SIV	Simian Immunodeficiency Virus
S/L	slow/low
TAR	Tat Activation Response
Tat	Trans-Acting Transcriptional activator
TCR	T Cell Receptor
Trs/Rev	Regulator of virions expression
TSP	Tropical Spastic Pararesis

Vif	Viral infectivity factor
Vpr	Viral protein r
Vpu	Viral protein u

Abstract

HIV pathogenesis relies on infection of $CD4^+$ cells. Viral entry into these cells initially utilizes CCR5 to infect monocytes and macrophages, and may later use CXCR4 to infect T cells. Later stage HIV infection progressing towards AIDS is characterized by high viral load and $CD4^+$ T cell depletion. In HIV-1 subtype C infections, there is rarely a switch to CXCR4. However, T cells, which are predomoninantly $CXCR4^+$ cells, still decrease in number. To shed some light on the cause of T cell depletion, HIV-gag specific FISH and HIV V3 loop sequencing were used to assert the location of the virus, and the chemokine coreceptor used, respectively. The majority of $CD4^+$ cells in blood and bone marrow extracts of late stage HIV^+ patients possessed positive signals. T cells were present in large number compared to monocytes/ macrophages in both compartments. Sequencing indicated CCR5 was still the preferred chemokine coreceptor. These findings support the hypothesis that the virus is able to enter T cells, despite a lack of coreceptor switching. It also raises the questions of whether CCR5 is upregulated on T cells, or other mechanisms are used by the virus to gain entry into T cells.

Chapter 1: Introduction

1.1 General background

Acquired Immune Deficiency Syndrome (AIDS) was first recognized in 1982 as a communicable disease that can be transmitted via sexual contacts, intravenous drug use, and blood transfusion (Barre-Sinoussi et al, 1983). However, its causative agent, the Human Immunodeficiency Virus (HIV) was only determined in 1983/4 to be a retrovirus that targets the helper/inducer T cells (Barre-Sinoussi et al., 1983; Gallo et al., 1984; Sarngaharan et al., 1984), much like the two other exogenous retroviruses that target humans, characterized since 1980 (Feinberg et al., 1986). Consequently, the virus was initially named Human T Lymphotropic Virus (HTLV) –III, as it was thought to have close association with the other two HTLV species. Although this has been shown to be the case - immunological and nucleic acid data had revealed HIV does belong to the HTLV family, and have closer association to HTLV-II than –I (Schupbach et al, 1984; Sarngadharan et al., 1984) – the etiology of HIV is very different to the other two HTLV members. HTLV-I has been implicated as the cause of adult T cell leukemia as well as HTLV-I-associated myelopathy/tropical spastic pararesis (HAM/TSP) (Barmak et al., 2003); HTLV-II is associated with a T-cell variant of hairy cell leukemia (Kalyanaraman et al., 1982). Both of these viruses have been found in AIDS patients as well, suggesting some possible etiological links between the three HTLV species.

Aside being known as HTLV-III, HIV also went by the names lymphadenopathy virus (LAV) and AIDS-associated retrovirus (ARV). As researchers revealed more of the virus' characteristics, it became more obvious that it shares a number of similarities with members of the nontransforming cytopathic lentivirus family of the retroviruses (Barre-Sinoussi et al., 1983). This is where scientists have placed it at present. With respect to the other members of the lentiviruses (e.g. sheep visna virus, equine infectious anemia virus, feline immunodeficiency virus), HIV possesses similar morphological, biological, molecular and pathogenic characteristics. Perhaps the closest member to the HIV is the primate variation of the virus, called Simian Immunodeficiency Virus (SIV). There are multiple SIV variants, originating from different species. HIV-1 is most closely related to the virus found in chimpanzees (SIV_{CPZ}), which suggest cross-species transmission from the chimpanzee. But SIV-positive chimpanzees have rarely been documented, making it difficult to determine the pathogenesis of SIV_{CPZ}. However, SIV-positive chimpanzees do not seem to suffer from the same deleterious effects as HIV-infected humans, which suggest that SIV_{CPZ} may itself have been the product of another cross-species transmission (Sharp et al., 1999).

Another retrovirus was also found in West African AIDS patients in 1986 and was determined to have similar biological and morphological properties to HIV-1; it was subsequently named HTLV-IV, or HIV-2. However, while the envelope glycoproteins of HIV-2 differ greatly to HIV-1, there is a serological commonality between HIV-2 and SIV. This suggests that HIV-2 is more closely related to the simian virus than HIV-1 (Clavel et al., 1986), and have led to the hypothesis that HIV-2 infection is a zoonosis, a

process whereby a disease of animal is transmitted to man and becomes established in human population. The zoonosis theory for the rise of HIV-2 has been generally accepted as a valid one, due to the amount of genetic evidence linking it to the primate form of the virus. It remains uncertain how HIV-1 appeared and adapted itself to human. HIV-2 infected individuals do exhibit clinical syndromes very similar to those infected with HIV-1 (Clavel et al., 1987), but the rate of death is much lower than compared to HIV-1 positive individuals (Poulsen et al., 1997). Early studies of HIV-2 had also suggested that co-infection of HIV-1 and HIV-2 can slow down the progression of the more aggressive HIV-1, but that been shown not to be the case with a more recent study that had a larger cohort (Schim van der Loeff et al., 2001).

HIV-1 has a number of genetic strains, classified into three clades, M (main), O (other), and N (non-M, non-O). Within these classifications, there are also divisions of subtypes as well as variants created from two or more existing subtypes, called circulating recombination forms (CRFs). HIV subtypes are defined as having genomes that are at least 25% unique. Eleven subtypes and thirteen CRFs have been identified so far, and almost all subtypes and a number of CRFs can be found in Africa. Naturally, most of the HIV subtypes are within the M clade, and HIV-1 subtype C is the most prevalent strain of the virus at present. It contributes to approximately 50% of the new infections worldwide, although it is most commonly found in Africa and parts of Asia. It usually spreads via heterosexual contacts. The second most prevalent strain is subtype A, which accounts for the 30% of new infections, while the most common subtype found in America and Europe Australia is subtype B, which seems to spread by homosexual contacts or

intravenous drug users (Berkelman et al., 1989; Quinn, 1990). According to UNAids's 2004 report, there are a total of 39 million people infected with HIV at the end of 2003, of which of which 4.9 million were new infections; at the same time 3.1 million had died from AIDS and associated diseases (http://www.unaids.org). While this seems to be a small decrease, compared to the 2003 figure (40 million infected, 3 million deaths) it is not indicative of the decline of the pandemic. Rather, it may be in indication of the infection rate hitting a plateau, while death rate are starting to increase as more and more infected individuals reach AIDS.

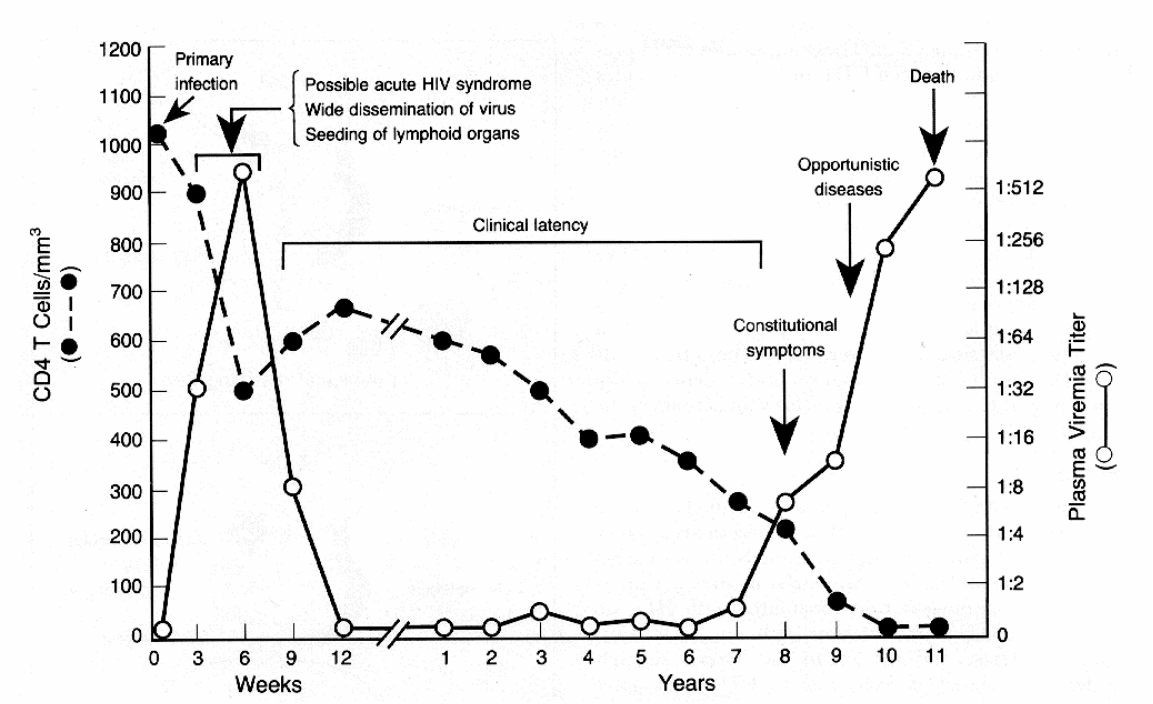
The most badly affected region by HIV-1 at the present is Sub-Saharan Africa, where it accounts for approximately 65% of all of the three figures quoted above. But more specifically, it is southern Africa that is worst hit, and most countries in the region are on top of the list of the world's highest HIV-infected populations. Botswana and Swaziland tops the list, with around 38% of adults (15-49 years old) infected with HIV, followed by Lesotho, Zimbabwe, South Africa, and Namibia (http://www.unaids.org). It should be stated, however, that percentage of population affected is somewhat misleading, as it does not reflect the absolute number of infected individuals, and therefore does not mean these countries has the highest total number of HIV/AIDS cases. Indeed, when compared with the highly populous Asian countries, where 1% prevalence may means *10 million* people are infected with the virus, as opposed to 38.8% prevalence equating 220 000 people in Swaziland. Having said that, data shows that South Africa presently has the most number of documented HIV^+ cases, which is approximately 5 million individuals.

Therefore it should be kept in mind that while the various numbers represent one facet of the HIV problem, it should not be seen as the complete picture to the HIV epidemic.

Beyond the genetic classifications, the virus is further divided by the mechanism of their entry into the host cells. Almost all HIV infections begin using the chemokine receptor CCR5 and are hence termed R5-using viruses (Schuitemaker et al., 1992; van't Wout et al., 1994). Later in the infection, some virus, particularly those of the subtype B strain, have the tendency to switch chemokine receptor usage from CCR5 to CXCR4, or use both receptors at the same time. These variants are termed X4-using (or R5X4) viruses, and are much more virulent, as well as being able to induce the formation of large, multinucleated cells *in vitro*, otherwise known as synctia formation (Bjorndal et al., 1997; Tscherning et al., 1998). This ability to induce synctia formation has also led to alternate names for the virus. The R5-using virus is considered to have a non-syncytium inducing phenotype, while the X4-using virus has a syncytium-inducing phenotype. A third class of differentiation between the two groups looks at their growth rate *in vitro*: The non-syncytium-inducing/R5-using virus has a slow/low (S/L) rate of growth, while the syncytium-inducing/X4-using virus has a rapid/high (R/H) growth rate (Asjo et al., 1986).

X4-using viruses has been shown to be more virulent due to their ability to cause synctia *in vitro*, and is a strong indicator of decline in CD4 cell count and disease progression (Koot et al., 1993; Shafer, Aguiniga & Merigan, 1995). It does not, however, increase the immediate risk of death (Richman & Bozzette, 1994). Furthermore, the X4-using virus very rarely causes synctia *in vivo*, which leads to question as to why and how it is more

virulent in the infected system. This accelerated disease progression could be due to the use of additional chemokine coreceptor that has allow infection of previously unassailable cells ($CD4^+$ T cells), and cause the immune system to deteriorate even faster.



<u>Figure 1.1</u>: Progression of HIV infection towards AIDS, with specific reference to the important events that are markers in pathogenesis and clinical status of the infected individual (From Pantaleo et al., 1993). The numbers present on both axes are general values, as the progression of the disease can vary from individuals.

The mutations may also enhance the binding affinity of gp120 when interacting with CXCR4 when compared to CCR5, which increase the efficiency of entry by the virus into cells. This change in receptor usage occurs due to mutations that appear in the third

hypervariable loop of gp120, and has been associated with the phenotypic change in subtype B (De jong et al., 1992; Shizimu et al., 1999).

The clinical progression of HIV infection towards AIDS has been clearly documented by research studies and reviews (Morrow, Park & Wakefield, 1994; Pantaleo, Grazlosi & Fauci., 1993). Clinically, the progression can be put into two broad categories: acute and chronic phase. After a primary infection, 40 to 90 percent of infected individuals present an acute mononucleosis-like or flu-like syndrome 3 to 6 weeks after the initial infection. Symptoms include fever, malaise, maculopapular rash, and oral ulcers, to name a few. Due to the similarity of symptoms between viral-based flu and the initial HIV infection, and the fact that there is no detectable HIV-1 antibody at this point, it is next to impossible to determine if an HIV infection has occurred, unless there are circumstantial evidences alluding to the possibility of HIV infection. A high level of viraemia is also present at this time, accompanied by body-wide dissemination of the virus. This leads to an immune response against the virus within a week to three months of the first physical symptoms, which is the point where seroconversion takes place – that is, the presence of anti-p24 and anti-160gp antibodies is detected in the body. The HIV-1-specific antibodies also bring about a rapid decline of detectable viraemia in the body, although that doesn't seem to stop the viral replication fully (Morrow, Park & Wakefield, 1994; Pantaleo, Grazlosi & Fauci., 1993).

Once seroconversion occurs, there is still a latent stage of variable time period where there are few clinical symptoms, except the slow increase of viral load and decrease in

$CD4^+$ T cells in the blood. The chronic phase of the disease reaches its end when the CD4 count of the infected individual declines to two hundred or less per microlitre of blood. This is classified as full-blown AIDS, at which point, the individual becomes very susceptible to opportunistic infections like tuberculosis. Other common diseases that affect AIDS patients include thrush, Kaposi sarcoma, *Pneumocystis carinii* Pneumonia (PCP), Non-Hodgkin Lymphoma (NHL), cryptococcal meningitis, and mycobacterium avium. At this late stage, the virus can also be found in the central nervous system, leading to peripheral neuropathy, or dementia and other neurological disorders in more severe cases. Therefore, HIV infection does not directly cause death, but weakens the immune system sufficiently so that when opportunistic diseases invade the body, the pathogen overwhelms the body and cause death. Figure 1.1 describes the progression of the infection with respect to the viraemia and $CD4^+$ T cell counts, with reference to certain key events, described above (Morrow, Park & Wakefield, 1994; Pantaleo, Grazlosi & Fauci., 1993).

1.2 Genetic and structural makeup of HIV

HIV has a ~10kb proviral genome, encoding for fifteen proteins and enzymes. While not all of these are essential for the proliferation of the virus, the less important gene products do aid in a more efficient process thereof. The genome is flanked on both sides by long terminal repeats (LTR) that are identical to one another and regulate both HIV replication and viral expression. The latter function of LTR affects the three major viral genes: the

gag (group-specific antigen) gene that encodes for the structural proteins, *pol* (protease, reverse transcriptase, and integrase) gene that produce the enzyme complex reverse transcriptase-protease-endonuclease, and the *env* (envelope) gene that give rise to both the internal and external glycoprotein envelope (Morrow, Park & Wakefield, 1994).

1.2.1 The major viral genes

The *gag* gene produces a polyprotein precursor ($Pr55^{gag}$) which undergoes proteolytic cleavage to produce the matrix protein, the capsid protein, and the nucleocapsid protein. The matrix protein is a 17kDa molecule that associates itself with the inner surface of the viral membrane as well as the icosahedral capsid of the core. It has also been suggested that the matrix protein is required to incorporate the envelope proteins into mature virions, and aid in virus entry. The 24kDa capsid protein was produced from the central domain of $Pr55^{gag}$ and creates the shell of the cone-shaped core of the virus that houses the genomic viral RNA and enzymes; it also contains motifs needed for proper viral assembly. The carboxy terminus of the gag polyprotein also seems to have functions in assembly, as well as producing the nucleocapsid proteins. The nucleocapsid is a 9-kDa protein that associates with viral RNA in the virion core; it also has a role in virion assembly. A small proline-rich peptide (p6) is produced from the carboxy terminus of the gag polyprotein, and affects the efficiency of virion budding on the cell surface. There are also other small peptides (e.g. p9) that are produced from the terminus, but these have not been found with a function at present (Morrow, Park & Wakefield, 1994).

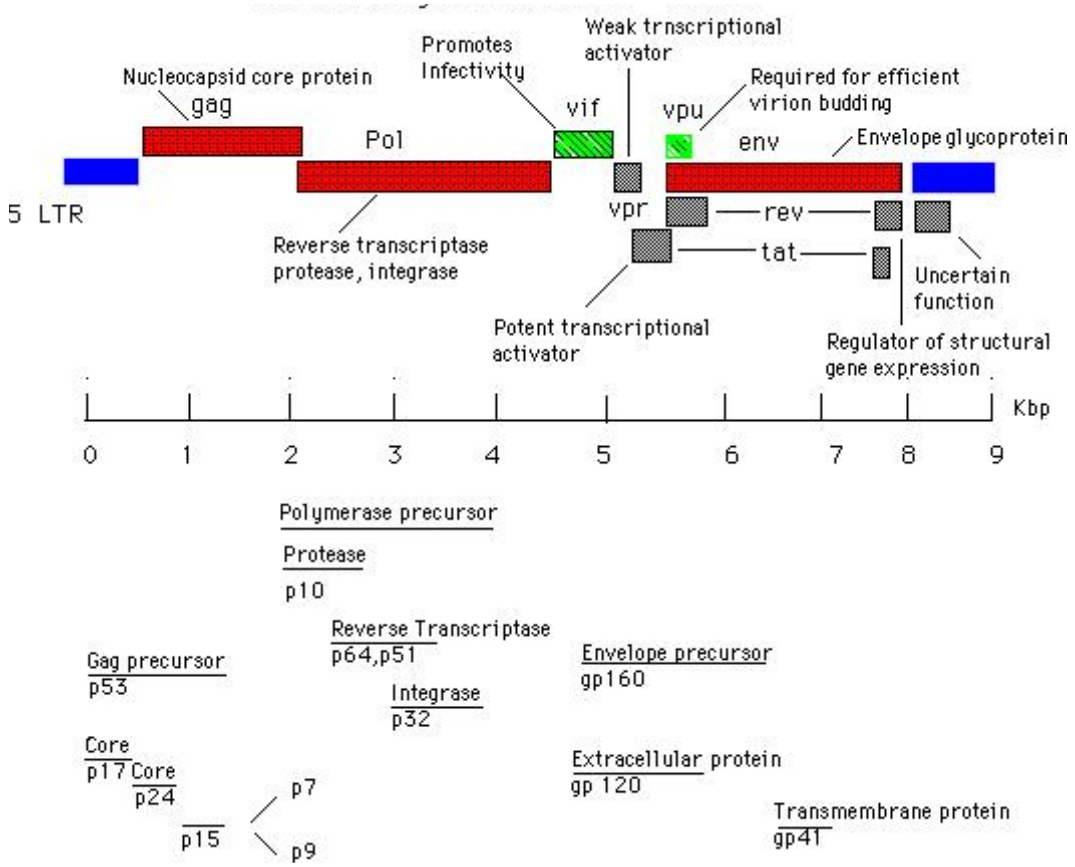


Figure 1.2: Organization of HIV genome. The approximately 10kb RNA sequence is flanked by Long Terminal Repeats (LTRs), and genes for the necessary proteins and enzymes are situated between them.

(Taken from http://uhavax.hartford.edu/bugl/hiv.thtm).

The *pol* gene is responsible for the three viral enzymes, and they are produced as a polyprotein, much like the gag polyprotein. The initiation site and 200-odd nucleotides of the 5'-end of *pol* overlaps with the 3'-end of *gag*, and the production of the pol polyprotein requires ribosomal frameshifting (the ribosome slipping back by 1 base) in order to create the correct codons. The 160 kDa pol precursor polyprotein is therefore a fusion protein created from both the *gag* and *pol* genes. The HIV protease is the first to

be produced from the polyprotein by autoprocessing event, and it needs to be fully active before the other two enzymes can be produced – it is responsible for the cleaving events on the pol polyprotein to release the reverse transcriptase and integrase.

The gag precursor polyprotein also requires the protease's activity to be processed. Considering its compulsory presence in a number of early protein formations, it is perhaps not surprising mutations in the protease can cause the production of noninfectious HIV virions. The reverse transcriptase (RT) is a dimeric enzyme, initially composed of two identical 66kDa (p66) polypeptides. The homodimer then undergo asymmetric endoproteolytic cleavage, where the one subunit is reduced to a 51 kDa (p51) polypeptide. Although the p51 subunit has demonstrated no enzymatic activities, its removal terminates certain activities on the other subunit. The viral RT has a number of catalytic activities, including DNA polymerase activity that copy both DNA and RNA templates and ribonuclease H activity, which degrades RNA in a RNA-DNA hybrid duplex. The HIV integrase is a 32 kDa protein and is essential for the integration of retroviral DNA into the host genome. It also possesses a number of functions, including DNA binding, specific DNA endonuclease activity, DNA integration and recombination (Morrow, Park & Wakefield, 1994).

The *env* gene produces yet another polyprotein precursor, which give rise to two mature form of envelope glycoprotein. The precursor is an 88 kDa polypeptide that is cleaved during translocation to the rough endoplasmic reticulum, where glycosylation and folding occurs. It then undergoes oligomerization to give stability to the protein. The

glycosylated precursor is now 160kDa (gp160), and is cleaved in the Golgi to produce a 120kDa surface glycoprotein, and a 41kDa transmembrane glycoprotein. These are transported to the plasma membrane to be incorporated into the budding virions. This process is quite inefficient, as only 5-15% of the gp160 precursors are cleaved to produce the two glycoproteins. The two envelope proteins are noncovalently linked to one another as trimers, present as spikes on the viral membrane; the strength of the association varies between strains. Both proteins are essential for the viral infection to occur, and this will be looked at in more detail (Morrow, Park & Wakefield, 1994).

1.2.2 The accessory genes

Six other genes exist in the proviral genome as well, and produce what can be classified as accessory or auxiliary proteins. They all possess regulatory functions and are able to influence the pathogenesis of the virus. Initially, most of these were thought to be noncompulsory proteins, and was merely there to enhance or aid the viral infection, but studies have determined that these proteins are just as important, if not essential, to the virus' life-cycle. The trans-acting transcriptional activator (*tat*) gene is among the first genes to be expressed and produce the most important of the accessory proteins. Tat's principle function is to bind to the tat activation response (TAR) element on the 5' end of the nascent mRNA and recruits cellular factors that promote processivity of transcription from the viral long terminal repeats (LTRs) (Turner and Summer, 1999). However, the mechanism of interaction is not fully elucidated between the tat and TAR element, and there have been indications that tat can still bring about transcription in the absence of TAR (Morrow, Park & Wakefield, 1994).

Tat has been singled out as the viral protein that seems to cause the majority of changes in the physiological behaviours of cells in the infected system. These include the upregulation of CXCR4 expression on erythroid cells (Gibellini et al., 2003), upregulation of inflammatory mediators and induction of monocytes into the brain (Pu et al., 2003), to name a few. It is unknown why tat protein is able to influence so many mechanisms outside the virus, although its presence throughout the body can certainly be attributed to the fact it is a soluble protein, and thus able to move into any area by following the movement of bodily fluids.

The *trs* or *rev* (regulator of virion expression) protein is required for viral replication, and is involved in the controlling the movement of RNA between the nucleus and cytoplasm. Rev binds the rev response element (RRE) on nascent unspliced mRNA, which recruits nuclear shuttling proteins to export the molecule and then import the processed mRNA. The rev protein is thus able to induce cytoplasmic expression of highly spliced mRNA (tat, rev, and nef) in early synthesis, and later, the unspliced genomic mRNA (gag and gag-pol) and singly spliced mRNAs (env, vpu, vif and vpr) mRNAs. It is also able to upregulate HIV synthesis via a trans-acting anti-repression mechanism (Feinberg et al., 1996), as well as reduces expression of doubly spliced RNA encoding for the viral regulatory proteins.

Two of the virus' regulatory/accessory proteins, Vpu (viral protein U) and Nef (negative factor), downregulate the surface expression of the CD4; they also have a similar but lesser effect on the expression of the major histocompatibility complex I. Vpu associates

itself with CD4 in the endoplasmic reticulum and recruits it to undergo proteolysis, whereas the Nef protein mediates endocytosis on removing the CD4 already on the cell surface (Emerman & Malim, 1998). However, recent data seems to indicate the function of Nef is more poised toward enhancing HIV-1 pathogenicity *in vivo* (Stoddart et al., 2003), even though it seem have been shown to be nonessential to viral replication *in vitro* (Herbein et al., 1998). While the reason for the virus to target CD4 for removal remains unclear, it is possible that these are mechanisms to both enhance infection as well as avoiding immune surveillance. Within the cells, CD4 and the envelope protein are both produced in the endoplasmic reticulum, and the removal of vpu can prevent premature binding of env. The downregulation of CD4 receptor expression would also prevent reinfection of the cell by budding virions (Turner & Summers, 1999).

The *R* or *Vpr* (viral protein R) gene as it has come to be known as, produces a protein that cause the nuclear localization of the preintegration complex (PIC), containing integrase, matrix, reverse transcriptase, and Vpr in early synthesis. Vpr also have the ability to delay host cells for extended periods of time in theG_2 phase, thus providing sufficient time for the virus to replicate (Emerman & Malim, 1998). The last accessory protein, Vif (viral infectivity factor) protein seemed to have strong effects in the infectivity of the virus *in vitro*, as its absence leads to the virus being 1500-2000-fold less infectious, and only allows the infection of transformed T cell lines *in vitro* (Mariani et al., 2003).

1.3 Mechanism of HIV-1 infection

HIV requires three components for successful invasion of the host cells. The first is its own envelope glycoprotein, gp120. The other two are the CD4 receptor and the chemokine receptor on the host cell. As previously stated, the envelope glycoprotein are made up of the gp120 surface protein and the gp41 transmembrane protein, which are non-covalently bonded to one another one the viral membrane as trimers (Cilliers & Morris, 2002). Gp120 has 5 conserved regions (C1-C5) that makes up its core, as well as 5 variable loops (V1-V5); it contain binding sites for both CD4 and the chemokine receptors. Gp41 possess 3 domains, one of which interact with the matrix protein, the second is the transmembrane domain, and lastly the external ectodomain (Morrow et al., 1994). CD4 is a 50kDa membrane glycoprotein that acts as a transducer in activation of peripheral T cells and differentiation of thymocytes. It is expressed on the surface of T cells, monocytes, dendritic cells and microglia in the brain. It primary function is in MHC II restricted interactions in immune responses, where T helper cells interact with antigen presenting cells (APCs) like dendritic cells to mount an immune response against the target antigen (MeSH database, 2004). Chemokine receptors are members of the seven-transmembrane domain receptor superfamily.

1.3.1 Chemokine receptor and HIV

The chemokine coreceptor is the most varied out of the three components. These receptors are members of the seven-transmembrane domain receptor superfamily. They are also known as G-protein coupled receptors, as the intracellular signals they transmit

are done so via the dissociation of heterotrimeric GTP-binding proteins. The receptors are 340-370 amino acids in length and positioned across the bilipid membrane, where the amino-terminal and three loops are presented extracellularly. The classification of chemokine receptors are closely modeled on the chemokine/s they interact with. All the chemokines have a conserved cysteine-based motif near their amino-termini, and differs in the number and sequential arrangement of the cysteine. Known classes of chemokine at this point include CXC, CC, C, and CX_3C, where the receptors for the first two classes (CXCR, CCR) are most relevant to HIV infection (Martin-Garcia et al., 2002).

Although CD4 was identified as the primary receptor fairly early on, it was found that the receptor only supported viral entry on a human cell. This suggested that a second, species-specific coreceptor was required to allow viral entry. It was proven to be correct when an orphan G-coupled receptor, fusin, was found to allow HIV-1 infection when coexpressed with CD4 (Feng et al., 1996). The discovery was quickly followed by the identification of CC-CKR-5, an alternative coreceptor that allows infection by the macrophage-tropic strain, as opposed to fusin which allowed infection by the T-cell tropic strain (Deng et al., 1996; Dragic et al., 1996). The two receptors were later renamed to CXCR4 and CCR5, respectively, to reflect their functionality. To this date, CXCR4 is only know to bind the chemokine SDF-1 (stromal derived factor-1), whereas CCR5 have a number of ligands, including MIP (macrophage inflammatory protein) -1α and -1β, as well as RANTES (regulated upon activation, normal T cell expressed and secreted) (Alkhatib et al., 1996).

It is important to look at the chemokine receptors in some details, as they are *the* components that can provide some natural resistance against viral entry. The first indication of possible HIV-1 resistance was when some highly exposed, but uninfected Caucasians were found. It was determined that a 32bp deletion on the CCR5 gene (CCR5Δ32) led to the production of a truncated protein, which fails to present itself on the cell surface. This mutation is somewhat common in those of Caucasian descent (15-20% of population is heterozygous; 1% homozygous mutant), but is rare in other populations (Liu et al., 1996; Samson et al., 1996). The heterozygous individuals have a decreased likelihood of infection by the R5 virus, and it is almost impossible for those with homozygous mutant to be infected. However, this protection only extends to the R5 virus, and these individuals can still be infected by dual-tropic, or X4 viruses. This was demonstrated sufficiently well by cases of CCR5Δ32 homozygous individuals being infected by X4 virus, with exclusive and persistent CXCR4 usage (O'Brien et al., 1997; Theodorou et al., 1997). While the average number of CCR5 receptor on any given macrophage has yet to be quantified, it has been shown that a high number of CCR5 can better facilitate the entry of the virus, and also determine the virus' postentry efficiency (Reynes et al., 2001; Lin et al., 2002).

It is unknown what controls the degree of expression of CCR5 on cell surface, but it would be of great interest to determine the mechanism, and perhaps suppress CCR5 expression. This should theoretically have little effects on the cells *in vivo*, as CCR5 is a redundant receptor – its ligands, MIP-1α, -1β and RANTES are able to utilize other chemokine receptors. Perhaps the most convincing evidence is that homozygous

CCR5Δ32 individuals seem to suffer from no ill-effects from the absence of the receptor (Liu et al., 1996). The influence of chemokines as competitive inhibitors against viral binding has also been looked at in CCR5 and its ligands. It was found that RANTES was the most potent inhibitor, followed by MIP-1α and -1β. The presence of these chemokine *in vivo* seldom approach the high concentration required to fully inhibit viral interaction with the receptor, but they are able to provide some deterrence to the process (Alkhatib et al., 1996).

There are also other less influential polymorphisms on CCR5 that have been found to affect viral infection or transmission. These include the CCR5P1 allele that enhance AIDS progression (Martin et al., 1998), CCR5-59353C that slows CD4 decline, and CCR5-59356T that increase perinatal transmission (Kostrikis et al., 1999), to name a few. While there are no mutations on CXCR4 to speak of, a polymorphism on its sole ligand, SDF-1 (SDF-1α) can increase the chemokine production and compete against the virus for binding. CCR2 also possess a mutation that alters the valine at position 64 to isoleucine. This has an impact on disease progression, where heterozygous individuals progress to AIDS 2-4 years slower, has little effect on transmission, and is most effective in early infection (Smith et al., 1997; Kostrikis et al., 1998; Mulherin et al., 2003).

Although CXCR4 and CCR5 are the dominant chemokine coreceptors used, some HIV-1 isolates have been found to utilize other chemokine receptors including CCR2b, CCR3, CCR8, Bonzo (STL33), BOB (GPR15) and GPR1 (Alkhatib, et al., 1997; Alkhatib, et al., 1997a; Bjorndal et al., 1997; Farzan et al., 1997; Horuk et al., 1998; Pohlmann,

Krumbiegel & Kirchhoff, 1999). All of these alternative coreceptors are rarely utilized, and only used in late stage of the infection. It may be that these coreceptors allow certain selective advantages, such as infecting other $CD4^+$ cells besides T lymphocytes and macrophages, and broaden the range of susceptible cells and site of infection. However, there have been little evidence of these advantages from *in vivo* studies as of yet.

1.3.2 Fusion mechanisms

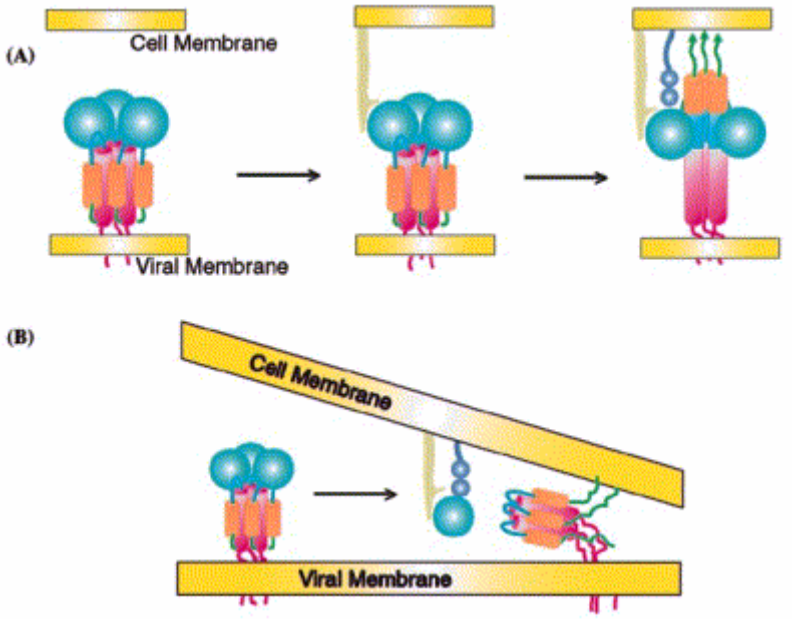
Even though the three components that are required for viral entry have been known for some time, the precise mechanism that cause fusion between the virus and host cellular membrane remains unclear. The initial steps of interactions between gp120, CD4 and chemokine receptor are fairly well understood. Study by Kwong et al. (1998) showed that 22 CD4 and 26 gp120 amino acid residues are required for interactions. Approximately 3 CD4 molecules needs to bind to a gp120 protein to induces a conformational change, where the V1/V2 loop are displaced to fully reveal the chemokine receptor binding site. Four to six chemokine receptors are then needed to bind to complete the conformation change of gp120. From this point, the specifics are somewhat unclear. The most widely accepted model at present is based on the fusion mechanism of the influenza virus. This is plausible as the ectodomain of gp41 possess certain structural similarities to the fusogenic form of haemagglutinin, the protein responsible for fusion on the influenza virus membrane (Turner & Summers, 1999; Kilby & Eron, 2003). Both proteins contain a central, three-stranded parallel coiled-coil, and each monomer is made up of two helices in anti-parallel arrangements.

When both CD4 and chemokine receptors have bound to gp120, and conformation change has occurred, it provides energy to gp41. The hydrophobic fusion peptides at the amino terminus of gp41 use the energy to insert itself into the host cell membrane, making the glycoprotein integral to both the viral and cellular membrane. The trimeric coiled-coil then bends back on itself, forming a six-helix bundle where the gp41 fusion peptide and the transmembrane domain are at the same end. The folding of the coiled coil thus brings the two membranes into proximity where the viral bilipid membrane mixes with those of the host membrane, and allows the content of the virion entry (figure 1.4) (Jardetzsky & Lamb, 2004). Other viruses, including orthomyxovirus, filovirus, and paramyxovirus, also utilize similar fusion protein to gain entry. However, the fusion mechanisms mechanism used by the other viruses are pH-dependant, using acid from endosomes to affect and enhance the process of lipid mixing between the two membranes, whereas the HIV-1 fusion process is pH-independent. Furthermore, there are other structural differences that seem to inhibit the necessary orientations of the gp41 ectodomain for fusion following the haemagglutinin model. An alternative model has thus been suggested, where upon the successful binding of CD4 and the chemokine coreceptor, the virus shed the gp120 proteins. This allows gp41 to orientate itself parallel to both membranes, and allows the insertion of the 'fusogenic peptides' into the cellular membrane (Caffrey et al., 1998; Turner & Summers, 1999).

1.3.3 Post entry cell-HIV interactions

Once the virus fuse with the cells through the use of CD4 and chemokine coreceptors, it introduces a reverse transcription complex into the cytosol of the cell. The viral reverse

transcriptase (RT) then produce viral DNA in association with the complex. The viral genome, as well as the viral proteins integrase (IN), reverse transcriptase, matrix (MA) protein, and Vpr protein are incorporated in a preintegration complex (PIC) and imported into the nucleus. Within the nucleus, the viral DNA is covalently integrated into the host genome by the viral integrase. This can be considered the early phase of the HIV life cycle, as the process may pause here indefinitely depending on the status of the infected cell (Zack et al., 1990; Saksena & Potter, 2003).



<u>Figure 1.3</u>: Potential mechanisms for CD4 and chemokine receptor-induced fusion of the viral and cellular membranes. (a) Spring-loaded mechanism similar to that proposed for hemagglutinin, where conformational changes in the TM ectodomain lead to a major displacement of the N-terminal fusogenic peptide toward the cellular membrane. (b) Shedding mechanism, where CD4 and chemokine binding result in

the loss of SU proteins, enabling reorientation of the TM and membrane fusion (from Turner & Summer, 1999).

Replication of the virus begins with the production of unspliced and spliced viral mRNA in the nucleus, and are then exported for translation. Spliced genes coding for tat, nef, and rev are produced first, as they provide regulation for subsequent steps of translation. Rev, in particular, is important for translation, as it directs splicing, and allows the production of unspliced (gag and gag-pol proteins), as well as singly spliced (Env, Vpu, Vif and Vpr) proteins (Morrow et al., 1994).

Gp160, the envelope precursor, is synthesized in the endoplasmic reticulum (ER). The protein is cleaved into the transmembrane (gp41) and surface (gp120) proteins, which remains covalently associated. To prevent premature binding of env to CD4, which is also produced in the ER, Vpu is involved to facilitate CD4 degradation. Functional CD4 on the cell surface is also removed by endosomal degradation, initiated by the Nef protein.

Formation of immature virions relies on 1000-1200 copies of gag being assembled on the cell surface. This process is aided by the gag-pol polyprotein. The gag collective then encapsulate two copies of the unspliced viral genome. After budding, the gag-pol polyprotein is cleaved to produce enzymes as well as the matrix, capsid (CA) and nucleocapsid (NC) proteins. These structural proteins arrange themselves in the maturation process to give rise to a functional, infectious virion. The production of mRNA, translation, actions of inhibitory proteins, and eventual production of mature,

infectious virions can be considered to be the late phase of the virus life cycle (Morrow, Park & Wakefield, 1994; Turner & Summer, 1999). This section is presented graphically in figure 1.4.

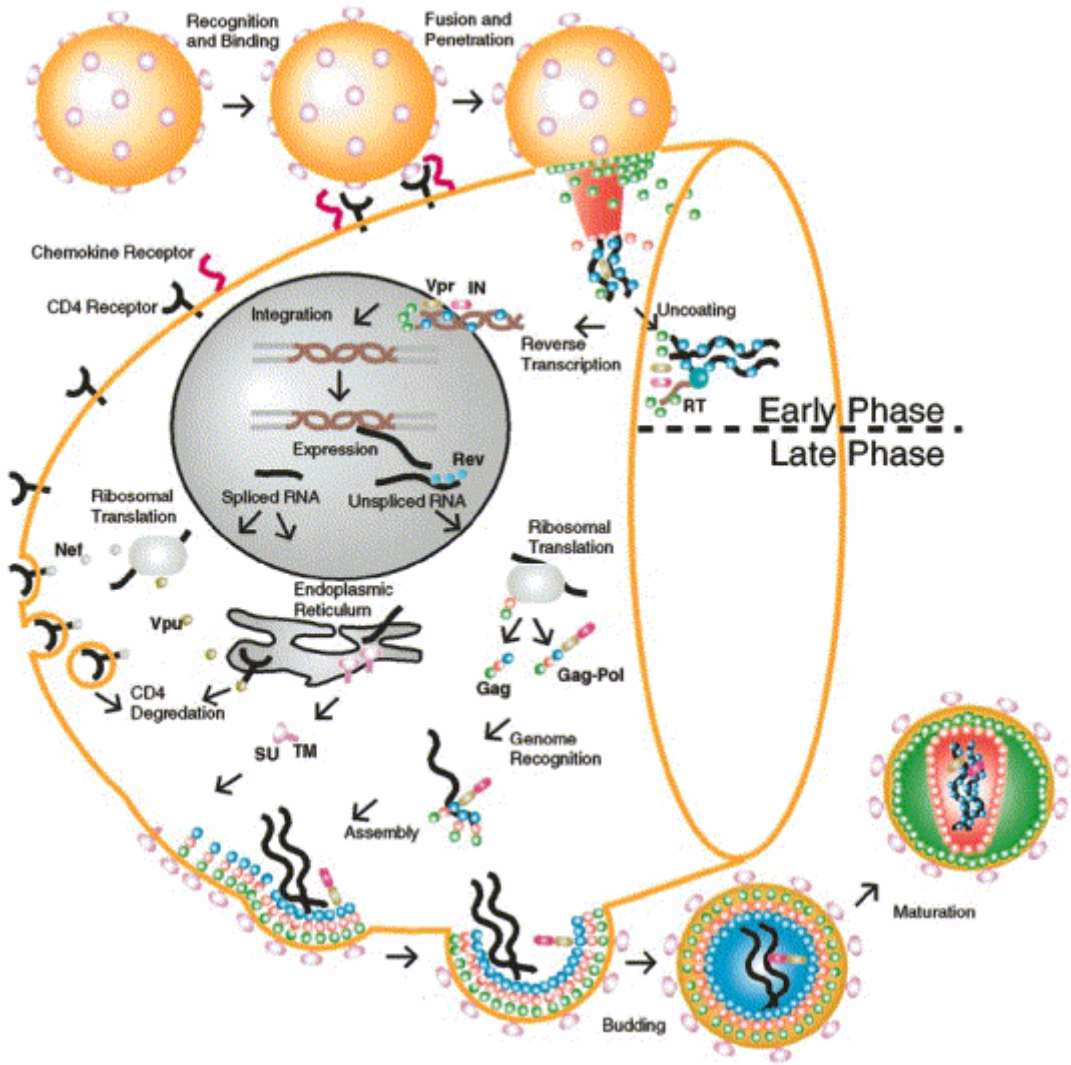


Figure 1.4: a schematic of the basic steps in HIV life cycle (From Turner & Summer, 1999).

1.4 Interactions with host cells

On the molecular level, HIV poses a number of interesting questions in its interaction with various cell populations of the immune system. These unresolved issues has led to the problematic definition of the mechanisms by which the virus cause AIDS, and while hypotheses has been proposed to address them, none have yet to be shown to be fully accurate under close scrutiny, or backed up by sufficient experimental/clinical data.

1.4.1 Macrophages and T cells: primary targets of HIV-1

The first and perhaps the most important question is what is the extent of interaction between the virus and the $CD4^+$ cell populations. It has become a given that HIV infect $CD4^+$ cells – both the macrophage and T cell populations – but the consequences of the infection vary quite drastically in these two populations.

Macrophages are any mononuclear phagocytes that are present in tissues, which process and present antigens to T cell receptors. They arise from hematopoietic stem cells in bone marrow, and pass through a number of immature stages (monoblast, promonocyte, monocytes) before entering the blood stream. The monocytes circulate for about forty hours when they then enter the nearest tissue, and fully mature into macrophages. Due to this very general classification, there are a large number of cells that can be classified as macrophage, while possessing more specific names on accounts of their function, location in the body, or morphology. These include histiocytes, kuppfer cells, osteoclasts, peritoneal and alveolar macrophages (NDI foundation, 2003; MeSH database, 2004). But

in such a broad spectrum population, only a small percentage of macrophage is infected (Chun et al., 1997). At any one point during the infection, macrophages represent no more than 10% of productively infected cells.

In theory, infected macrophages is able to produce virus throughout their existence (Blankson, Persaud & Siliciano, 2002). They also do not appear to vary greatly from uninfected cells, as their half-life remains consistent to normal macrophages, and suffer less cytopathic effects *in vitro* (Ho, Rota & Hirsch, 1986). However, the presence of the virus appears to influence macrophage activity, causing decreased efficiency of phagocytosis, and perhaps the initiation of consequent immune responses (Crowe, 1995).

Even though (infected) macrophages seems to suffers little in the presence of the virus, the interaction of infected macrophages with other cells of the immune response do have undesirable consequences, particularly in light of the fact that as an APC, it is involved in nonspecific immune responses and would present the virus to vulnerable $CD4^+$ T cells. Furthermore, it can produce chemokines like MIP to attract more monocytes, macrophages and T cells, to enhance the response to the pathogen. In HIV infection, this merely provides more viable targets for the virus to attack. Furthermore, infected monocytes-derived macrophages (MDMs) have been shown to be the culprit in introducing the virus to the cells of the central nervous system (CNS). In particular, the virus is able to invade the microglial cells that express all the necessary receptors (CD4, CCR5/CXCR4), albeit at lower level than the peripheral cell populations. The invasion of CNS is the cause of HIV-related encephalopathy in children, and AIDS dementia

complex (ADC) or HIV-associated dementia (HAD) in adults (Crowe, 1995; Martin-Garcia et al., 2002). These disorders can proceed from mild symptoms like peripheral neuropathies, to full-blown dementia in advance stages of HIV infection/AIDS if unchecked by antiviral drugs.

The infected cells in CNS are usually considered to be replication competent reservoirs. Neuropathogenesis from HIV infection also appears to be the doings of the brain macrophages rather than the other CNS cell populations. Activated brain macrophages releases a number of cytokines and chemokines that are considered neurotoxic, including tumour necrosis factor (TNF) –α and -β1, interleukin (IL)-1 α and β, arachdonic acid metabolites, and N-methy-D-aspartate (NMDA) agonists like glutamate, cysteine and quinolinate. Viral proteins released from infected macrophages also have a similar effect. The neurotoxins stimulate astrocytes, causing release of glutamate, leading to neural injuries, and the injuries cause the neurons themselves to release more glutamate, thus propagating the cycle (Crowe, 1995; Martin-Garcia et al., 2002).

$CD4^+$ T cells are the other dominant cell population that is targeted by HIV, and is much harder hit as the disease progresses. While macrophages are little affected by the virus, $CD4^+$ T cells number plummets dramatically, and is associated with clinical immunodeficiency at the end of viral infection. T lymphocytes are involved in cell-mediated immunity, and are divided into subsets according to their functions: helper/inducer, effector/suppressor, and cytotoxic cells. The $CD4^+$ T cells are of the first group, involved in major histocompatibility class (MHC) II restricted interactions, as well

as the induction of a number of major immunological responses (see figure. 1.4). Defects or depletion of T cells would have severe repercussion on the factors they secrete. The helper/inducer T cells are also divided into two subsets. T helper 1 cells secrete interleukin-2 (IL-2), gamma-interferon (IFN-γ), and interleukin-12 (IL-12). They are able to eliminate antigen-presenting cells, possess lymphokine-mediated effector activity and support cellular immune responses. Th2 cells secrete interleukins-4, -5, -6, and -10 (IL-4, IL-5, IL-6, IL-10). This allows the cells to influence B-cell development, antibody production and able to enhance humoral responses (Douglas et al., 2003; Fauci, 1988; Sarih, Maataoui & Benlimane, 1996; MeSH database).

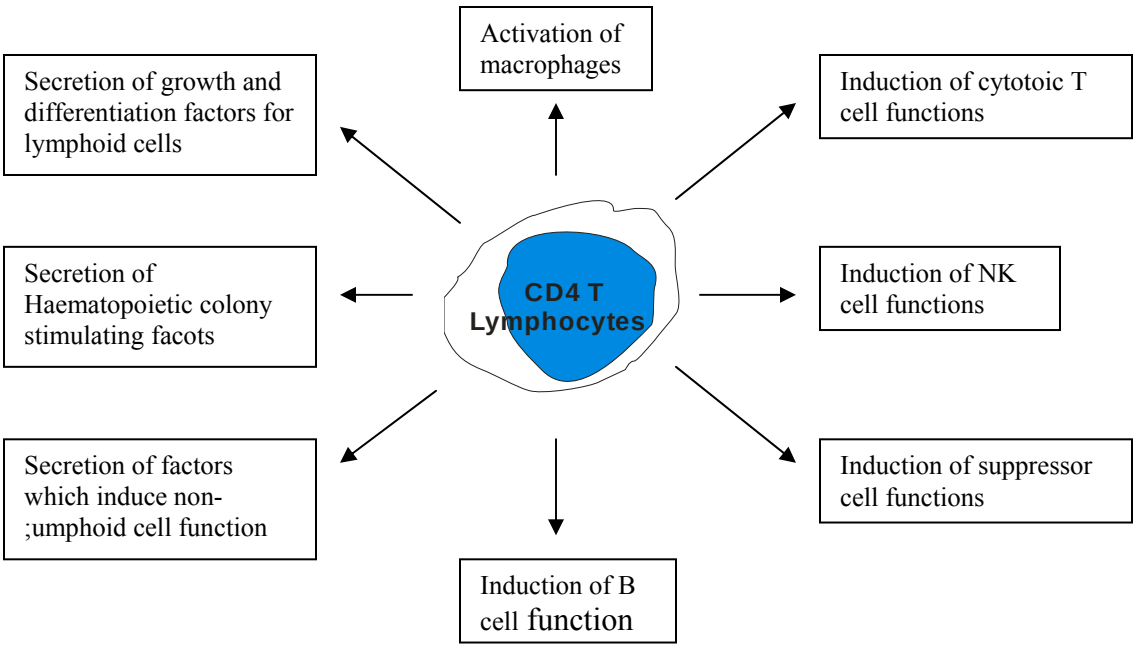


Figure 1.5: CD4+ T cells' role in the immune system. It is critical involved in induction of a number of immune responses, as well as certain non-lymphoid cell function. These functions are mainly achieved by T cell secretion of soluble factors. (Adapted from Fauci, 1988).

It has been suggested that Th1 cells are more adversely affected than Th2 cells in HIV infection, a situation detectable even before the CD4 counts begin to decline (Clerici et al., 1989; Imami et al., 2002). The depletion of Th1 cells may explain the inability of the immune system to deal with the virus. Cell mediated response is the primary pathway to deal against viral infection. Without Th1 cells, elimination of virus would obviously become more difficult. However, there are contrasting data to this observation, which showed no significant change in cytokine expression to favour those secreted by Th2 cells (Sarih, Maataoui & Benslimane, 1996; Douglas et al., 2003). The shift in cytokine profile during HIV infection thus remains a point of contention. There are also indications that the balanced survival of both Th1 and Th2 is necessary for long-term non-progressors (LTNP). This is not surprising as the existence between the two helper T cell populations is the norm in healthy individuals, and would allow the immune system better management of the virus.

Aside from the quantitative perturbation of T cell depletion, there also exist functional abnormalities. T cells that interact with soluble antigens are deficient from early stages of HIV infection. Those that remain are also less responsive to exposure to antigens and mitogens. The suppression of CD4 expression on infected cell contributes to the functional anomalies as well (Fauci, 1988). Interaction of CD4 with MHC II molecules is one of the defining features and function of T cells; a decline in CD4 receptors will therefore decrease the efficacy of these cells.

$CD4^+$ T cells can also be differentiated according to their active status. This in turn affects the cells' vulnerability of HIV-1. The population is usually dominated by naïve resting T cells, which express CD45RO and CCR5. These cells then circulate until they encounter antigen. They then become activated T cells, expressing CD45RA and CXCR4 (Berkowitz et al., 1999; Blaak et al., 1999). The cells also undergo blast transformation and proliferation to respond to the pathogen. A percentage of the surviving T cells then reverts to a resting state, but as a memory T cell, expressing CXCR4. A small subset of memory T cells express CCR5 also exists. The resting memory T cell is dormant until it encounters the same antigen in the future, and allows a much quicker immune response to the pathogen.

1.4.2 Mechanisms of $CD4^{\pm}$ T cell depletion

The definitive clinical classification of disease progression into AIDS is the profile of $CD4^+$ cell count in the blood as time goes by. When the infected individual has less than 200 $CD4^+$ cells per microlitre of blood, the infection is considered to have entered AIDS (a normal individual typically has a CD4 count of 1000 and upwards). But despite the fact that decline, and eventual depletion of $CD4^+$ T cells is a clear manifestation of the severity of the infection, the precise mechanism/s by which the virus achieve this remains unknown. The present group of theories in existence to answer this question can be roughly divided into two camps, those of direct elimination mechanisms versus the indirect mechanisms.

In the first scenario, it is suggested that $CD4^+$ T cells in the blood are eliminated as a direct consequence of the cells being infected by the virus, and those uninfected remains out of harm's way. The infected cells may perish through a number of mechanisms: **1**). cell lysis as new virions bud off, thus stripping the host of its cellular membrane (Zagury et al., 1986). **2**). cytotoxicity brought about by the accumulation of unintegrated viral DNA (Pantaleo et al., 1993). **3**). elimination through autoimmune clearance or cytotoxic antibodies (Fauci, 1988). **4**). It may also be possible that the replicating cells are induced into G_2 arrest and eventually undergo apoptosis induced by the viral protein Vpr (Ameisen, 1992), or apoptosis signals via the receptors (e.g. CD4 and TCR) of the infected cells (Groux et al., 1992).

Direct killing mechanisms	Indirect killing mechanisms	Alternative mechanisms
• Cell lysis • Cytotoxic effects through accumulation of unintegrated viral DNA • Autoimmune/cyto-toxic responses targeting infected cells • Apoptosis of infected cells	• Synctia formation • Autoimmune responses targeting uninfectd cells • HIV superantigen • Apoptosis of uninfected cells	• Sequestration of $CD4^+$ T cells in lymph nodes • Return of cells to lymph node followed by apoptosis • Chronic infection-induced immune activation

<u>Figure 1.6</u>: The possible mechanisms by which $CD4^+$ T cells may be depleted in an HIV-infected system.

However, these direct kill-mechanisms does not stand very well to the experimental data, which shows a very low level of viral expression in $CD4^+$ T cells in both lymph nodes and peripheral blood *in vivo* (less than 0.01% of mononuclear cells) (Harper et al., 1986; Embretson et al., 1993). This is possibly due to the fact that while HIV can enter the resting lymphocytes, reverse transcription does not proceed fully, leading to an abortive infection, which may be considered as latent infection of these cells (Adams et al., 1994). Considering greater than 98% of the lymphocytes *in vivo* at any one time are resting, it is logical that only a small fraction of the cells are capable of actively producing the virus. Furthermore, HIV principally infect activated T cells, yet the depletion of cells starts with naïve $CD4^+$ and $CD8^+$ T cells (Chen & Cloyd, 1999). Consequently, direct killing mechanisms cannot account for the drastic decrease in the $CD4^+$ T cell population.

The indirect theories suggest that the virus eliminate surrounding uninfected T cells. This could occur via **1**). synctia formation between the infected and uninfected cells (Lifson et al., 1986). **2**). autoimmune killing of uninfected cells (Golding et al., 1988). **3**). selective elimination mediated by an HIV-encoded superantigen (Finkel et al., 1995) – although it's more likely that the viral superantigen increase infection of cells, or **4**). induction of apoptosis in uninfected cells by HIV proteins (e.g. Tat).

While indirect killing mechanisms seems quite plausible in terms of mathematics, as most $CD4^+$ T cells remain uninfected, there has been little support for any of these hypotheses, both *in vivo* and *in vitro*. Synctia formation rarely occurs in patients, and then only appears in the latter part of the infection (Somasundaran et al., 1987); autoimmune

response against uninfected T cells is an infrequent event (Cham et al., 1991), and no convincing data is present to verify the existence of a HIV superantigen (Boldt-Houle et al., 1993). Although high level of tat and gp120 proteins can induce apoptosis, no strain of HIV produces such high levels *in vivo* during productive infection (Li et al., 1995; Wang et al., 1994).

In light of these unsatisfactory models, new theories have been proposed to resolve the *impasse*. Wang and colleagues (1997) showed that HIV-1 binding or entering resting peripheral lymphocytes activates the expression of L-selectin (CD62L). This is the receptor by lymphocytes for homing to lymph nodes; it has also been demonstrated to enhance the lymphocytes' homing effect in SCID mice. It is possible that due to abortive infection of resting lymphocytes, $CD4^+$ T cells becomes strongly attracted to return to peripheral lymph nodes, thus decreasing the number of circulating $CD4^+$ T cells. Cloyd et al. (2001) went one step further to suggest that the abortive infection of resting T cells not only enhance CD62L expression, but also upregulate the pre-apoptotic Fas molecule.

When these T cells enter the lymph nodes, a second signal received through the homing receptor causes them to go into apoptosis (Chen & Cloyd, 1999). A similar scenario is when antigens activate the immune response within lymphoid tissues. $CD4^+$ T cells become sequestered within these tissues while the activated $CD8^+$ T cells are allowed into circulation. Although the $CD4^+$ T cells do not actually die in this scenario, there is impression of depletion, as number in the blood decrease. This switch of dominant T cell population would therefore lead to the façade of a massive $CD4^+$ T cell depletion and the

related symptom shown as an inversion of CD4/CD8 ratio seen in the late stage of HIV infection (McCune, 2001).

Another theory is that the depletion $CD4^+$ memory T cells is driven by infection-induced immune activation. Studies of different infected groups have shown that the turnover rate of $CD4^+$ T cells is independent of the viraemia level. This suggest $CD4^+$ T cell depletion in the chronic phase of the infection is related to the overall activation and turnover of T cells, not the rate of viral replication (Grossman et al., 2002). As mentioned in the previous section (1.4.1), naïve $CD4^+$ T cells (and memory T cells that recognize the antigen) are activated by antigens. After proliferation and eliminating the antigen, the process also removes a high percentage of the activated T cells. Replenishment of resting memory T cells relies on the conversion of activated cells that has survived.

It is likely that HIV eliminates a significant percentage of the activated T cells, beyond the usual percentage that survives an immune response. This allows only a very small number to be converted into resting memory cells. Furthermore, production of naïve T cells from the thymus have been found to be reduced in HIV patients, with an even more prominent deterioration of the thymus when the X4-using virus is in play (Berkowitz et al., 1998; Berkowitz et al., 1998a). In association with the chronic immune activation in these individuals, the process of replenishing resting T cells are impaired due to a decrease of replacements. This would progressively lead to a smaller and smaller memory T-cell population for future activation, which will eventually disappear.

Even with these alternative pathways for removing the $CD4^+$ T cells from the peripheral blood, it seems unlikely that any single mechanism is fully responsible for the depletion. All these theories deal with the elimination of T cells in the blood, but do not really address the source of the T cells, the thymus. There are indications that the thymus is negatively affected in HIV infection. This becomes more prominent if it involves a X4-using virus, as mentioned above. The deficiency in T cell production, in combination with any of the mechanisms mentioned above would be a more plausible scenario for the disappearance of T cell. $CD4^+$ T cell depletion is therefore likely to be a multi-facted process, that limits production of new T cells, and at the same time, eliminate circulating T cells.

Perhaps the biggest discrepancy in the description of $CD4^+$ T cell depletion lies with the subtypes that possess no ability to enter the majority of these cells. It has been mentioned previously (section 1.1) that HIV-1 can change chemokine coreceptor usage. HIV-1 subtype is the most proficient strain to undergo this change, where around 50% do eventually become X4-using and SI phenotype. Less prevalent strains like A and E, also has a tendency to change, albeit at a lower prevalence than subtype B. This allows the virus entry into the majority of $CD4^+$ T cells and supplies accountability of the virus for $CD4^+$ T cell depletion.

This does not explain the depletion that occurs in Subtype C infections. Switch in chemokine coreceptor usage is a rarity, usually around 6-7% of the case (Abebe et al., 1999; Peeters et al., 1999). But it may be as high as 14-16%, depending on the

geographical region and sample size of the study (Tien et al., 1999; Cillier et al., 2003). If HIV relies on using CXCR4 for infecting the majority of T cells, then there is an inconsistency in subtype C pathogenesis. This leads to the question of how the virus is eliminating the $CD4^+$ T cells, and whether direct infection of $CD4^+$ T cells are necessary in this event. While this is particularly important for understanding subtype C infections, it would also be helpful to explain infection involving other subtypes that also do not under this change.

1.4.3 Other cells susceptible to HIV infection

There are also other cells that are affected by the virus, although they do not feature as significantly in the HIV pathogenesis as macrophages and $CD4^+$ T cells. $CD8^+$ T cells, which constitute the remainder of the T cells, are also affected to an extent, but not to the disastrous proportion that its $CD4^+$ counterpart suffers. Although chronic immune activation is a hallmark event in HIV infection, leading to the whole T cell population having a shorter half-life (Hellerstein et al., 1999), $CD8^+$ T cells does not appear to suffer depletion. However, $CD8^+$ T cells does become activated for longer periods of time, and also possess shorter telomeres. This may allow these cells more time to fulfil their immunological functions, as well as increase proliferation to keep up with the shorter half-life.

The population has also been found to coexpress CD4, and the level of expression can match those of mature $CD4^+$ T cells (Imlach et al., 2001). But they do not suffer the cytopathic effects from the virus (Mercure et al., 1993), which would allow better

maintenance of the population over time. While they are unlikely to be the primary target of the virus – they represent less than 0.01% of PBMCs, and their infection relies on a primary infection of $CD4^+$ T cells – infected $CD8^+$ T cells is able to transmit the virus to $CD4^+$ T cells. Thus, the possibility of infected $CD8^+$ T cells is a cause for concern, as it means the population would need to be taken into consideration for treating the disease

The dendritic cells (DCs) of the mucosal are another cell population affected by HIV-1. These cells are present on the squamous epithelial (including the epithelial and genital mucosae), and are primarily represented by the immature dendritic cells (iDCs), or Langerhans cells (LCs). They are considered to be specialized antigen capture/processing cells, and are speculated to be the first targets of HIV in sexual transmission (Lin et al., 2000). This is quite plausible in light of the macaque SIV model, where the first targets after vaginal challenge are the mucosal dendritic cells, and iDCs accounts for up to 90% of infected cells after 18 hours in the macaque (Lin et al., 2000; Izmailova et al., 2003). The iDCs' function is to acquire foreign proteins by phagocytosis or pinocytosis and reduce them into antigen for presentation on the cell surface. As they express CCR5, iDCs are drawn to sites of inflammation where chemokines like MIP-1α, -1β and RANTES are being produced. Once they picked up sufficient number of antigens (Ags), they migrate to secondary lymphoid tissues using the draining lymphatic system.

The dendritic cells mature by the time they arrive in the lymph node, now expressing co-stimulatory molecules to attract and trigger the appropriate T cells. CCR5 is downregulated at this point, whereas CXCR4 is upregulated. Either way, dendritic cells

(DCs) are viable targets throughout infection - they are more resistant to the T-tropic viruses, although reasons for this remain unknown (Rowland-Jones, 1999). Dendritic cells' role as antigen carrier has also been found to be further exploited by HIV. The process enhances the capture and presentation of the virus to the T cells, and thus further facilitates the virus dissemination in an infected system. Furthermore, integrated proviral DNA have been found in dendritic cells, suggesting that DCs may be sites of production infection *in vivo* (Zambruno et al., 1995). However, it appears that replication is incomplete, and is only productive in mature dendritic cells so as to transmit the virus to the T cells (Rowland-Jones, 1999).

The follicular dendritic cells (FDCs) of the lymphoid tissues are also affected by the virus. Follicular dendritic cells are present in the B cell area (primary follicles and germinal centres) of lymphoid tissues, and are unlike the typical dendritic cells that present antigens to T cells. FDCs only hold the antigen in forms of immune complexes on the cell surface for long periods of time and present the antigen to B cells in immune response (Entrez-Pubmed MeSH, 2003).

The FDC network in the lymphoid germinal centres was earlier mentioned as one of the reason why there is a large increase of viraemia in late HIV progression/AIDS. The decrease of FDCs is one of the main causes of lymphadenopathy. Although lymphadenopathy is a worrying symptom by itself, it is made worse by the fact that the FDCs in the germinal centres acts as a filter to retain the virus, either in immune complexes, or as virions on surface of follicular dendritic processes. This filter can limit

the amount of viraemia in circulation, but it also concentrates the virus to a region where any susceptible cells migrating through the lymphoid tissues may become infected. In addition, HIV is able to use the follicular dendritic cells for replication, thus causing the eventual disintegration of the FDC network. The disorganization of the network would thus lead to less efficient trapping of the virions, as well as the release of the previously trapped viruses (Fox et al., 1991).

Lastly, other mucosal cells, particularly in the GIT, can be infected by HIV-1. The upper GIT is an important site of viral entry in vertical transmission *in utero* (amniotic fluid and blood exposure) and post-partem (breastfeeding), as well as in homosexual transmission (oral-genital contact). Although it was found that intestinal macrophages do not express CCR5 and cannot be infected by R5 viruses *in vitro*, intestinal epithelial cells (IECs) do express CCR5 and CXCR4, and is able to support both R5- and X4-using viruses *in vitro*. But despite this, primary IECs only express CCR5, and not CXCR4. While there are no CD4 present on the IECs, as previously mentioned, the virus is able to use an alternative receptor, galactosylceramide, in order to gain entry into these cells. These IECs have also been shown to selectively transfer the R5 virus for further infections (Meng et al., 2002). The observation of an alternative receptor to CD4 raises the question whether HIV-1 can continue with infection even without its primary receptor targets *in vivo*.

1.5 HIV latent reservoir

One of the best studied and documented area in HIV pathogenesis is the latent reservoirs that are formed in the resting $CD4^+$ T cells. It is this obstacle that stands in the way of the full eradication of HIV in this age of HAART. A viral reservoir is a cell type or site within the body in which some form of the virus accumulates and persists with greater stability than does the main pool of actively replicating virus (Blankson, Persuand & Siliciano, 2002). Cells that fall into this category would be those that have been mentioned in the previous section – brain macrophage, dendritic cells, lymphatic follicular dendritic cells, mucosal cells and so forth – which are replication competent. A latent reservoir refers to cells harboring the virus, and while they are not actively producing virus, they retain the capacity to do so (Simmons & Siliciano, 2004).

The knowledge of HIV latency was initially confined to what is termed preintegration latency, where the virus enter resting $CD4^+$ T cells and produces a small amount of viral DNA. But without any activation signals (e.g. contact with an antigen), the proviral DNA loses its ability to actively transcribe after a few hours, although the genetic material may persist in the cell for up to two weeks (Bukrinsky et al., 1991; Zack et al., 1990). It was then found that the latency of the virus also exists in a much long-lived form (post-integration latency), where reverse transcription has occurred in the activated $CD4^+$ T cells and the viral DNA is integrated into the host cellular genome. The activated T cells would then return to a resting state carrying the viral genome (Chun et al 1995; Chun et al., 1999). But this event has been found to be rare, as the cells must survive *both*

cytopathic effects of the virus and cytolytic host effector mechanisms to be able to revert to the resting state. An activated, infected $CD4^+$ T cell dies quickly to viral cytopathic effects, or cytolytic T lymphocytes (CTLs), and has a half-life of one day.

Consequently, the frequency of resting T cells carrying replication competent provirus is very low: around 5 per 10^6 cells in the lymph nodes, and 7 per10^6 in the blood (Chun et al., 1997). There is, however, also a third state, where activated $CD4^+$ T cells carry the integrated virus. This can be considered the intermediate step between infection and post integration latency, and is only detectable in the late infection, while nonexistent in long term survivors (Chun et al., 1995). This seems to support the fact that activated T cells that become infected have difficulties entering the resting state due to the reasons stated above. It is only when the virus has caused sufficient damage to dthe immune system, that detectable level of latently infected $CD4^+$ T cells are seen.

The mechanisms and actions of the latent reservoir have also been explored in great detail. It was initially thought that the production of virions by the reservoirs was restricted by the status of the cell, as the rescue of the replication competent viruses requires stimuli that trigger global T cell activation. However, new data revealed that the determinant of replication may involve other factors like the nuclear preintegration complex (PIC), where the viral genetic material is actively transported into the cell nucleus in the absence of cell division (Bukrinsky et al., 1992). This could explain the replication of HIV-1 in nondividing cells such as monocytes/MDMs and dendritic cells. The slow kinetics of the resting T cells would also impede the replication, as the import of nucleotides and other

necessary component needed for reverse transcription would be severely hampered by the low energy level of the cell (Pierson et al., 2002). Events like viral mRNA splicing and subsequent nuclear export may also influence the efficiency of the latent reservoir in producing new virions. So, while reliance of HIV replication on proliferation of T cells cannot be totally dismissed, it seems likely that the limiting factors and controlling components to the latent reservoirs is multi-faceted.

An impressive amount of effort has thus been directed to determine the point at which these latent reservoirs are established, and what interactions exist between the reservoirs and HAART. Even when HAART was implemented ten days after the initial infection, post integration latent reservoirs were still established. Furthermore, once these latent reservoirs form, present therapy can do very little to eliminate them. This is due to the fact that antiretrovirals target active replication, and there are no such activities occurring in latent reservoirs. Or at least, the extremely low levels of replication occurring in these cells (also know as 'cryptic' replication), is sufficiently minimal to escape the effects of the drug.

The cells acting as latent reservoir also behave extremely different to the active infected T cells. A typically virus-producing cell and plasma virus has a half-life of 1-2 days (Ho et al., 1995; Wei et al., 1995). This can be even shorter under antiretroviral treatment, where free virus lives for less than 6 hours, while infected cells survive for about a day. None of these data apply to latent reservoirs, though. The infected resting T cells have been found to be incredibly stable, with a mean half-life of 43.9 months. With this in min,

it was predicted that 73 years of continuous HAART would be required to fully eradicate the virus. Even the most optimistic lower limit predicts around 21.5 years (Finzi et al., 1999). Although another group has estimated the latently infected T cell half-life to a much shorter 6.2 months, it would still require 6-10 years of therapy (Zhang et al., 1999).

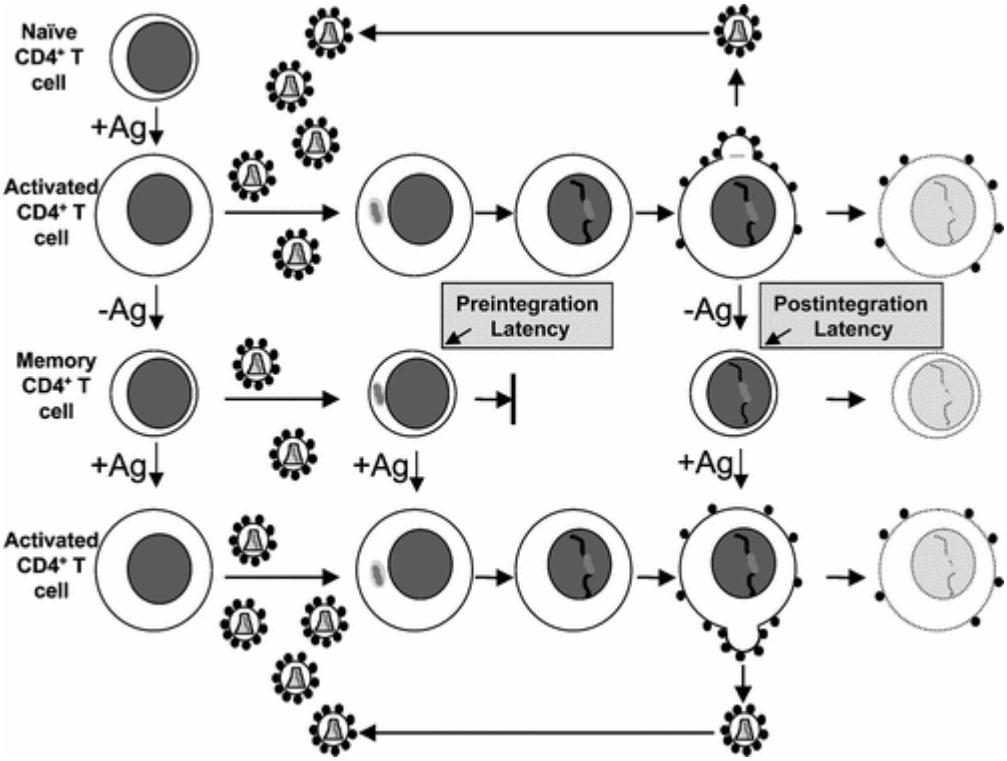


Figure 1.7: The variations of latent reservoir that can be established by HIV-1 in CD4+ T cells (From Blankson, Persaud & Siliciano, 2002). The leftmost side depicts the normal functions of the $CD4^+$ T cell, alternating between resting (small), and activated (large). R5-viruses can infect activated $CD4^+$ T cells, or some memory T cells with sufficient CCR5 density (preintegration latency), and the viral genome become integrated in the former. If the activated $CD4^+$ T cells survive to go into a resting state, postintegration latency is established.

Even if one of these data sets is accurate, one is still presuming the latently infected cells will not undergo the occasional cycles of proliferation. A typical memory T cell has a intermitotic half-life of about 6 months, but some continue to maintain the memory population through homeostatic proliferation. This can be achieved in the absence of antigen and probably without activation of the latent HIV genes. Such a process would lead to the existence of integrated virus beyond the 6 months of memory T cells lifespan (Blankson, Persaud & Siliciano, 2002). It also assumes continuous low level of replication that escape HAART would not contribute to infection activated T cells. The memory population would thus continue to receive infected cells, which replaces those that had died out. There is also the assumption that the reservoir remains small throughout the disease, as calculations were based on only 10^5-10^6 cells acting as latent reservoirs. The most recent long-term study by Siliciano and colleagues (2003), showed that even after 7 years of 'successful' treatment, there seems to be minimal decay of the latent reservoir. This is likly be the norm, rather than the exception. Aside from the seemingly impossible task of eradicating the latent reservoir, what is just as disconcerting is that these reservoirs may carry 'archival' strains from a time before treatment, and is thus able to defy treatment (Pomerantz, 2001). Therefore, it is of utmost importance that study into the latent reservoir continues, so as to determine ways to better contain, to eliminate these cells.

1.6 Fluorescent *in situ* hybridization

Fluorescent *in situ* hybridization, or FISH, was originally developed by Pardue and Gall, and John et al., independently of one another in 1969. It has become an established and useful molecular technique, and a number of variations have been introduced, particularly in conjunction with other molecular techniques, like PCR and flow cytometry. The basic principle of an in situ hybridization is to design and produce labeled – radioactive or non-isotropic labeled – nucleic acid probes. The resulting probes are placed onto the tissues or cells in question, and allowed to hybridize. The sample is then washed to rid of the excess, unbound probes, and then examined under a light or fluorescent microscope for radioactive or fluorescent probes, respectively.

The manufacture of probes and its nucleic acid depends on its target and the techniques available. For DNA probes, it can be produced as synthesized DNA oligonucleotides; nick translation to produce DNA fragments; random primers to generate random sized fragments; or PCR to make longer fragment. The production of RNA probes is a bit more standard: the sequence of interest is subcloned to have RNA polymerase promoter flank the fragment (T7, SP6, or T3). The resulting plasmid is then digested to produce varying length of RNA probes, where 200-600bp is the optimal length. There is now an alternative to subcloning to produce RNA-specific probes – PCR derived RNA-specific DNA probes. This is effectively the same principle used to produce the DNA specific probes, merely adding in the extra step of reverse-transcription PCR (RT-PCR) so as to produce the DNA template required (Cone & Schlaepfer, 1997; Knuchel et al., 2000).

The most obvious advantage to FISH is that it is an inherently specific technique, as the probes designed should only target the sequence of interest, much like PCR primers. Secondly it allows the visualization of the target in the context of its typical environment. However, it is a technically demanding process, and a number of things need to be taken into consideration before a good result can be achieved. Firstly, the preparation of certain probes types/species is cumbersome and/or time-intensive: radioactive probes are more sensitive, but as the probes are radioactive, it is more cumbersome to work as extra protections are required. Subcloning generation of the target RNA and subsequent degradation is also cumbersome. Secondly, there is a general problem of low sensitivity: *in situ* hybridization typically has difficulties detecting a target that has less than twenty copies in the cells/tissues (Knuchel et al., 2000; Patterson et al., 1993). The techniques can suffer from poor resolution, particularly if the quality of the biological sample. Methods of tissue preservation – like paraffin embedded slides – and subsequent treatment of the sample also needs to be carefully controlled to ensure good hybridization occurs.

This is particularly important when dealing with RNA-FISH. Due to the relative instability of RNA, the sample need to be as fresh as possible and process within 2-4 hours in order to preserve the integrity of the RNA. The slides then have to be stored at -70°C until needed. The technique is also time consuming and labour intensive. The probes needs to be first produced in one of the manner described above, then the slides needs to be screened to ensure that there are sufficient cells for hybridization. Optional

step may be required (e.g. then treated with DNase to remove potential background targets) to optimize the hybridization. Once the hybridization is complete, the slides need to be put away in the incubation overnight (10 hours or more) to allow maximum hybridization between the target and the probes.

FISH has principally been used to detect chromosomal aberrations, like rearrangement of genes (e.g. BCR-ABL), or extra chromosome (e.g. trisomy 21). In HIV research, FISH is rarely used by itself, and has been utilized in conjunction with PCR, and flow cytometry. These variations are principally used to overcome some of the flaws of FISH: PCR-driven FISH can amplify low copy numbers of the target sequence sufficiently to be detectable (Patterson et al., 1993), while flow cytometry can speed the counting of positively labeled cells to matter of seconds (Borzi et al., 1997). Although these combinations of techniques are useful, they are not necessary in this study. The issue of low copy number is somewhat irrelevant in the detection of actively replicating reservoirs, as any active site of replication should have an abundance of viral mRNA present in the cell. And while flow cytometry is a nice application by enhancing the counting process, manual counting is still viable here, even though it is more time intensive.

1.7 Prediction of HIV-1 chemokine coreceptor usage

The initial approaches to monitor the change from NSI to SI were to inoculate PBMC with the virus or cocultivation of PBMC with infected cells, and to evaluate the amount of synctia formation in the cultures. Alternatively, it can be determined by culturing the virus with the MT-2 cell lines (a transformed T cell line), as these cells only support the replication of the SI virus (Dong et al., 2005). Other indicator cell lines that express certain chemokine coreceptor only were also used. Their infection indicates the use of the receptors. However, these culturing assays merely provide a positive or negative result with respect to the phenotypic characteristics, and do not reflect the genetic changes which may significantly enhance understanding of the viral pathogenesis.

More recently, the gp120 of the HIV envelope protein – and specifically the V3 loop - has been found to be the key determinant in the virus' ability to induce syncytium. Position 306 and 320 in gp160 (reference virus HXB), which fall into the V3 loop (and equates to position 11 and 25 in the V3 loop), seems to possess an absolute correlation in the naturally occurring virus' capability of inducing syncytium. The NSI variants are either negatively charged or without charges, while the SI virus has positively charged amino acids at either one or both of these positions (Fouchier et al, 1992; De Jong et al., 1992). At least one of the two amino acid positions is therefore always associated with X4-tropic, dualtropic (R5X4), or the SI phenotype. This phenotype-genotype correlation is also independent of any changes in the other parts of gp120. Mutation has been found to occur in the V2 and C4 region, and may have association to the selection of chemokine

coreceptors (Hoffman et al., 2002). However, these variances that occur outside the V3 loop seem to have much less importance in comparison with the two amino acid substitution. There were also indications that a purely visual examination of the amino acid composition at positions 11 and 25 is actually one of the more dependable methods for predicting the SI and X4-using phenotype, based only on motifs (Resch et al., 2001). This defining characteristic correlates the V3 loop genotype with the phenotype, which has also come to be known as the 11/25 rule.

The second method of phenotype prediction is the Position Specific Scoring Matrix (PSSM). The PSSM is derived from a profile, which is a collection of aligned sequences that contain the domain of interest and list the frequencies of amino acid in each position for the resulting protein. The PSSM utilize this information and provides a log-odds score for finding a matching amino acid in the sequence provided, against the reference data produced by the profile. In the PSSM created by Mark Jensen at the University of Washington, the score range indicate the R5 viruses have a low value, dualtropic R5X4 viruses have intermediate score, and X4 viruses possess the highest numbers (Jensen et al., 2003). This method relies on the full sequence of the V3 loop as input in order to obtain a score on the PSSM. There is also a second PSSM – the sinsi matrix - that calculates the score in terms of NSI/SI phenotype, with a similar pattern (NSI has low value score; SI has the high value scores). The dividing range for the sinsi scoring matrix is slightly different compared to the x4r5 matrix, as the intermediate scores on the x4r5 matrix may reflect differently on the sinsi matrix.

The third and final method used for phenotype prediction was the basic calculation of the total charge of the V3 loop. It takes into account all of the amino acids and their charges. Only 5 out of the 20 amino acids possess charges at the physiological pH: arginine (R) and Lysine (K) carries a +1 charge; aspartic acid (D) and glutamic acid (E) carries a -1 charge. Histidine carries a +0.5 charge, although a number of prediction software considers it as having a +1 charge. A low overall charge of the V3 loop (typically 5 or less) is an indicator of NSI/R5 phenotype, while any V3 amino acid chain with +6 or higher charges are usually the SI/X4 phenotype (Cilliers et al., 2003).

One other region that was highly conserved in the V3 loop is the chain of 4 amino acids starting from position 6 (NNTR). This stretch is an N-linked glycosylation site, and its presence (and subsequent glycosylation of the structure) has been linked with the use of CCR5 by the virus (Zhang et al., 2004). Its absence does not indicate the virus has switched to using an alternative chemokine receptor, but if the site is present, it is highly indicative of the virus using CCR5.

Aims

The pathogenesis of HIV-1 relies on $CD4^+$ T cells for successful replication and transmission. $CCR5^+CD4^+$ cells (macrophages, monocytes, activated T cells, and small subset of memory T cells) are the virus' preferred targets, and this may later include $CXCR4^+CD4^+$ cells (resting/memory T cells). $CD4^+$ T cells have been acknowledged to account for the majority of infected cells, with activated T cells being the virus' preferential target. This population, however, typically represent a minority of T cells. In the subtype C virus, a switch in chemokine coreceptor usage from CCR5 to CXCR4 is rare, even late in infection. Nevertheless, these infections are not exempt from$CD4^+$ T cell depletion and high viral load, which are key markers in late stag of the infection, progressing towards AIDS.

In order to assert whether T cell depletion is occurring through indirect killing mechanisms, we propose to investigate the site of HIV-1 presence in these patients.

To achieve this, there are 3 main objectives:

1. Collection of peripheral blood or bone marrow samples from HIV^+ patients.

   These individuals have to fulfil the following criteria: being in the late stage of the disease (200-300 CD4 cell counts or lower), ART-naïve, and not suffering from opportunistic diseases. Antiretroviral drugs greatly reduce HIV presence, thus

treatment naivety is a necessity. Acute opportunistic infections (OIs) will cause proliferation of white blood cells, leading to a distortion of the true number of cells carrying the virus. Absence of OIs is therefore required.

2. The design of a FISH probe that is specific to HIV-1 subtype C genome, and optimize it on HIV-infected cultures.

The hybridization will then be applied to slides of HIV^+ whole blood and bone marrow, as well as separated $CD4^+$ cell from these bodily fluids. The result will be able to tell us in which cells the viral genomic materials can be found.

3. Prediction of Chemokine coreceptor usage of virus extracted from the patients by V3 loop sequencing. The sequencing will also determine the subtype of the viruses and any genotypic changes of the region.

The correlation between the FISH and sequencing data will provide a picture of where the virus is situated in the knowledge of viral chemokine coreceptor usage.

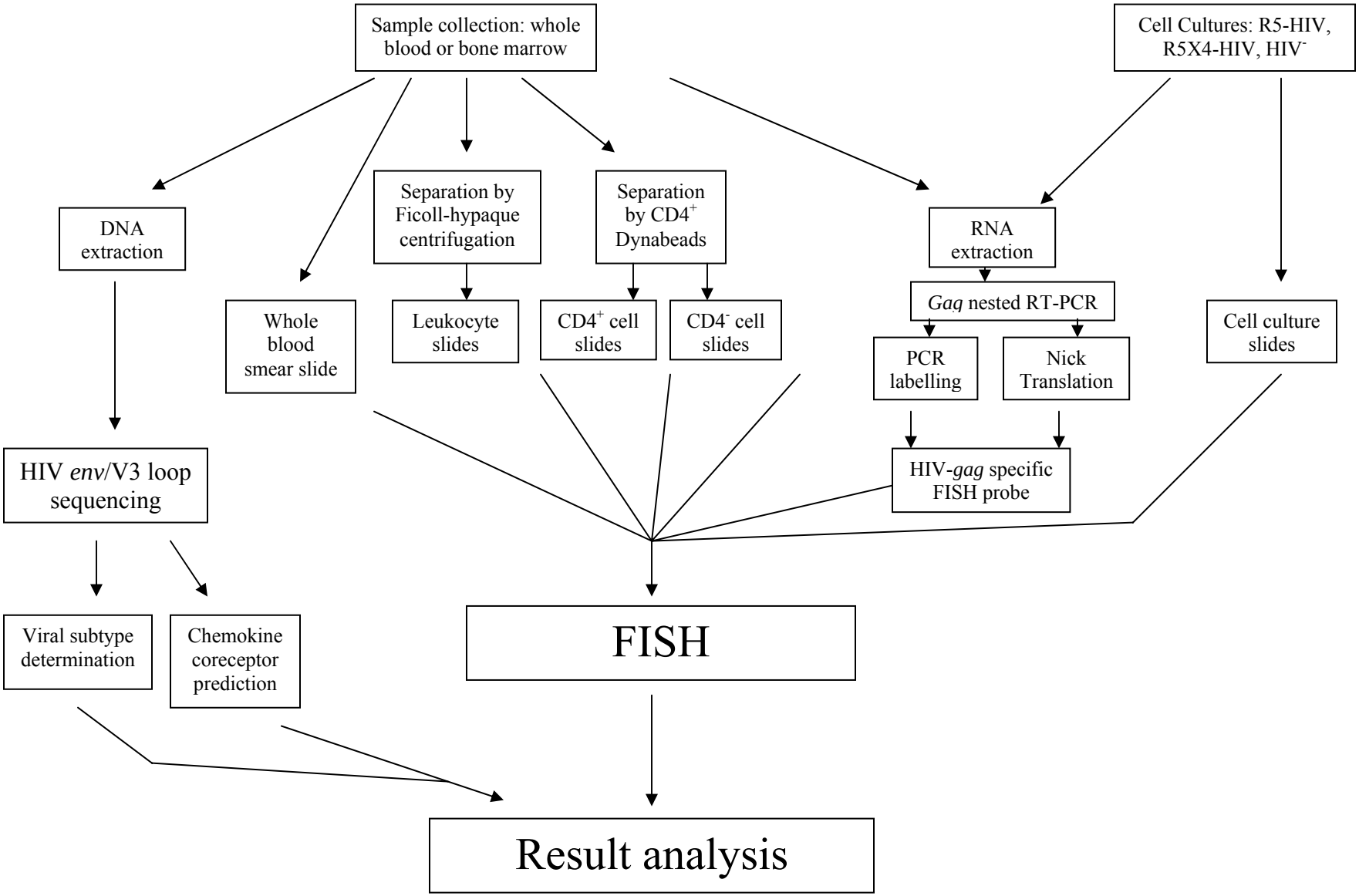


Figure 2.1: The outline of the experimental procedures for this study

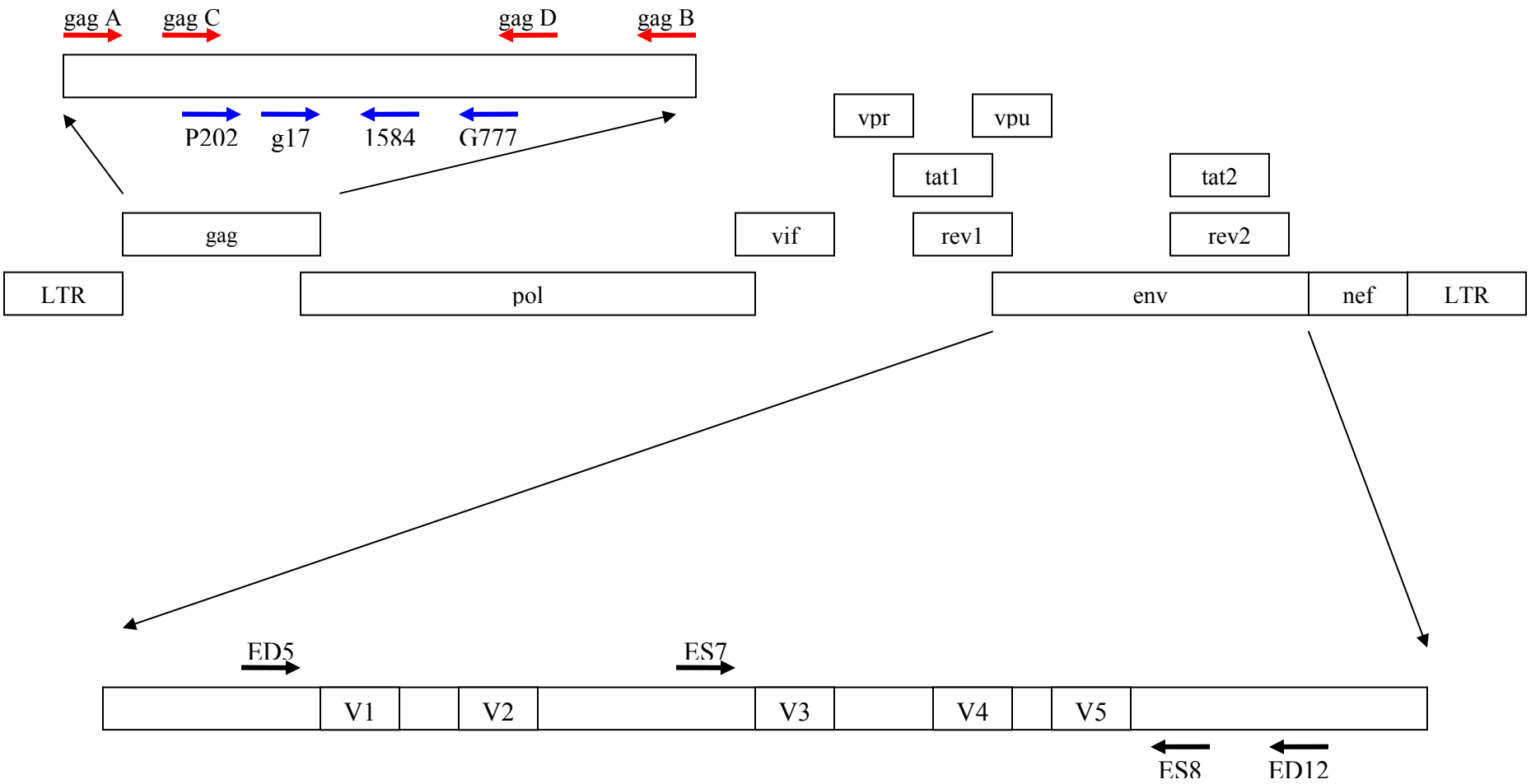


Figure 2.2: The organization of the HIV genes, with the PCR primers used in this study. The gag and env gene are magnified to show the relative positions of the primers used for producing the gag-specific probes and V3 loop sequence.

Chapter 2: Materials & methods

2.1 Sample collection, extractions, and lymphocyte isolation

2.1.1 Sample collection

Fresh blood was collected from patients at the Johannesburg General Hospital HIV clinic. Patients without any opportunistic infections, and low CD4 cell count (≤200) were asked to donate 3-4ml of blood. Once consent was obtained from an individual, blood was drawn into 5 ml EDTA tubes. The blood was processed within 2-4 hours after collection, to ensure the integrity of the RNA. Whenever possible, bone marrow extracts were collected in the EDTA tubes, following lumbar puncture, and treated in the same manner. 34 samples was collected, 28 of blood, 6 of bone marrow extract. Ethical clearance was obtained for this collection process (See appendix IV).

2.1.2 Cell Lines

These were obtained from the National Institute of Communicable Disease (HIV unit, courtesy of Mr. Jabulani Nhlapo & Mr. Tonie Cilliers). Cell lines with the appropriate chemokine coreceptor expression (CCR5 or CXCR4) were infected with virus isolate with known phenotype and chemokine coreceptor usage. KC42, and U87.CD4 cell lines with CCR5 expression was used for culturing of NSI viruses, while GHOST.3 cell line with endogenous CXCR4 expression was used for culturing SI viruses.

The cultures were processed after 4-5 days of growth, and were used for RNA extraction, or making slides.

2.1.3 RNA extraction

Total RNA was extracted from the blood of all (34) the HIV-positive patients collected and/or cell line cultures (KC42 infected with R5-virus, or U87 infected with dualtropic virus). The extraction was done using the Gentra® Purescript RNA isolation kit and follows the provided protocol (Gentra System, Minneapolis, USA). RNA was stored at -70 ºC until needed.

2.1.4 DNA extraction

Genomic DNA was extracted from blood of the patients and cell line cultures using the Gentra® Puregene DNA isolation kit (Gentra System, Minneapolis, USA). Extractions were carried out as protocol provided in the kit. DNA was stored at -20 ºC until needed.

The mixed DNA/protein pellet, remnant from the RNA extraction, can also be used to extract the DNA. Cell lysis solution can be added to the protein-DNA pellet, and the extraction continues from that point on. The process was also described in the provided protocol.

2.1.5 Ficoll-hypaque separation

Ficoll-hypaque extraction allowed the separation of white blood cells from the fresh blood. Blood was diluted using Hank's balanced salt solution at a 1:1 ratio and pipetted up and down to ensure a good mix.

A specially designed 10 ml tube was used, where the 3 ml of the bottom was divided into a compartment to house the ficoll solution.

3 ml of ficoll was added to the tube, and it was made sure there was no air bubbles trapped in the compartment. The diluted blood was then placed it on top of the ficoll. The blood was spun at 400 g (1410 rpm) for 40 minutes at 18-20 ºC. The white blood cells were visible as a thin layer between the plasma and the ficoll. The white blood cells were removed by pipetting and placed in a clean tube with as little of the two layers as possible, as the plasma could cause platelets and plasma protein contamination, while ficoll had granulocytes. At least 3 volumes of the Hank's balanced salt solution were added to the lymphocytes and the cells were resuspended by pipetting. The sample was centrifuged at 60-100x g for 10 minutes at 18-20 ºC, and supernatant was then removed. 6-8 ml of Hank's balanced salt solution was added and centrifuged once more at 60-100x g for 10 minutes at 18-20 ºC. Supernatant was removed once again and the cells were resuspended in 1ml of Hank's Balanced Salt.

The suspension was used make slides directly after the separation. The remaining suspension was then stored at -70 ºC.

2.1.6 Dynabead cell separation

The Dynal (Invitrogen) Dynabeads were superparamagnetic polystyrene beads manufactured to the same size (4.5 μm±0.2 μm) and labeled with specific antibodies to isolate certain cell populations in the blood. In this case, Dynabead with CD4 monoclonal mouse antibody were used to isolate the $CD4^+$ cell populations from blood of HIV^+ patients, as the ficoll-hypaque method was to be lacking – there were inefficient recovery of the white blood cells.

As per the provided protocol, the Dynabeads were washed, mixed with blood and used to separate the $CD4^+$ T cells from the peripheral blood. The separated cells were stored in culture medium at 4 ºC until needed.

2.1.6.1 DETACHaBEAD

There was an alternative to leaving the cells attached to the Dynabeads, which was particularly useful, as the beads may interfere with any subsequent processes after cell separation. The variation was referred to as DETACHaBEADs. The DETACHaBEAD CD4 solution contained a polyclonal anti-Fab antibody specific for the CD4 antibody on the Dynabeads. When added, the DETACHaBEAD competed with the antibody-antigen binding at the cell surface and released the antibody and beads from the cell. The CD4 cells were left viable, unstimulated and without surface antibodies (Dynal Biotech, 2004). The separation of the cells and removal of Dynabeads were performed in lines with the provided protocol.

2.2 Fluorescent in situ hybridization (FISH)

FISH involved a number of procedures to complete, including slide preparation, probe manufacture, optimization. This was followed by the hybridization process itself, and the subsequent analysis under the fluorescent microscope.

2.2.1 Slide preparation

2.2.1.1 Cultured cell-line slides

Cells suspended in culture medium were spun down at 15 000x g for 20 seconds. Majority of the supernatant was removed, and cells were resuspended through vortexing. The cells were dropped from 20-30cm above the slide and air dried. The slides were fixed in methanol for 30 minutes at -20 ºC, and dehydrated in an ethanol series (70%, 90%, 100%) for 5 minutes each at room temperature. The slides were allowed to air dry and then stored away at -70 ºC until needed.

2.2.1.2 Slides from blood and separated fractions

Slides were made from the fresh blood obtained from the patient, as well as from separated lymphocytes. In the first instance, blood was placed onto a clean slide and then lightly smeared over the whole surface using the edge of a second slide. The slide was allowed to air dry, and then fixed in the following solutions: methanol for 10 minutes at room temperature, methanol for 30 minutes at -20 ºC, fix (methanol:acetic acid mix, 3:1

ratio) for 1 hour at -20 ºC. The slides were then dehydrated and stored in the same manner described in 2.2.1.1.

For both ficoll-separated and Dynabead-separated cells, the cell suspension was dropped onto a clean slide from some distance above the slide, and then steamed for 60-80 seconds. Alternatively, the cell suspension smeared over the surface of the slide in the manner described for the blood smear slide. They were allowed to air dry, and then dehydrated and stored as mentioned in 2.2.1.1.

2.2.1.3 DNase I treatment

Before the slides can be used for FISH, it was necessary to treat them with DNase I to remove as much of the DNA as possible so as to eliminate excess background signals. 10 µl of DNase I (10ng/ml) was first heat treated at 95 ºC for 10 minutes, and then diluted with 990 µl of Sabax water. 10-15 µl of the dilution was placed onto a coverslip and then placed onto the slide. The slide was then incubated upside down (area with coverslip on the bottom side) for 30 minutes at 37 ºC. The coverslip was then removed by gentle tapping, and the slide was washed in 1X PBS to rid the DNase I enzyme.

2.2.2 Probe manufacture

2.2.2.1 Reverse-transcription Polymerase Chain Reaction (RT-PCR)

RT-PCR was required to produce the cDNA needed for manufacture of the FISH probe. RT-PCR uses total extracted RNA and amplifies the viral mRNA; more specifically, part

of the gag gene. The gag gene was chosen based on the reason that it is one of the few viral genes that is has a low mutation rate throughout all the subtypes and variants of HIV. Two different sets of primers were used to produce two gag-specific probes. These were used separately during hybridization. The Qiagen one-step RT-PCR kit (Qiagen, USA) was used to obtain the cDNA (see appendix I for primer sequences and RT-PCR parameters).

RNA obtained from cell cultures (mainly KC42 infected with R5-using viruses; R5X4 were also used) were used as template for RT-PCR and subsequent probe production.

2.2.2.2 Polymerase Chain Reaction (PCR)

PCR was used to in terms of PCR labeling for incorporation of fluorescently labeled nucleotides into the gag-specific fragments. It was also used to produce the alternative gag fragment used in nick translation.

The probe production was adapted from a nested PCR protocol (courtesy of Mia Coetzer, HIV unit, NICD). The PCR target the gag gene of HIV-1, and was able to amplify consistently for R5-using virus as well as the X4 strain. The first round PCR produces a ~1127 bp fragment, and the nested PCR gave rise to a ~466 bp fragment.

The second gag primer sets was also a nested PCR, courtesy of Dr Maria Papathanasopoulos (Department of haematology & molecular medicine, Wits medical school). The first round amplified a 2.5 kb fragment, and the second round produces a

fragment about 1.2 kb in size, spanning the complete gag gene. (See appendix I for primer sequences and PCR protocols).

Both sets of primers were first tested using RNA obtained from both R5 and R5X4 cell cultures, and then subjected to PCR labeling or nick translation to produce the necessary probes. Negative controls with sabax water replacing the RNA/DNA were run cocurrently to the samples to detect for contamination.

2.2.2.3 PCR labelling & probe precipitation

The labeling process was a straight adaptation of the nested round for the first set of gag PCR. The only alteration was the volume of FITC-dUTP (optimized to 0.5-0.6 µl per PCR reaction) was added, and corresponding volume of water was subtracted from the total volume.

The labeled fragment is visualized in the same manner as normal PCR (see appendix I). The product would be the same size, although it would now be a visibly green-yellow colour. There were typically bands of fluorescence towards of the bottom of the run, brought about by the unincorporated fluorescent nucleotides.

The probe was then precipitated to concentrate it and maximize its efficiency. 1/10 volume of the sodium acetate (pH 3.5) and 2.5x volume of 100% ice cold ethanol was added to the product and then allowed to sit at -70 ºC for at least 30 minutes. It was spun at 13 000x g (or higher) for 30 minutes at 4ºC. The probes should be visible as a small

yellow pellet at the bottom of the tube at this point. As much of the supernatant was removed from the tube without disturbing the pellet; 100 µl of 70% ice cold ethanol was then added and slowly pipetted up and down to wash the pellet. The probe was spun at maximum speed for 10 minutes at 4 ºC. The supernatant was removed, and the probe was allowed to air dry for 10-30 minutes. The probe was resuspended in a volume of hybridization buffer equal to the volume before precipitation. It was then left overnight on a shaker for homogenization in the dark. The pellet may still remain after the night, but would be easily dispensed with by gentling pipetting the solution. The probe was stored at -20 ºC and aliquoted when needed.

The precipitated probe used per slide was optimized to 2.5-3.0 µl, diluted with a further 10 µl of hybridization buffer per slide.

2.2.2.4 Nick translation

The second set of nested *gag* primers amplified the complete *gag* gene, which is about 1.5 kb in size.

The new probes were tested using the same protocol as before. While it detected the viral genetic material, it was inefficient in penetrating the cell nucleus. Thus the probe manufacture, and hybridization conditions were altered in parts to try and accommodate the entry and hybridization of the new probe (For a complete summary of the troubleshooting process, please see Appendix A).

Nick translation was the chosen method for labelling and manufacturing FISH probes from the second sets of *gag* primers. FITC was substituted for Spectrum Green – dUTP in this process. Dnase I, DNA polymerase I, dNTPs, Spectrum Green – dUTP, and appropriate buffers were incubated (see appendix I). The fragment size was checked on a 2% agarose gel. If it was satisfactory, the probe was inactivated. Precipitation of the probe was in the same manner as previously described, with the additional input of human C_0t 1 DNA.

2.2.3 FISH

2.2.3.1 Hybridization

The slides to be hybridized were removed from the -70 ºC, allowed to thaw at room temperature, and then subjected to the following conditions for FISH.

Denaturation of slides (70% formamide):	5 minutes @ 76 ºC
Dehydration (70%, 90%, 100% EtOH):	3 minutes each @ RT
Denaturation of probe:	5 minutes @ 80 ºC

The slide was allowed to air dry a little before the probes were applied to the coverslip and placed onto the slide.

The coverslip was sealed with rubber cement and allowed to set, then the slides were placed into an incubator at 37 ºC. Incubate and hybridize overnight (10-16 hours).

2.2.3.2 Washing

The coverslip was removed from the slide by gentle tapping on the side and the slides were washed in the following:

50% formamide	5 minutes @ 42 ºC, 3 times
2X SSC:	5 minutes @ 42 ºC
2X SSC + TWEEN 20:	5 minutes @ 42 ºC
2X SSC + DAPI:	15 minutes @ RT
2X SSC + TWEEN 20:	2 minutes @ RT

The slide was allowed to air dry, and an anti-fading reagent (Vectorshield Mounting, Vecotr laboratories) was applied to the coverslip and then placed over the area of hybridization. Rubber cement was used to seal the coverslip and allowed to set.

2.2.3.3 Slide Analysis

Once the rubber cement has set, the slides were examined under the fluorescent microscope (Olympus). The cells were observed under 60X magnification, and 100X magnification. The images were also captured using the Genus imaging software (Applied Imaging, USA).

2.3 envelope gene V3 loop sequencing

DNA extracted from patient blood (28) and bone marrow (6) were subjected *env* nested PCR. The first amplified regions of the env protein from V1 to V5; the nested PCR produce a product that is approximately 600 bp in size, spanning V3 to V5 of the *env* gene (see appendix I for primer sequence and protocol). The fragment was visualized on 2% agarose gel to confirm its presence before moving to the next step. Two PCR reactions were performed for each sample, with a forwards or reverse primer to obtain complementary fragments that can be combined during sequencing analysis. All the samples were processed in this manner.

The products were put through chain termination to produce the varying lengths of labelled oligonucleotide chains, and then cleaned up using the sodium acetate method. The DNA was then placed through an ABI prism sequencer (ABI systems, USA) and the raw data (chromatogram) of the procedure were obtained. The data was cleaned up and analyzed using the Sequencher software (Gene Codes, USA). BioEdit© (obtained from http://www.mbio.ncsu.edu/BioEdit/bioedit.html), a sequence alignment editor, was used further clean up, and to convert the nucleotide sequences to amino acid.

2.3.1 Phylogenetic analysis

Coherent nucleotide and amino acid sequences were converted into FASTA (.fsa or .fas) format, and subjected to a number of software for the determination of the subtype as well as to check for any possible contamination in sequencing. The sequences were

initially aligned using the Clustal X software (obtained from ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/), and realigned to sequences from other major subtypes using BioEdit software. BioEdit was also used to determine the areas of ambiguity, i.e. where there may be a deletion or insertion in relation to the reference sequences. The Sequencher® sequence alignment software (Gene Codes, Michigan, USA) for the final cleaned up and alignment of the forward and reverse sequences. The cleaned data showed 7 samples had insufficient areas of the forward and reverse strands complementary to each other. This did not allow alignment to create a coherent sequence that spanned the complete V3 loop. Consequently, only 11 samples produced coherent V3 loop sequences for analysis.

The fully edited sequences were put through the MEGA version 3.0 (Kumar, Tamura, Nei, 2004) for phylogenetic analyses, as well as to test for cross-contamination in the sequencing process. Maximum parsimony neighbour joining tree with bootstrapping was used to create a phylogenetic relation of the samples with respect to their V3 loop sequence.

The sequences were first tested for phylogenetic relations with a number of reference sequences. The references included other major subtypes from the M clade (A, B, D, G) as well as the sequences of two simian immunodeficiency virus (SIV) isolates from chimpanzees. Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 3.1 (Kumar, Tamura, Nei 2004). Maximum parsimony neighbour joining

tree with bootstrapping was used to create a phylogenetic relation of the samples with respect to their V3 loop sequence. Bootstrapping of 250 times was also applied.

2.3.2 11/25 amino acids

2 amino acid positions of the V3 loop, 11 and 25, are useful indicator of the chemokine receptor usage. Amino acids with basic charges (arginine, lysine, or histidine) found at either of these positions increases the overall charge of the V3 loop and are indicative of the virus being X4-using (Cilliers et al., 2003).

2.3.3 Position Specific Scoring Matrix (PSSM)

The amino acid sequence were put through the Position Specific Scoring Matrix, an online program (http://ubik.microslu.washington.edu/computing/pssm/) that calculate the likelihood of the samples being the two phenotypes (NSI or SI), or the tendency of using CCR5 or CXCR4. The PSSM is derived from a profile, which is a collection of aligned sequences that contain the domain of interest and list the frequencies of amino acid in each position for the resulting protein. A log-odd score is calculated be matching the amino acid sequence of the sample to the profile. In the PSSM created by Mark Jensen at the University of Washington, the score range indicate the R5 viruses have a low value, dualtropic R5X4 viruses have intermediate score, and X4 viruses possess the highest numbers (Jensen et al., 2003).A alternative PSSM can be used calculates the score in terms of NSI/SI phenotype, with a similar pattern (NSI has low value score; SI has the high value scores). The dividing range for the sinsi scoring matrix is slightly different

compared to the x4r5 matrix, as the intermediate scores on the x4r5 matrix may reflect differently on the sinsi matrix. The only input required is the amino acid sequence in FASTA format. Following submission, the program would produce a table (R5X4 matrix or SINSI matrix) with the relevant data.

2.3.4 Calculation of overall charge of V3 loop

5 out of the 20 amino acids are basic or acidic at the physiological pH, and thus can influence the overall charge of the V3 loop. These are arginine (R), lysine (K),aspartic acid (D), glutamic acid (E), and histidine (H). The first two are positively charged (+1), the next two acidic amino acids are negatively charged (-1), while histidine possess a 0.5 positive charge (although some software consider it to be +1). The overall charge of the loop can be calculated by taking account of the presence of all 5 of these amino acids and obtaining the negative or positive charge by the simple addition/subtraction. R5-using viruses typically possess +3 to +5 charge, whereas dualtropic and X4-using virus has +6 to +9 net charge (Cilliers et al., 2003). The process was applied to all the samples that produced a coherent V3 loop sequence to calculate their charges.

Chapter 3: Results

3.1 PCR products and labelling reaction

3.1.1 Gag PCR

The first round amplifies a 1.2 kb fragment, which could not be detected on agarose gel.

The second round/nested PCR produces a fragment approximately 466 bp (dependent on possible mutations), which can be visualized on a 2% agarose gel containing ethidium bromide.

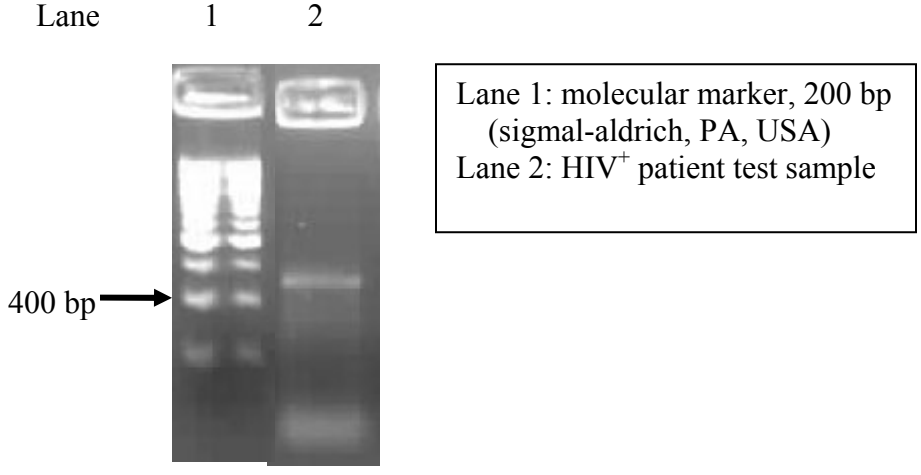


Figure 3.1: 2nd round product of the gag nested PCR. Lane 1 was the molecular marker (200 bp). Lane 2 was a HIV$^+$ patient sample, placed between 400-600 bp.

3.1.2 Gag PCR II

The second set of nested gag PCR amplifies the complete gag gene in the first round, which was about 2.5 kb in length, and could not be detected on agarose gel.

The nested PCR gave rise to a ~1.2 kb fragment, which was visualized on a 2% agarose gel with ethidium bromide.

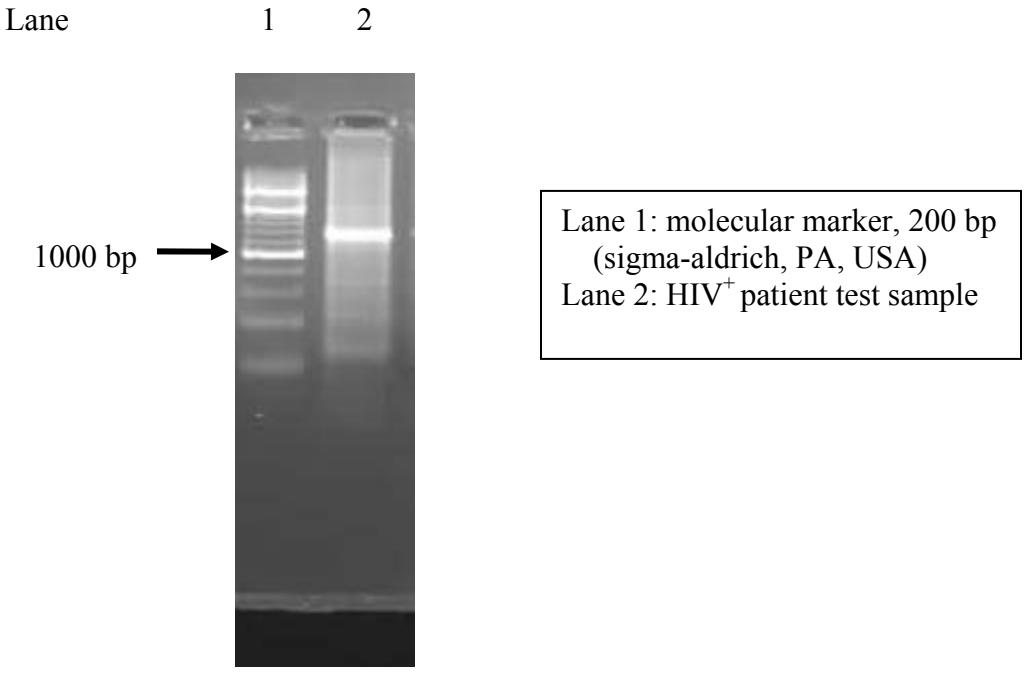


Figure 3.2: 2nd round product of the alternative gag nested PCR. Lane 1 was the molecular marker with 200 bp difference between each band. Lane 1 was a HIV$^+$ patient sample

3.1.3 Nick Translation

The random cutting activity of the DNase I used in nick translation meant that the second round product would be reduced to fragments of smaller sizes. Although there is a set time period (75 minutes) for nick translation, enzyme activity would vary from sample to sample due to DNA concentration. Therefore, it was necessary to visualize the size of the products on a 2% agarose gel with ethidium bromide before using them for FISH probes.

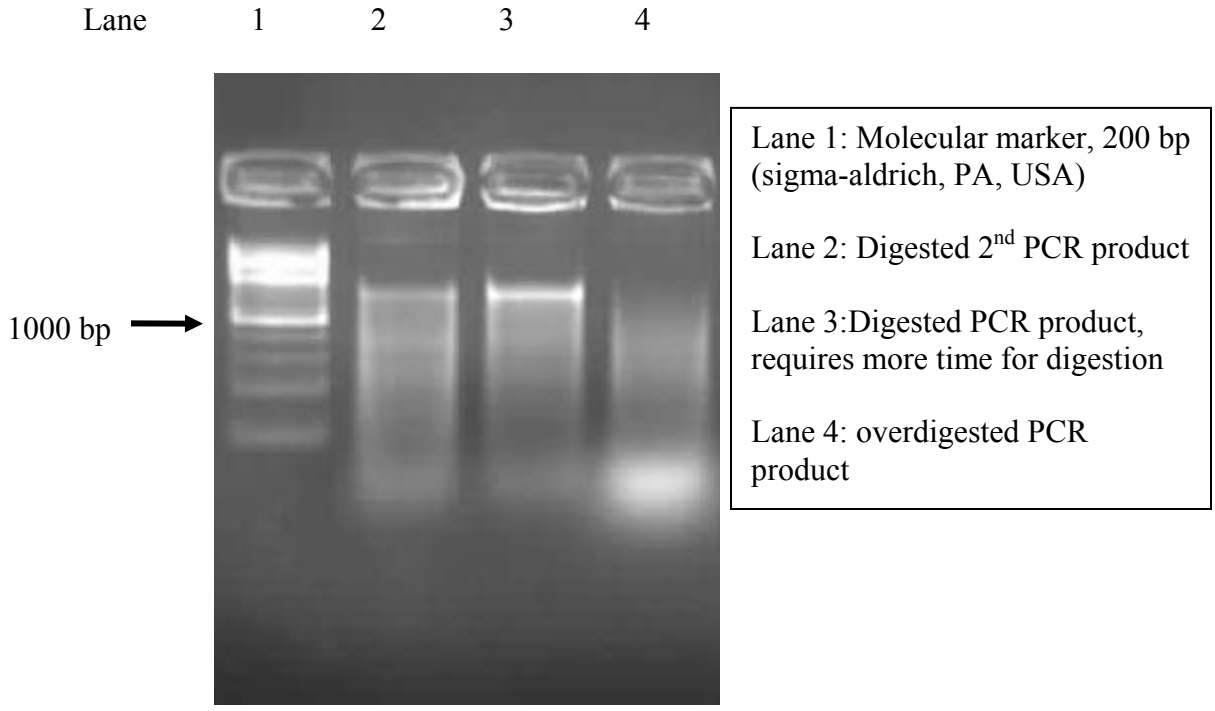


Figure 3.3: 2nd round gag products after digestion by DNase I. Lane 1 was the molecular marker, 200 bp difference between each band. Lanes 2-4 were digested samples.

In nick translation, the optimal size of digested DNA fragments would be between 200-500 bp. below 200 bp, the probe would lose its specificity, and become prone to stick to the glass slide, rather than entering the cells. Over 500 bp, the probe becomes too big to efficiently gain entry into the cells.

In the digestion of the nested products in figure 4.4, lane 2 was the best of the three, as there were minimal amount of the original fragment left, and a fairly even distribution of the digested fragments, producing a good range of fragments for probe labelling Lane 3 was acceptable as there was also good distribution of the digested fragments, although the intensity of the original product indicate that time allowed for digestion should have been extended. Lane 4 was the opposite of lane 2, which showed over-digestion of the products, which was clearly visible with a concentration of fragments near the dye front.

If the digestion was satisfactory, the remainder of the product was cleaned and resuspended in hybridization buffer. 4-5µl was used to hybridize one slide.

3.2 FISH results

The HIV-gag specific probes, labelled with either FITC or spectrum-green, were applied to a number of variations of slides to try and produce accurate pictures of the viral location. The process was first tested on KC42 and U87 cell lines, before use on material obtain from HIV^+ patients. All the patients were tested in this manner.

3.2.1 Positive control: R5-using virus culture

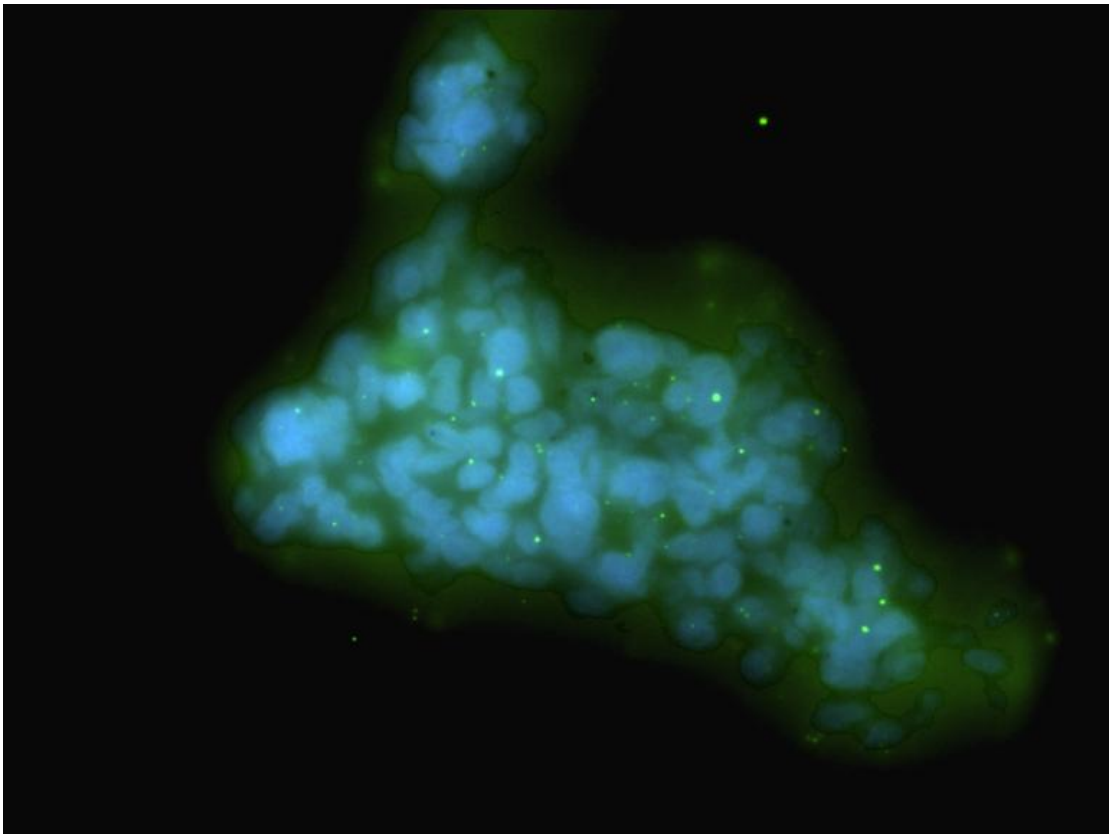


Figure 3.4: Hybridization of the FITC-labelled FISH probe to the KC42 cell line infected with STR5/Du179, a R5-using virus. The green dots were sites of successful hybridization. DAPI was used as counterstain to label the cell nuclei (100X mag.).

The occurrence of syncytium in the R5-virus infected cell line was a common feature, as the cells were free floating in the culture medium, and thus tend to congregate as infection progresses. Syncytia formation was especially common in the mid-late (from 5^{th} day) stages of the infection in these cultures. The presence of the fluorescent signals in and around the cell nuclei indicates the successful hybridization of the probes to HIV genome. The probe used was amplified from RNA extracted from the cells of the same culture, using PCR labelling.

3.2.2 Positive control: X4R5-using virus culture

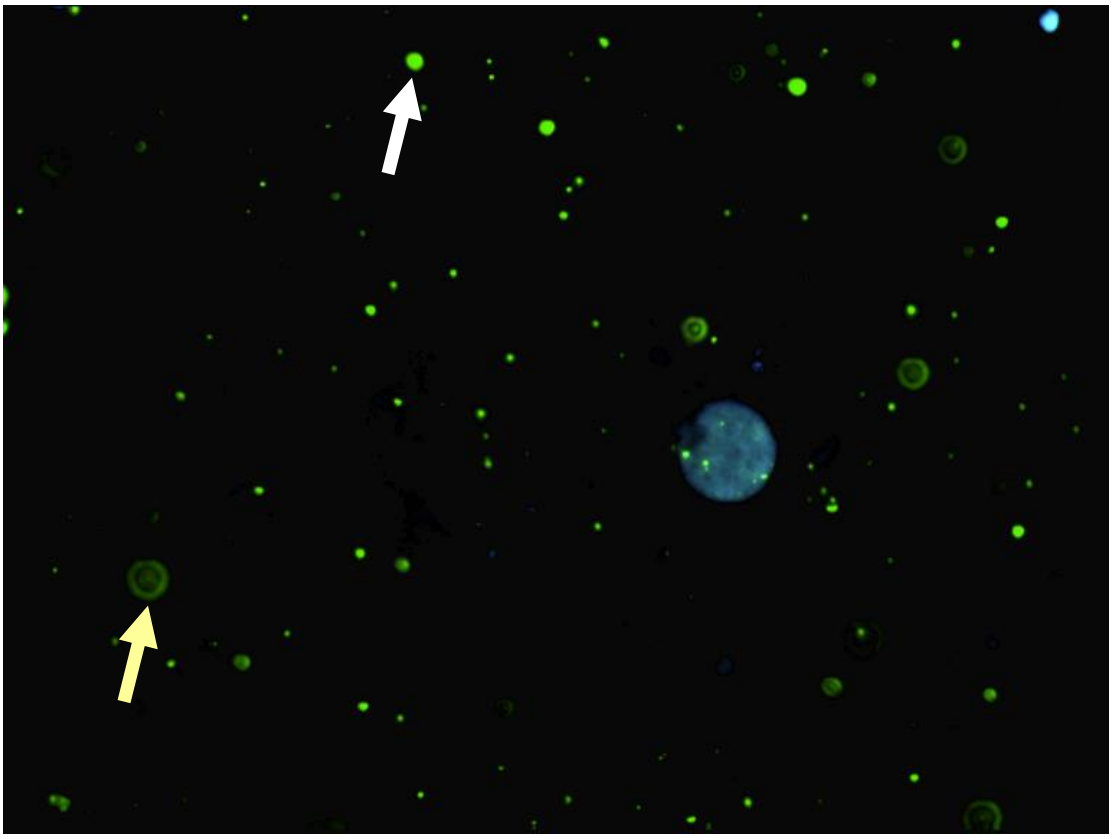


Figure 3.5: Hybridization of FITC-labelled FISH probes to the U87 cell line infected with CM9, a R5X4 (dualtropic) virus. The green dots were points where the probe has successfully hybridized to the HIV. DAPI was used as counterstain to label the cell nuclei. The arrows indicating signals that were at different planes of focus (100X mag.).

The R5X4 infected cell suffers cytopathic effects much faster than other of the R5 infected cultures. Dualtropic virus is highly virulent, even compared to the R5 virus that is virulent compared to the typical *in vivo* species. Consequently, it was somewhat difficult to locate any significant number of intact cells in these cultures. It can be seen that there was a large number of signals outside the cell nuclei, which was not surprising

as there were a high concentration of free-floating virions in these cultures. There were also a number of dots that was larger size, or possessed haloes around a point. This was due to the fact that the environment on the glass slide was a three-dimensional one; while the focus on one plane, it did not exclude the signals from other depth in the same field of focus. Therefore these larger dots of fluorescence were also positive signals, albeit somewhat out of focus and look bigger than what they suppose to look like. The multiple signals in the nucleus also seems to indicate that there are multiple integrants, or at the very least, multiple infection processes involving one cell. The probe used was manufactured using R5-using viral RNA, using PCR labelling.

The probe was shown to be specific to HIV genomic material in cell cultures, with both R5 and R5X4 cells. The same probe would be applied to the patient samples.

3.2.3 Positive control: HIV$^{\pm}$ patient whole blood

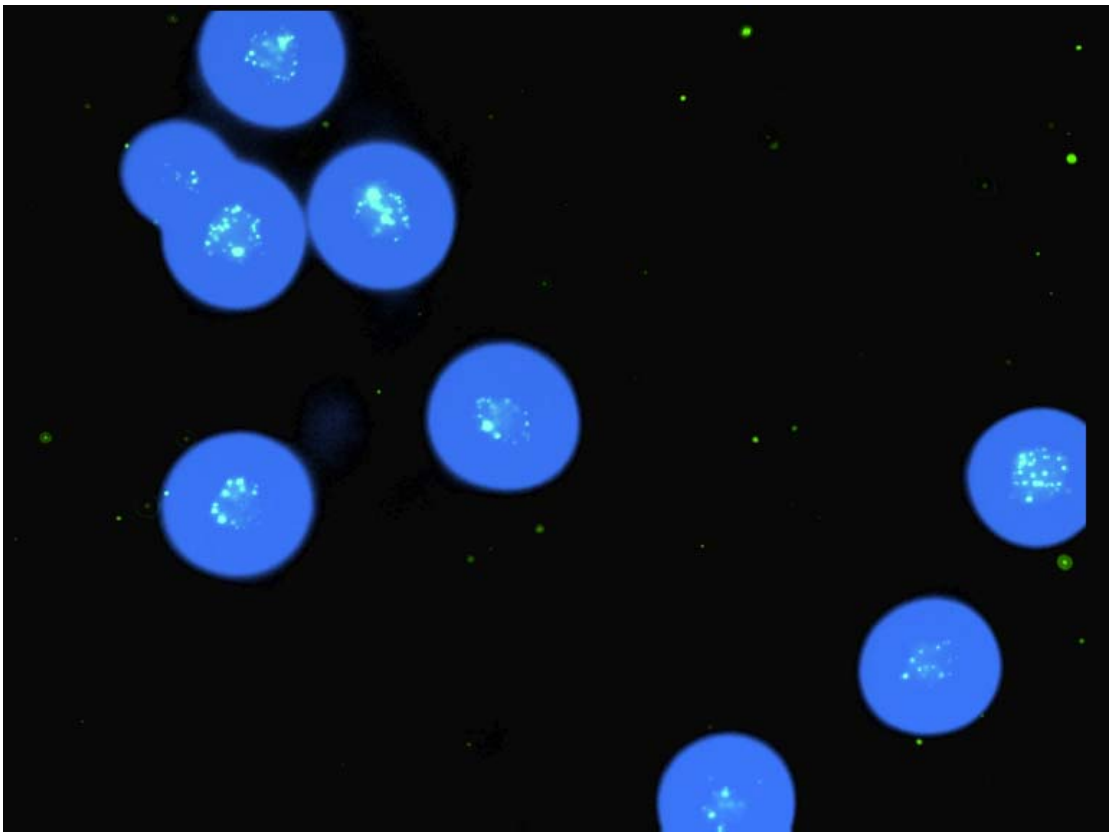


Figure 3.6: Hybridization of FITC-labelled FISH probe on a blood smear slide, made from fresh peripheral blood from HIV^{+} patient (100X mag.).

A large number of signals can be seen within the nuclei. The larger spots of fluorescence were due to the target being on the different plane of focus. Extracellular signals were also observed. This is a fairly clear picture of whole blood hybridization in comparison to earlier efforts, involving *in situ* hybridization (Embretson et al., 1993). Probe was manufactured from R5-using virus, with PCR labelling. There is a much smaller amount of extracellular signals compared to previous slide, which may be due to the difference in viraemia of an *in vitro* culture, versus an *in vivo* environment.

3.2.4 Negative control: HIV$^-$ whole blood

Figure 3.7: Hybridization of FITC-labelled FISH probe on a negative control blood smear slide, made from HIV- whole blood (100X mag.).

No presence of signals within the nuclei, and only one signal (more than likely an artefact) is observed outside. Probe manufactured from R5-using virus RNA, with PCR labelling.

3.2.5 HIV$^{\pm}$ PBMCs after ficoll-hypaque separation

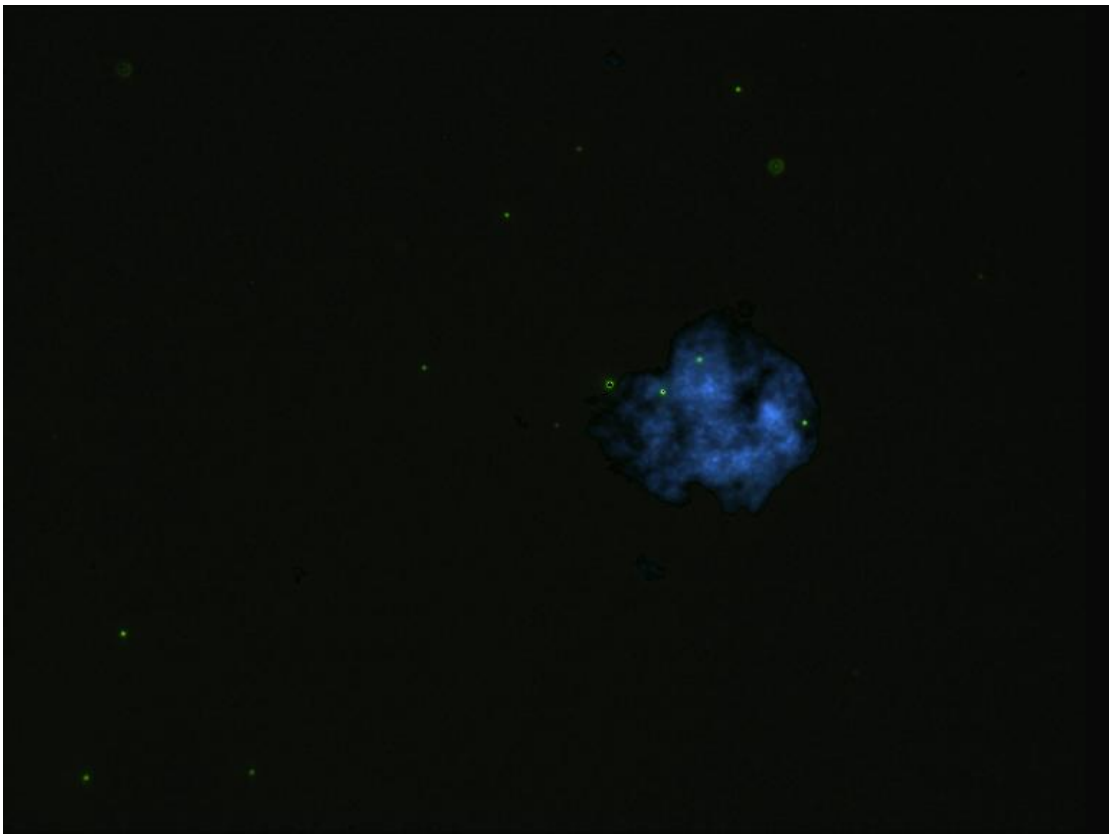


Figure 3.8: Hybridization of FISH probe on a ficoll-separated lymphocyte slide, from HIV^{+} patient (100X mag.).

Ficoll-hypaque separation often produced cells that have become diffused and have no coherent shape to the nucleus. Although signals can be detected, it was not possible to determine whether the viral material is in the vicinity of the nucleus through active transport, or the breakdown of the nuclear membrane during the separation process. Probe manufactured from R5-using virus RNA and PCR labelling.

3.2.6 HIV$^{\pm}$ bone marrow CD4$^{\pm}$ cells after Dynabead separation

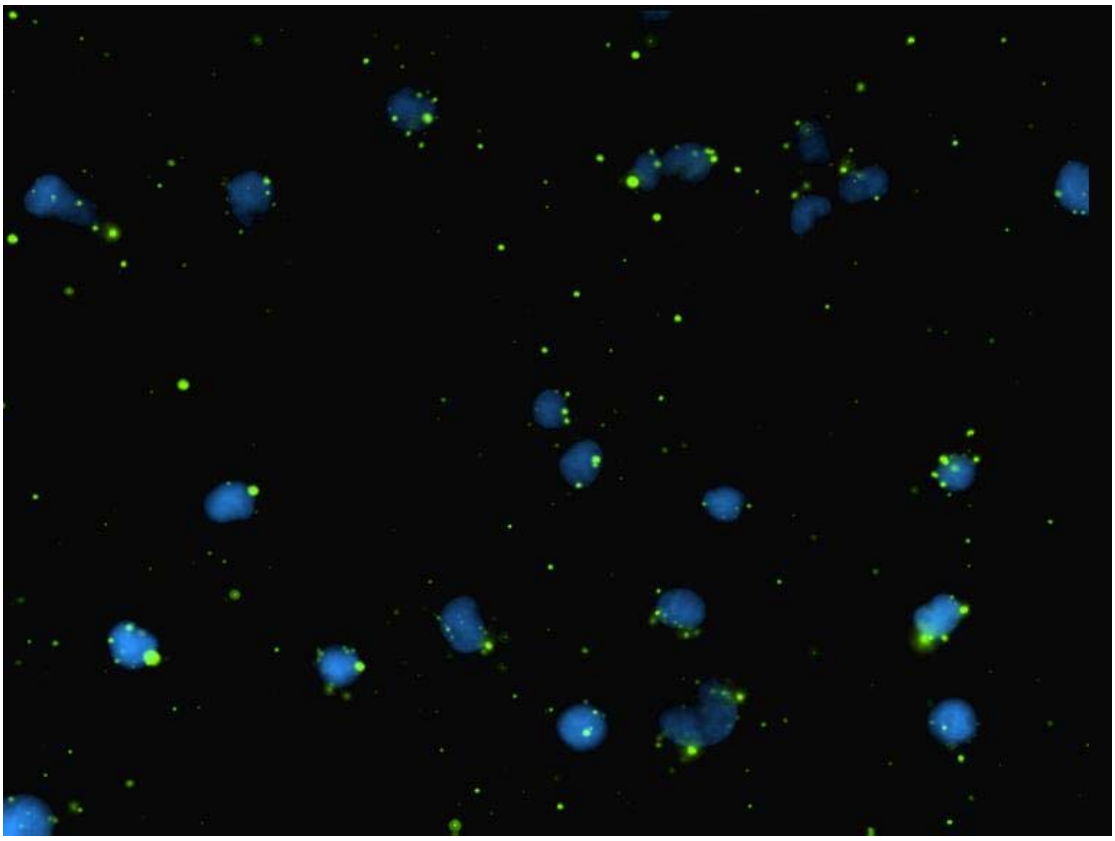


Figure 3.9: Hybridization of FISH probe, to bone marrow extract after Dynabead separation, from HIV^{+} patient (100X mag.).

These cells were CD4^{+}, as they were recovered using Dynabeads. The Dynabeads had been removed by DETACHaBEADs in this case. Signals were present in the nuclei as well as their periphery. There were also a number of signals in the inter-nuclear spaces on the slide. Probe manufactured from R5-using virus RNA and nick translation.

3.2.7 HIV$^{\pm}$ bone marrow CD4^{-} cells after Dynabead separation

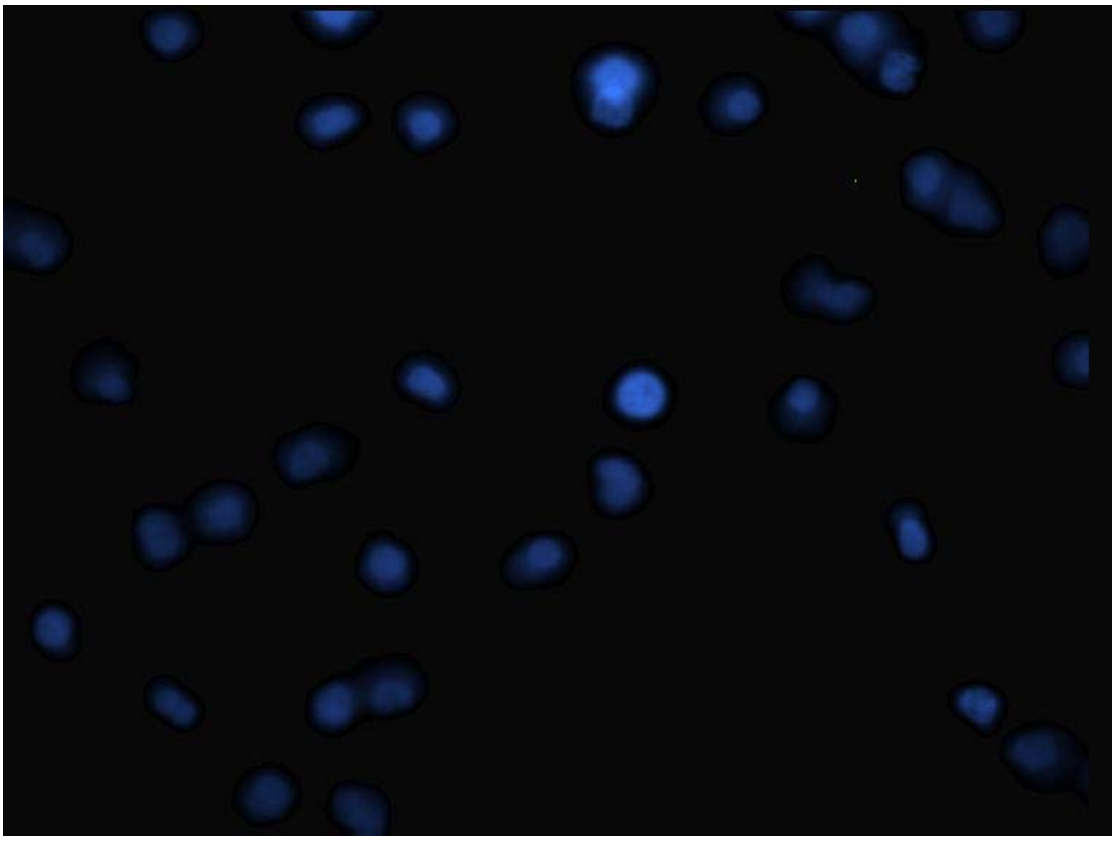


Figure 3.10: Hybridization of FISH probe, to CD4- cell from bone marrow after Dynabead separation, from HIV^{+} patient (100X mag.).

This was from the same sample of HIV^{+} bone marrow extract as the previous page, but with the removal of all the cells vulnerable to HIV infection and washing, it resembles a negative slide. It effectively demonstrates the high standard of purity and separation through the use of Dynabeads. Probe manufactured from R5-using virus RNA and nick translation.

3.2.8 Quantitative data from FISH analysis

Table 4.1: CD4 counts and amount of signals detected in the $CD4^+$ cell population of FISH slides from HIV^+ patients.

Sample no.	CD4 count	No. of T cells (total no. of signals)	No. of monocytes/macrophage (no. of signals)	Ficoll or Dynabead extraction
HIV02	138	50 (149)	0 (0)	Ficoll
HIV03	50	50 (110)	0 (0)	Ficoll
HIV15	193	50 (167)	0 (0)	Ficoll
HIV17	14	50 (216)	1 (9)	Ficoll
HIV19	14	50 (233)	0 (0)	Ficoll
HIV22	308	50 (102)	0 (0)	Ficoll
HIV24	109	50 (155)	0 (0)	Ficoll
HIV25	308	200 (411)	0 (0)	Dynabead
HIV26	109	200 (643)	0 (0)	Dynabead
HIV27	245	200 (396)	0 (0)	Dynabead
HIV28	200	200 (287)	0 (0)	Dynabead
HIV29	180	200 (752)	0 (0)	Dynabead
HIV30	230	200 (438)	0 (0)	Dynabead
HIV31	250	200 (405)	1 (12)	Dynabead
HIV32	146	200 (762)	3 (35)	Dynabead
HIV33	70	200 (795)	2 (19)	Dynabead
HIV34	300	200 (343)	0(0)	Dynabead
BMHIV1	200	200 (742)	6 (60)	Dynabead
BMHIV2	148	200 (620)	15 (102)	Dynabead
BMHIV3	173	200 (608)	11 (85)	Dynabead
BMHIV4	166	200 (729)	7 (78)	Dynabead

The table shows the amount of signals detected in the $CD4^+$ T cell and macrophage population. In the earlier samples, only 50 T cells were counted as the samples were separated using ficoll, and that method was ineffective in the amount of T cells it retained. It was difficult to find any significant amount of cells on the slides, consequently only 50 cells were counted. In the Dynabead separated samples, the yield was much higher, and thus a greater number of T cells could be counted.

With respect to the $CD4^+$ monocytes/macrophages, these numbers does not indicate the number of macrophages that was present on the slides. During the analysis of the slide, there was no intention to specifically seek out these cells to fulfil a quota, as the monocytes/MDMs were known to be of a low percentage. The monocytes/MDMs counted here were encountered as the counting for the T cell was proceeding, and thus were random. The differentiation between monocytes/macrophages and T cells were purely based on the expression of CD4, and the size of their nuclei. Expression of CD4 was confirmed by their recovery through the use of $CD4^+$ Dynabeads. T cell nuclei are smaller and more compact compared to the larger macrophage nuclei.

The number of signals on a slide from a patient who has low CD4 count was generally higher than those with a more encouraging CD4 count.

3.3 Sequencing results

A nested PCR targeting the *env*/V3 loop was used. The 2nd round product, around 600 bp, has to be visualized first on a 2% agarose gel before proceeding with the sequencing process itself.

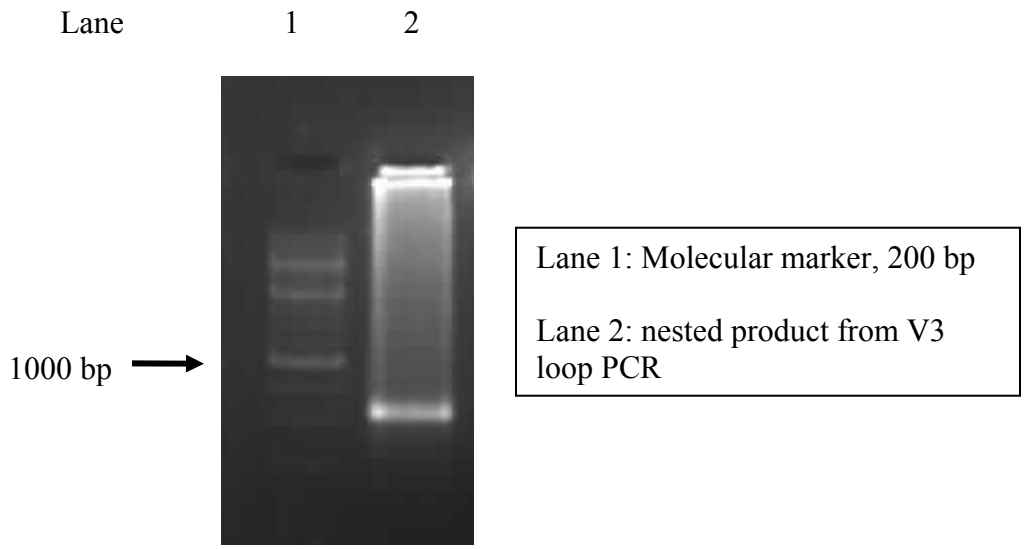
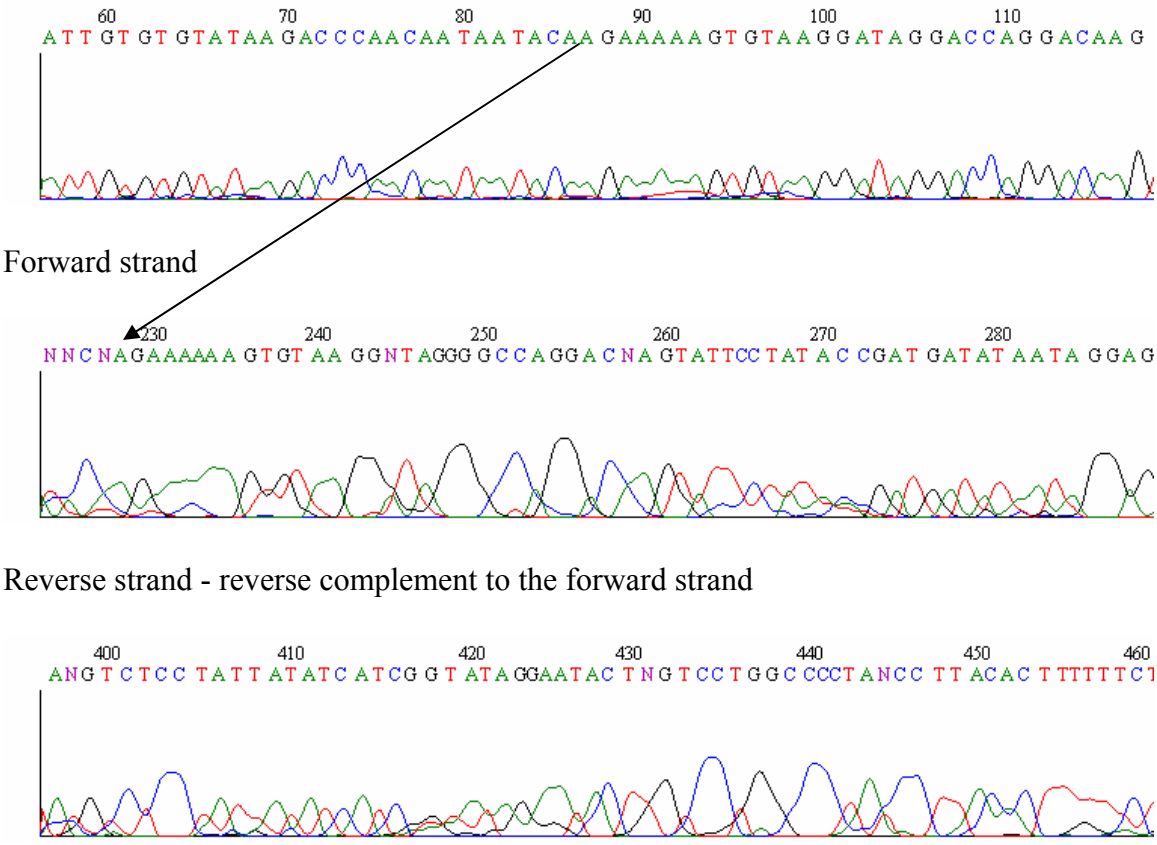


Figure 3.11: 2nd round product used for sequencing of the V3 loop.

Forward and reverse sequences of the HIV samples were obtained in the forms of ABI prism files and analyzed using the Sequencher software (Gene Codes Corporation, Michigan). Only samples with coherent forward and reverse sequence that can be combined, so as to produce the V3 loop sequence to resolve ambiguities, were examined.

3.3.1 V3 loop chromatograms

Results from the sequencing are produced as chromatograms, a series of peaks using four different colours to depict for the presence of a species of nucleotide (A, G, T or C) at any given position.



Forward strand

Reverse strand - reverse complement to the forward strand

Reverse strand (unmodified)

Figure 3.12: Portions of the chromatograms from sample HIV02. The arrow indicates the start of complementarity between the forward and reverse strand. The third chromatogram is how the reverse strand looked prior to reverse complementarity was applied.

The three chromatogram sections depict the beginning of the V3 loop from one of the HIV^{+} samples. The first chromatogram is the forward strand with part of the sequence prior to the V3 loop. The second chromatogram is the reverse strand. Using BioEdit© to orientate the reverse strand and to provide the complementary nucleotide, it provided a strand that matched the forward sequence. This is not to say there is a perfect match at every position: ambiguities were resolved using Sequencher. The third chromatogram was the reverse sequence prior to undergoing the reverse complementarity process.

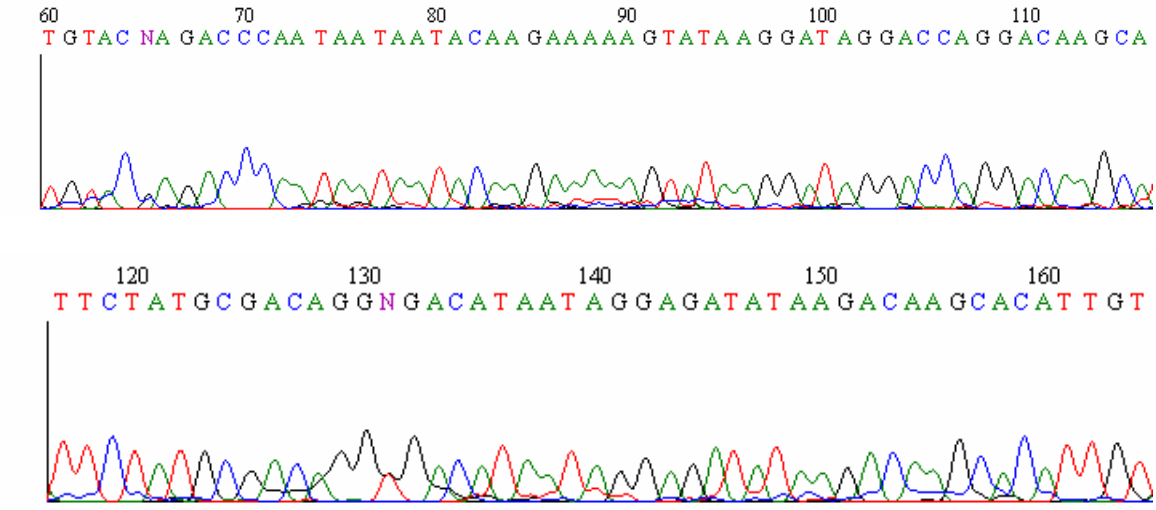


Figure 3.13: Complete V3 loop sequence, from BMHIV01.

Although a product around 600 bp in size was sequenced, the V3 loop was only a small section of the fragment. The V3 loop is of variable length due to deletions and insertions. The forward strand obtained from BMHIV01 was a good example of clean sequence, where there were no ambiguities in any of the nucleotide position in the V3 loop.

Table 4.2: The predicted amino acid sequence of patients in comparison to the consensus sequence of a typical HIV-1 subtype C V3 loop.

	1				5					10					15					20					25					30					35
Consensus sequence	C	T	R	P	N	N	N	T	R	K	S	I	R	I	G	P	G	Q	T	F	Y	A	T	G	D	I	I	G	D	I	R	Q	A	H	C
HIV02	-	I	-	-	-	-	-	-	-	-	-	V	-	-	-	-	-	-	V	-	-	.	-	D	-	-	-	-	-	-	-	-	-	-	-
HIV03	-	I	-	-	G	-	-	-	-	-	-	V	-	-	-	-	-	-	A	-	F	-	-	-	-	-	V	-	-	-	-	K	-	Y	-
HIV15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	V	-	-	-	P	-	-	-	-	-	-	-	-	-	-	-	-
HIV18	-	-	-	-	-	-	-	-	-	-	-	V	-	-	-	-	-	-	-	-	-	-	-	-	E	-	-	-	-	-	-	K	-	Y	-
HIV19	-	-	-	-	-	-	-	-	-	-	-	V	-	-	-	-	-	-	V	-	F	-	-	G	-	-	-	N	-	-	-	-	-	-	-
HIV21	-	V	-	-	-	-	-	-	-	-	-	V	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N	-	-	E	-	-	-	-
HIV22	-	-	-	-	H	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HIV24	-	-	-	-	H	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HIVBM01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HIVBM03	-	-	-	-	-	-	-	-	-	-	-	V	-	-	-	-	-	-	V	-	-	-	-	N	-	-	-	N	-	-	-	-	-	-	-
HIVBM04	-	V	-	-	K	-	-	-	-	R	-	M	-	-	-	-	-	-	V	-	-	.	N	-	-	-	-	-	-	-	-	-	-	-	-

18 samples were sampled, but only 11 produced coherent forward and reverse sequences that combined to provide unambiguous V3 loop sequences. The combined sequences generated the amino acid sequences listed in the table above.

The typical V3 sequence of a subtype C virus contains 35 amino acids, beginning with a CTRP motif, a crown GPGQ motif, and ending with a RQAHC motif. The motifs at the two ends of the V3 loop are typically quite well conserved, but there are still variations within them. The tyrosine to valine/isoleucine substitution at the second amino acid; glutamine to lysine/glutamic acid substitution at position 32 and the tyrosine to histidine subtition at position 35 are mutation that has been observed previously, and appears to have no effect on V3 loop functionality (Cilliers et al., 2003).

The GPGQ motif (highlighted in blue) is highly conserved in HIV, and each subtype posses different crown motif (Treurnicht et al., 2002; Cillier et al, 2003). There were no

changes at all in the region, which is consistent with existing data. The two amino acid positions highlighted in green are the two key positions – 11 and 25 – that have considerable influences on chemokine coreceptor usage. There were no changes at position 11, but mutation occurred at position 25 in 3 of the samples. The mutations detected at position 25 are not predicted to have major impact on the coreceptor usage, as they were not positively charged amino acids which can raise the overall charges of the loop.

Deletions (highlighted in purple) were also found in 2 of the sample. While X4-using viruses have a tendency to pick up insertions (particularly at position 13), R5-using viruses tend to lose amino acids instead. This does not detract from their ability from utilizing the chemokine receptors, but also seem to have little effects on the overall charges of the loop.

The N-linked glycosylation site starting from position 6 of the V3 loop (NNTR) was also examined, as glycosylation of this site is highly indicative of the virus using CCR5. This was observed in all the sequences obtained.

3.4 Phylogenetic analysis

The bootstrapped estimation of phylogeny produced by MEGA3 compared the sequences samples against a number of reference V3 sequences from other HIV-1 subtypes within

the M clade. The Chimpanzee SIV V3 loop was used as the outlier. The phylogenetic tree showed all samples were within the subtype C cluster, and no significant similarities except between HIV22 and 24. HIV22 and HIV24 were obtained from a husband-wife pair, which explains the high degree of homology between the V3 loops.

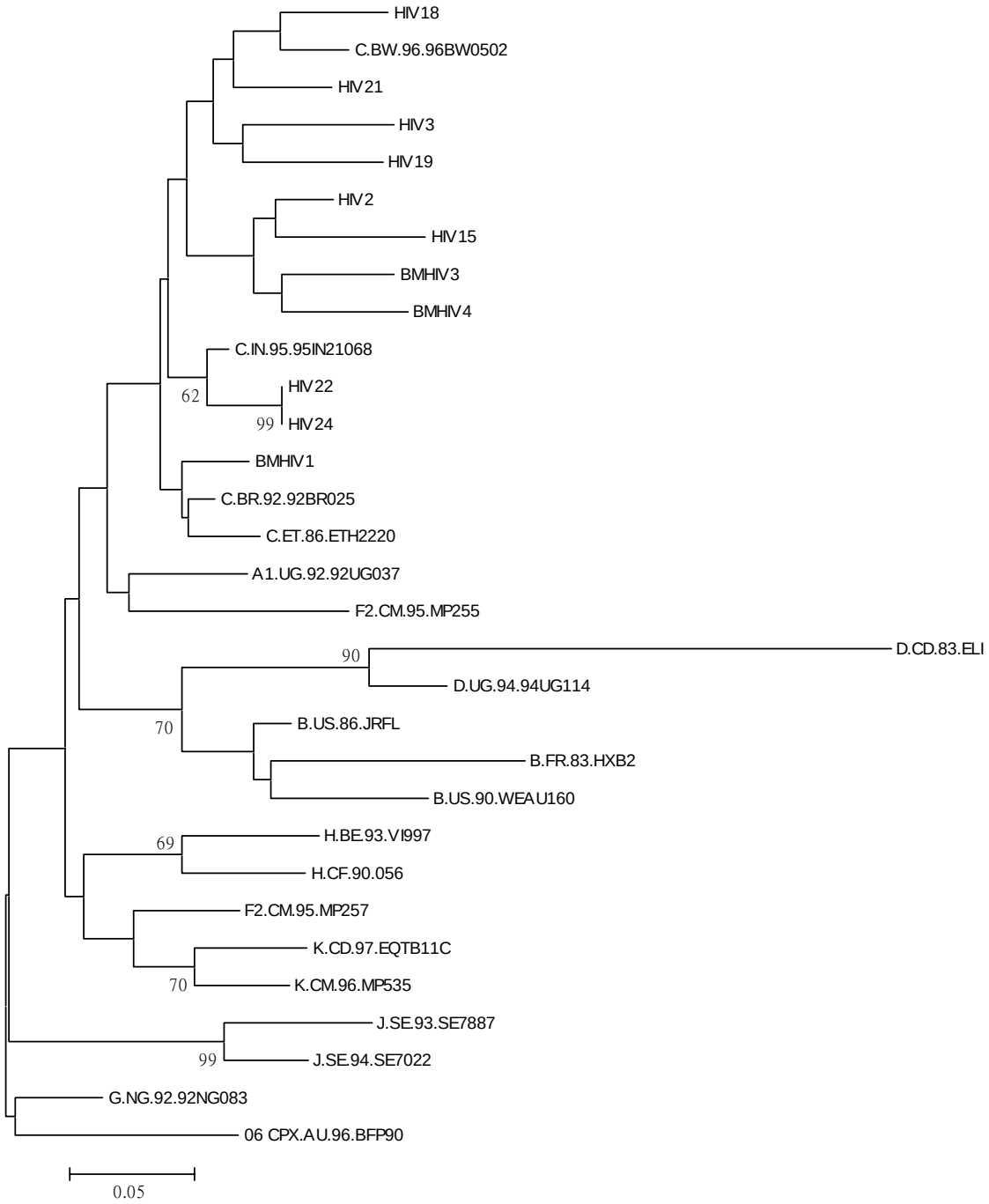


Figure 3.14: Phylogenetic tree (by neighbour joining method with maximum parsimony applied) and contamination analysis of the sequenced samples in relation to a number of subtype archetypes.

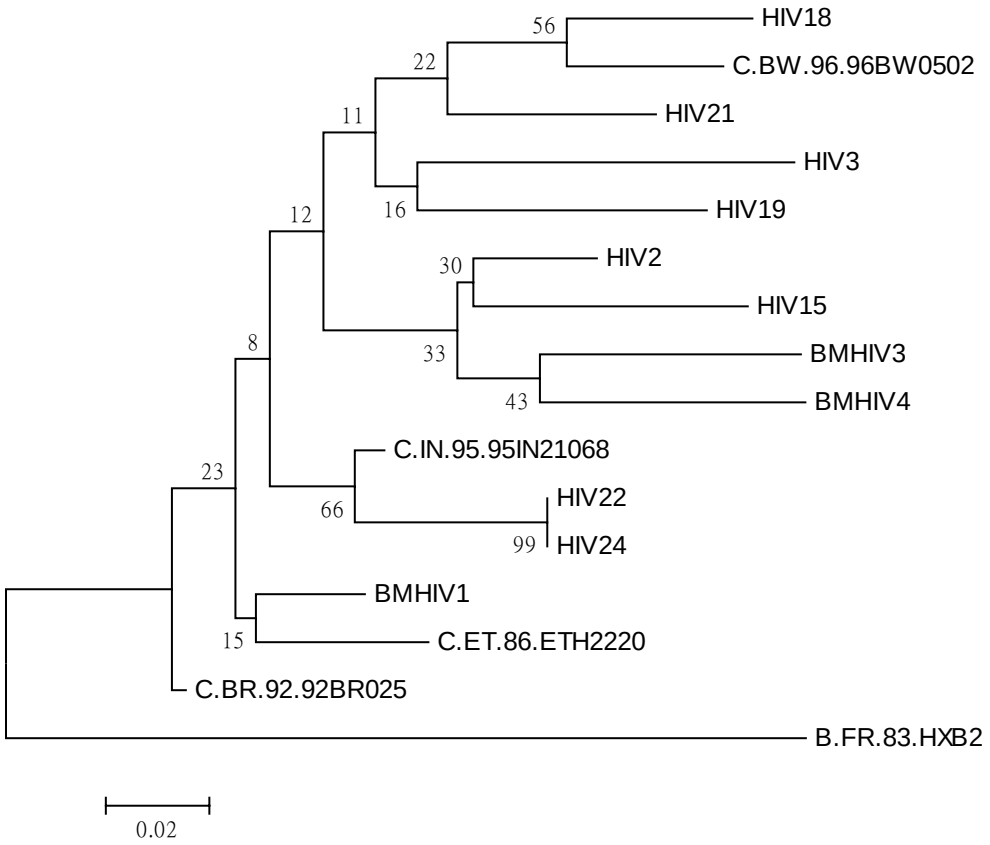


Figure 3.15: Phylogenetic tree (by neighbout joining method with maximum parsimony applied) of samples specifically within the subtype C cluster.

A tree using only the subtype C reference sequences to match against the sequenced samples. HXB2 of subtype B was used as the outlier. The bootstrap values were not sufficiently significant similar in any branches save the husband-wife pair. It is also not possible to determine if there are any similarity between the peripheral blood samples and bone marrow samples.

3.5 Prediction of chemokine coreceptor usage

Three methods can be used to predict the chemokine coreceptor used to virus isolates. The PSSM program not only provides the value for the matrix, but also looks at the amino acid at position 11/25, as well as calculates the overall charge in one complete table. It is therefore easier to look at all three together in the matrices below.

3.5.1 Position Specific Scoring Matrix (PSSM)

Table: 4.3: x4r5 PSSM matrix

name	score	pred	x4.pct	r5.pct	geno	pos.chg	net.chg	percentile
HIV22	-8.58	0	0.22	0.89	SD	7	5	0.55
HIV24	-8.58	0	0.22	0.89	SD	7	5	0.55
BMHIV1	-10.36	0	0.16	0.74	SD	6	4	0.33
HIV2	-5.18	0	0.33	0.96	SD	6	3	0.87
HIV15	-2.96	1	0.51	0.98	SG	6	4	0.63
BMHIV3	-4.15	0	0.42	0.98	SD	6	5	0.70
BMHIV4*	-3.23	0	0.51	0.98	SD	7	5	-10.00
BMHIV4*	-1.34	1	0.64	0.99	S-	7	5	-10.00
HIV3	-5.79	0	0.31	0.95	SD	6	4	0.88
HIV18	-5.91	0	0.29	0.95	SE	6	4	0.62
HIV19	-4.16	0	0.42	0.98	SG	6	6	0.69
HIV21	-4.53	0	0.36	0.98	SD	6	4	0.79

score: Sequence PSSM score
pred: 1: CXCR4 use predicted; 0: CCR5 use only predicted
x4.pct: Percentile of score within score distribution of known X4 viruses.
r5.pct: Percentile of score within score distribution of known R5 viruses.
geno: Amino acid residues at V3 sites 11 and 25; basic residues (R or K) at either or both of these sites is predictive of CXCR4 use.
pos.chg: Total number of positively charged (R/K/H) amino acid residues. Higher positive charge has been correlated with likelihood of CXCR4 use.
net.chg: Number of positively charged (R/K/H) amino acid residues, minus number of negatively charged (D/E) residues.
percentile: If this number is over .95, this sequence is most likely not a V3.

The column 'geno' indicate which amino acid is present at position 11 and 25 in these samples. None of them possessed amino acid of positive charges, which is indicative of R5-using viruses.

The column 'net charge' provided the overall positive charges for each V3 loop. All except HIV19 possessed +5 or less charges, which is a characteristic of R5-using viruses. Although HIV19 has +6 charges, it is not necessarily an indicator for X4-tropism. Taking into consideration of other prediction methods, and the typical value of 7 and over for CXCR4-using viruses, it is more likely for the virus to be R5-using.

With regards to the PSSM values, the scoring by the program indicated the majority of the samples were more likely R5-using viruses, save for HIV15, which was closer to being a X4-using virus. To test the validity of this, the sequences were run through the SINSI matrix.

There exists 2 BMHIV4 as there was a deletion present near position 25. As previously mentioned, position 25 is of importance as it is one of the key mutation for determining coreceptor switch. It was also required by the matrix for scoring, and thus the program took that into consideration and produced results for the variations that could occur.

Table 4.4: sinsi PSSM matrix

Name	score	pred	x4.pct	r5.pct	geno	pos.chg	net.chg	percentile
HIV22	-13.24	0	0.00	0.04	SD	7	5	0.55
HIV24	-13.24	0	0.00	0.04	SD	7	5	0.55
BMHIV1	-12.36	0	0.00	0.16	SD	6	4	0.33
HIV2	-3.88	0	0.10	0.99	SD	6	3	0.87
HIV15	-3.02	0	0.14	0.99	SG	6	4	0.63
BMHIV3	-6.25	0	0.06	0.92	SD	6	5	0.70
BMHIV4*	-5.72	0	0.07	0.94	SD	7	5	-10.00
BMHIV4*	-1.37	1	0.19	0.99	S-	7	5	-10.00
HIV3	-8.44	0	0.04	0.85	SD	6	4	0.88
HIV18	-7.73	0	0.04	0.88	SE	6	4	0.62
HIV19	-5.56	0	0.07	0.95	SG	6	6	0.69
HIV21	-11.62	0	0.01	0.30	SD	6	4	0.79

score: Sequence PSSM score
pred: 1: CXCR4 use predicted; 0: CCR5 use only predicted
x4.pct: Percentile of score within score distribution of known X4 viruses.
r5.pct: Percentile of score within score distribution of known R5 viruses.
geno: Amino acid residues at V3 sites 11 and 25; basic residues (R or K) at either or both of these sites is predictive of CXCR4 use.
pos.chg: Total number of positively charged (R/K/H) amino acid residues. Higher positive charge has been correlated with likelihood of CXCR4 use.
net.chg: Number of positively charged (R/K/H) amino acid residues, minus number of negatively charged (D/E) residues.
percentile: If this number is over .95, this sequence is most likely not a V3.

The SINSI matrix uses similar principles as the R5X4 matrix to calculate whether the virus is a synctium inducing virus, as opposed to a CXCR4-using virus, and so is an alternate form of comparison to the R5X4 matrix results.

The SINSI matrix agreed with the R5X4 matrix, save for HIV15, which it now considered as an NSI virus. This could indicate the virus was possibly at an intermediate evolutionary stage, where the virus was becoming dualtropic, thus leading to the mixed results from the two matrices.

Table 4.5: Manual calculation of V3 loop charge

	1 5 10 15 20 25 30 35	Charge
Consensus sequence	C T R P N N N T R K S I R I G P G Q T F Y A T G D I I G D I R Q A H C	
HIV02	- I - - - - - - - - - V - - - - - - V - - . - D - - - - - - - - - - -	+2.5
HIV03	- I - - G - - - - - - V - - - - - - A - F - - - - - V - - - - K - Y -	+4.5
HIV15	- - - - - - - - - - - - - - - - - - V - - - P - - - - - - - - - - - -	+3.5
HIV18	- - - - - - - - - - - V - - - - - - - - - - - - E - - - - - - K - Y -	+3.5
HIV19	- - - - - - - - - - - V - - - - - - V - F - - - G - - - N - - - - - -	+3.5
HIV21	- V - - - - - - - - - V - - - - - - - - - - - - - - - N - - E - - -	+2.5
HIV22	- - - - - H -	+4.0
HIV24	- - - - - H -	+4.0
HIVBM01	- - - - - - - - - - - - - - - - - - A - - - - - - - - - - - - - - -	+3.5
HIVBM03	- - - - - - - - - - - V - - - - - - V - - - - N - - - - N - - - - -	+3.5
HIVBM04	- V - - K - - - - R - M - - - - - - V - - - . N - - - - - - - - - -	+5.5

The one significant caveat in the PSSM matrix result is the rounding off of the charges the V3 loop carries. The software considers histidine (H) residues as carrying one full positive charge, as oppose to 0.5 positive charge. While this makes little difference to most of the sequence, for V3 loops that has 5, or 5.5 positive charges, it is significant,

because a V3 loop that carries 5-6 positive charges, may belong to either CCR5- or CXCR4-using viruses. HIV19 and HIVBM04 are two samples that falls into this ambiguous range, both carrying 5.5 positive charges. In such a situation, overall positive charge for the V3 loop becomes irrelevant, and would require other data in order to make a reliable prediction.

Chapter 4: Discussion

There remain many challenges to prevent and treating human immunodeficiency virus infections, despite over two decades of research and investigation (Fauci et al., 2005; Fauci, 2005). The long period of disease progression has made it difficult to assess the virus' characteristics at times of infection and transmission – long term studies (5-10 years studies) are required to obtain results correlating time of infection and disease progression. There is also still no satisfactory animal model that can fully mimic the infected human system. To resolve these problems, researchers have been working on producing data regarding the virus over short periods in cell cultures. This can provide a fairly informative representation of the life and interactions of HIV in real infections. This study investigates the location of the virus in late stage infection without the interferences of ART, or opportunistic infections like tuberculosis. More specifically, in this undisturbed setting, we are trying distinguish between the two $CD4^+$ cell populations of macrophages and T cells, to determine where the primary sites of infection, which may reflect sites of active replication.

4.1 FISH, optimization and observations

The use of FISH in HIV research has been scarce, particularly as a standalone technique. In spite of the advantage of localizing the target without severely disrupting the cellular

environment, it suffers from a low sensitivity threshold. Probe sensitivity is usually assured by the coverage of a large expanse of the target and flanking areas. Alternatively, a target of small genomic length can be compensated for if it has a high copy number, which would ensure successful detection. Unfortunately, with the administration of ART, neither condition exists. HIV genome is a mere 10 kb fragment, too small to be efficiently hybridized. The advantage in high copy number is at the same time eliminated by the antiretrovirals. At best, the used of FISH as a standalone has been confined to cell line experiments (Cone & Schlaefer, 1997; Deichmann, Bentz & Haase., 1997; Knuchel et al., 2000), or used in combination with PCR or flow cytometry to compensate (Borzi et al., 1996; Embretson et al., 1993; Patteron et al., 1993).

Although ART has been introduced to South Africa as part of the HIV/AIDS treatment program, there remain a large number of infected individuals that need to be examined and placed onto therapy. In this situation, it has been possible to obtain whole blood or bone marrow from ART-naïve patients, who were in the late stages of the infection, admitted prior to antiretroviral treatment. High copy numbers of the virus thus still exist in these samples, and FISH can be used for detection in this instance. It would have been useful to use a protocol from one of the studies mentioned above, but these experimentations were not well suited for this study. They were typically performed in a highly controlled environment, using cell lines and *in vitro* experimentations, and this was not the setting that was required here. Consequently, it was necessary to devise a novel FISH variation for our use.

4.1.1 Optimization of HIV-specific FISH

The probe was designed to target the *gag* gene of the virus. It is a region of viral genome that is rarely mutated in all contagious variants compared to other regions (Desai et al., 1986; Engelbrecht et al., 2001). This would ensure the detection of the virus, no matter which subtype or quasi-species was present. Although the template for the probe was the viral RNA, the probe itself was DNA. A DNA-based probe is easier to manufacture, handle and store, as it is more stable and not as easily degraded.

The probe and the hybridization process were first tested on cell cultures, infected either with R5- or dualtropic HIV-1 subtype C viruses. The probe was able to hybridize to the virus in both cultures, and was detectable under fluorescent microscopy. The negative controls using the cultures without the virus showed no hybridization within or nearby the nucleus. A few signals were present outside the nuclei, typically at least one field away from any nucleus, and only present in small number (less than 10 signals per 200 cells counted). They were therefore unlikely to be associated with the cell cytoplasm. The rare signals in the control cultures were likely to be artifacts, and were a good indication of the fact that the probe is highly specific to the HIV genome. The low number of signals in the negative control is expected as the target is much smaller than the typical size, and would invariably lead to background signals of similar sizes and intensity (Raap, 1998).

The protocol was then transferred to material from the peripheral blood. Hybridization to whole blood smear slides and subsequent detection was successful, although there was a

high degree of background observed on these slides. This was possibly due to the presence of other cells and components of the blood. The probes were getting trapped by these elements, and created false hybridization signals. The free floating virions were also an obstacle. Due to the limitation of the microscope when looking at a three dimensional environment, these signals may appear to be associating with a cell, which would distort the results obtained. The negative control slides made from whole blood of HIV^- individuals, however, had minimal signals again. It was thus more probable that the cell free virions were the leading cause of the large number of signals outside cell nuclei in whole blood.

It was therefore necessary to lessen the impact from the cell free virions and components of the blood, as well as to concentrate the white blood cells. The first attempt was by centrifugation, to separate the blood into fractions by mass weight. The blood would separate into serum, white blood cells, and red blood cells, and the leukocytes were extracted by pipetting. The effects of the process on the two parameters were minimal. The blood separated adequately, but manual pipetting still led to a fair degree of contamination from other blood components, and an incomplete recovery of the lymphocytes. Ficoll-hypaque centrifugation was then attempted. The combination of ficoll and centrifugal force would create a clearer differentiation between the various blood fractions. This concentrated the leukocytes more than the pure centrifugation of the whole blood. But as it still relied on manual pipetting, the method remains affected by incomplete recovery of leukocytes. This was also impeded by the very low concentration of lymphocytes in the blood of the HIV^+ individuals admitted into the study. In contrast

to healthy HIV^- controls in the same procedure, HIV^+ samples rarely had sufficient lymphocytes to fully cover the surface area of the tube used. It was therefore extremely difficult to receover all the lymphocytes without drawing the neighbouring fraction as well. If contamination was the issue, only a small amount of lymphocytes can be drawn from the tube. Small number of cells makes the screening of the slide an uncomfortable experience, with the cells few and far in-between. There was also the problem of broken cells in the HIV^+ samples. Few lymphocytes seemed to be able to maintain their integrity following the ficoll-hypaque separation, and even the negative controls suffered to some extent. The reason for this remains unknown, but may be due to the centrifugation process.

Separation through centrifugation and the inherent mass difference of the blood components was thus insufficient. The recovered lymphocyte would also still be a combination of $CD4^+$ and $CD4^-$ cells, and it would be necessary to further isolate the former cell population, in order to observe the interactions between those cells and the virus. To resolve both these problems, Dynabeads (Invitrogen) were used, which were polymer beads that were uniform in size, and superparamagnetic in nature. They can be labelled with ligands specific to the cell population that one would like to recover, or remove, from the biological environment. The Dynabeads used was targeting the $CD4^+$ cells only, and the efficiency of the process was highly impressive. The blood that remained after the $CD4^+$ cells removal, were used to make slides and subjected to FISH. There were no signals associating with the nuclei of the cells present on those slides. When FISH was performed on the $CD4^+$ cell population, the signals were abundant in

both the nuclei and the intracellular spaces. The separation was also more effective as there were a greater concentration of intact and infected cells. There was also a small increase in the number of monocytes observed. The process was as effective on the bone marrow extracts. Under fluorescent microscopy, the only possible way to discern the $CD4^+$ monocytes from T cells was through the different size of the nucleus, as no other cellular structures were visible. T cells have a small concentrated nucleus (a more intense blue circle), while the macrophages possess a larger, more diffused nuclei. The macrophage nucleus was typically twice the size of the T cell nucleus, and also much lighter in the DAPI staining – they were faint enough for the computer software to dismiss as background. While it was possible to count the number of signals, it was not possible to capture the DAPI staining for the macrophages. This is the main reason for the lack of infected macrophage images, as the computer analysis software was not able to capture the cell nuclei properly.

There is an issue of visually defining the location of the signals when examining the labelled slides. There is no clear cut way to differentiate a true viral integrants against a virus that is sharing the same locality as a cell nucleus, except by moving through the complete plane of the slide, and determine the plane of focus (p. o. f.) for the signals and the cell nucleus. There is typically one clearly defined plane of focus for the cell muclei, and all signals detected within the DAPI-stained areas can be considered to be inside the nuclei. The signals would be highly concentrated dots of green if they're in focus. As the plane of focus is moved, there may be other signals that come into focus while the

nucleus loses its visual clarity. These would be more difficult to interpret, and would rely on the expertise – and to some degree – the subjectivity of the person reviewing the slide. Therefore, becase of the different number of signals that can exists at different depth, dependent on the field of focus, it is not possible to objectively state the number of signals per cell and justify it through single frames of image capture. It may be possible to superimpose multiple images of one cell to obtain a full picture, but the resulting image would be a highly obscure one, due to effect of fluorescent signals in and out of focus. The number of signals obtained for cells during the process of counting, is an accumulation thereof, and still images cannot be expected to be fully representative of the environment under the microscope.

In this study, the plane of focus was not significantly shifted when quantifying the number of signals relating to the cell nucleus. When the shift in plane of focus has caused the nucleus to become indistinct on its borders, this is considered the edge of the nucleus, and signals that are not in focus by this point are disregarded.

The fact that the hybridization environment is a three dimensional one, which can only be represented as a two-dimensional one on paper, also makes it slightly difficult to make satisfactory comparisons to other FISH-HIV papers, even if one is to disregard the different in probe manufacture, hybridization methods, viral subtypes, conditions used to culturing the infection, and so forth. But there is without a doubt that detection of the viral signals is greater abundance than any previously published data (Lawrence et al., 1990; Embretson et al., 1996). Whether this is a true reflection of repeated

infection/supre-infection, or the way HIV-1 subtype C infection behaves in late infection, may be argued, as there are no other data thus far to support, or dissuade the observations made in this study.

4.1.2 Quantitative analysis of $CD4^{\pm}$ monocytes/macrophages and T cells

The HIV^{+} slides were then analyzed quanlitatively. The $CD4^{+}$ T cells were the dominant population detected in peripheral blood and bone marrow – less so in the latter environment - with the monocytes/macrophages being very few in numbers (less than 1%). The T cells and monocytes/macrophages can be visually differentiated with the size of their nuclei – the former has much smaller, brightly DAPI stained nuclei, whereas the latter is three to four times the size of T cell nulei, and is typically faintly stained with DAPI. Almost all (±99%) of the viable and structurally intact T cells detected harboured at least 2 positive signals, and averaged 5-10 signals in each nucleus, with an upper limit of 40 signals detected. With regard to the macrophages/monocytes, even when concentrated through the use of Dynabeads, they were still a very small population in comparison to the T cells. In the peripheral blood, they were almost non-existent, and those that were detected would only be the monocytes that were in the process of maturation. Only in the bone marrow extract can the monocyte-derived-macrophages be considered fully mature. In both environments, the macrophages were approximately 1% of the total number of T cells present on a slide, if not less. The peripheral blood and bone marrow had thus presented a very small, infectable population of cells outside $CD4^{+}$ T cells.

The blood and bone marrow are therefore unlikely areas where the M-tropic viruses are replicating. The lymph nodes would be more probable locations where there is a substantial number of MDMs for HIV to infect and replication under this observation. However, both macrophage and T cells possessed signals within their nuclei, which demonstrated their status as infected cells, and possibly as reservoirs. The positive infection of the cells cannot be disputed as the signals were observed in the nuclei. Although there are a number of blocks to active replication of the virus, like the absence of PICs and low nucleotide concentration (Bukrinsky et al., 1992; Spina et al., 1995; Han et al., 2004), there has been no evidence of blocks in the integration process, once the viral genetic material has been transported into the nuclei. If the viral cDNA can enter the nuclei of the infected cells, then it can be integrated. Considering the existence of points of blockade, which prevent the successful production of HIV virions, a pure visualization of the presence and location of HIV genetic material in the nucleus is therefore insufficient proof of replication.

The macrophages detected, all possessed no less than 10 signals within their nuclei, but they never approached the upper limit numbers detected in the T cells. The highest number observed in macrophages was 29. It has been suggested that macrophages do not possess a block in the export of the preintegration complex that allows the transport of virus genome material into the nuclei as opposed to the resting T cells. They also possess greater kinetics which allows greater efficiency in import of materials like nucleotides that would be necessary in transcription. This should theoretically mean that there are more copies of the viral genome in the nuclei, as there were no impediments to viral entry,

but that was not the case. It may be possible that there is less need for the viral genetic material to repeatedly infiltrate the macrophage nuclei to ensure successful integration and replication, due to higher kinetic of the cells.

In contrast, the T cells have multiple blocks to the import of the viral genome into the cell nuclei, as mentioned above (Bukrinsky et al., 1992; Spina et al., 1995; Han et al., 2004). When the blocks in the T cells are removed, the viable material that hasn't degraded would be able make their way into the nucleus, and causing the large number of signals in the compartment. Alternatively, signals inside and near the nuclei may be describing the presence of new viral material. They have accumulated within the nuclei or cytoplasm due to the slow kinetic nature in the resting T cells and a lack of export signals.

The examination of $CD4^-$ cell populations left behind by the Dynabead isolation, followed by FISH, also confirmed the fact that HIV-1 was not able to infiltrate these cells. Haemopoeitic progenitor cells in the bone marrow would fall into this group, where around 90% are $CD34^+CD38^+$ cells, which have a minimal expression of CCR5 on their surfaces. The remaining 10% are $CD34^+CD38^-$ cells that have a low level of CCR5 expression. Other $CD4^-$ cells in the bone marrow typically express CXCR4 (Berkowitz et al., 1998a). So even with viable chemokine coreceptors present on these cells, the virus actually cannot substitute CD4 with other surface receptors to gain access into cells; or at least, not in the peripheral blood and bone marrow tissues.

4.2 Phylogenetic analysis of samples and prediction of chemokine coreceptor usage

The incongruence between the presence of the virus in the CXCR4-expressing T cells and the supposed rarity of X4-using viruses in the South African HIV^{+} population, predominantly infected by HIV-1 subtype C, have posed an interesting dilemma in this study (Tscherning et al., 1998; Peeters et al., 1999; Morris, et al., 2001). It was therefore necessary to determine the chemokine coreceptor usage by the viruses in the sample obtained. While the relationship that exists between the switch of chemokine coreceptor usage from CCR5 to CXCR4, and the viral phenotype change from NSI to SI, is not an absolute one. The presence of the X4-using virus is a good indicator of the virus mutating to a SI phenotype (Dong et al., 2005). The SI virus is also typically much more aggressive, and leads to rapid disease progression as well as decreased survival time. In monitor seropositive cohorts, as well as for evaluating antiretroviral therapy trials, the appearance and presence of X4-using virus has been used as a marker for these two conditions (Richman et al., 1991; Koot et al., 1993).

4.2.1 Phylogenetic analysis

18 samples were submitted to sequencing using the ABI prism 2100 sequencer, but only 11 samples produced coherent V3 loop sequences for analysis. Phylogenetic and molecular evolutiondary analyses of these samples were conducted using MEGA version 3.1 (Kumar, Tamura, Nei 2004).

The software provides a number of options for the creation of phylogenetic trees, although in this analysis, only the neighbour-joining method was used to produce the tree. This method devised by Saitou and Nei (1987) construct phylogenetic trees by defining pairs of nearest nodes (which leads to minimal number of branches in the tree) as neighbours, and repeats the process until all the samples has been paired off in such a manner. The lengths of the tree branches were then calculated and provide a value as to how similar the neighbours were. It also inferred the amount of evolution and similarities that occurs between the samples. The advantage of the neighbour-joining method was that it does not assume a constant rate in the samples' evolution, which makes it a more realistic model. Compared to other models of analysis that used the principle of minimum evolution, the neighbour-joining method is also less time consuming, uses less computational procedures and remains accurate in its predictions (Kumar, Tamura and Nei, 2004).

Here, MEGA 3.1 was not only used to observe the phylogenetic relationships, but also to check if there was any cross contamination of samples that may have occurred during the sequencing process. There were no patterns of association between the samples themselves that suggested a possibility of contamination. There were two samples that were genetically identical to each other (samples HIV22 and HIV24), but this can be explained by the fact that they were a husband-wife pair, one of whom had transmitted the virus to the other. With respect to the phylogeny of the samples and the reference sequences, all the samples were most closely related to the subtype C reference samples. These were all clustered within the same branch in which the subtype C reference

sequence was placed. This short fragment was insufficient to produce evolutionary information that pertains to the virus outside of the loop, due to the lack of additional genetic information.

Bootstrapping was also applied to the analysis, which attempts to determine probability distribution in the data group by randomly sampling the group. Bootstrap is a resampling process that quantifies uncertainties by calculating standard errors, confidence intervals, and significant tests (Soltis & Soltis, 2003). It has only become viable through the advent of computers, and in comparison to traditional resampling methods, it required fewer assumptions, and may provide more accurate answers. There is disagreements concerning at what point the bootstrap value is considered significant, but at least 95% bootstrap are usually considered to be significantly between neigbouring samples. There are certain situations, though, where 70% bootstrap proportion is deemed to correspond to a 'real' clade. In the sequenced samples, only the husband-wife pair presented significant correlation, which was expected in this case. Although a straightforward phylogenetic tree looks almost identical to its bootstrapped counterpart, save for the fact that latter possess values at each branch (see figure 4.1). Bootstrap is usually set between 100-1000 resampling, and there is no consensus as to how many times resampling is optimal for a sample population. The sequences were resampled 250 times in this instance, which is theoretically sufficient for such a small sample size.

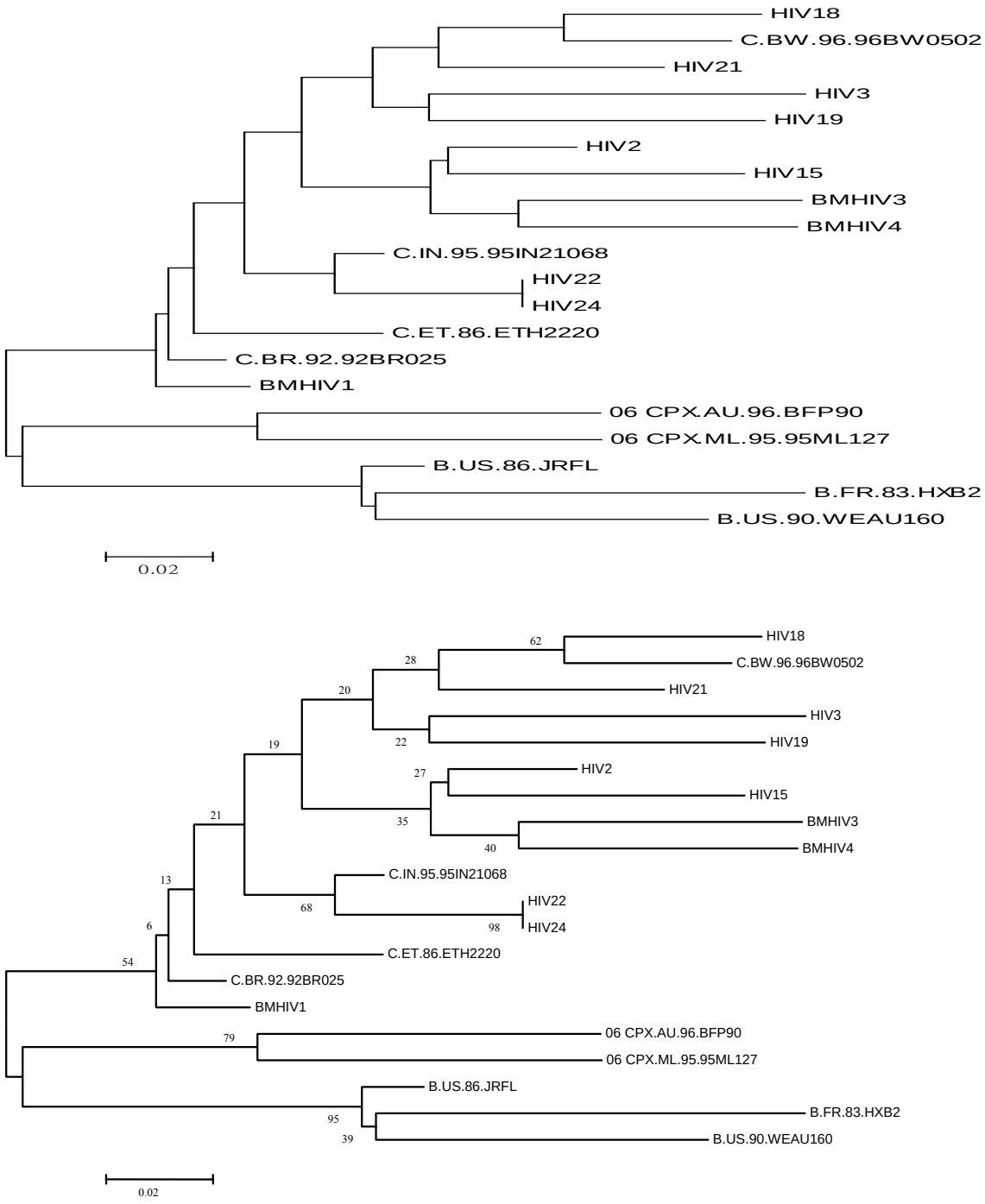


Figure 4.1: patient V3 loop sequence with subtype B, C and chimpanzee reference sequences. The first tree (by neighbour joining) is without bootstrapping, while the

second tree (also by neighbour joining) is bootstrapped 250 times. The values at each branch are the estimated genetic similarities between the neighbours.

Due to the short length of the V3 loop sequence, it is again difficult to obtain significant data from bootstrapping. This could be remedied by sequencing more of the flanking region of the V3 loop; alternatively, distance matrix may be utilized to determine the genetic similarities between the samples. Nonetheless, there are other ways to determine subtypes, as will be discussed. The only disadvantage is that one cannot estimate the genetic similiarities between samples.

4.2.2 Analysis of the V3 loop amino acid sequence

The samples' V3 amino acid sequence can also provide some validation to their subtypes, purely by visual comparison of the sequences obtained. There are three motifs that exist within the V3 loop that showed noticeable differences between the various subtypes, and they are conveniently situated at the beginning, the middle, and the end of the sequence. The first motif is a four amino acid sequence that denotes the start of the V3 loop, typically 'CTRP'. The variations often occur at the second position, including valine (V), and isoleucine (I), both of which appeared in the samples, as well as alanine (A), cysteine (C), and tyrosine (Y). There were no variance with the first, third or fourth amino acid positions, and was also consistent in the reference sequences. The second motif is at position 15-18, and is the crown of the V3 loop. This region is less mutable than other regions of the loop and highly consistent. With the exception of subtype B, which almost always possesses a 'GPGR' motif, most of the other subtypes have 'GPGQ', with a few

'GPGR' variant and even more rarely, 'GPGK'. All the samples had the 'GPGQ' sequence in this motif, making it more likely that they are subtype C than B, which is the second most frequent subtype in South Africa. The third motif at the end of the V3 loop, is typically a 'RQAHC' amino acid sequence in all the subtypes. There are many more variations on the end of the V3 loop compared to the other two motifs. Two variations that appeared in the sequenced samples were 'RKAYC' and 'REAHC'; the former appeared in the references in more obscure subtypes like F and J, while the latter does not occur at all, although the individual substitutions do occur in the reference subtypes.

The other region of interest in the V3 loop was from position 5-10, which typically carries the 'NNNTRK' sequence. The amino acid in position 6, which is also in this motif (also known as g15, as it is the 15th N-linked glycosylation position in the gp120) has been acknowledged as being a site of N-linked glycosylation, and this modification has been found to have different preferences for the chemokine coreceptors (Pollakis et al., 2001), The lack of a glycosylated amino acid at position 6 allows more efficient use of the CXCR4 receptor by the protein. CCR5 functions better with a glycosylated amino acid (Polzer et al., 2002). The sequences were subjected to another online software that accounts for the number of N-linked glycosylation sites (http://www.hiv.lanl.gov/content/hiv-db/GLYCOSITE/glycosite.html; Zhang et al., 2004). It was found that all of the samples possessed a glycosylated asparagine, thus making them more prone to association with CCR5. So while a purely visual inspection of the three motifs in the V3 loop cannot provide definitive proof that a sequence belongs to a specific subtype, it does allow a limited comparison which narrows down the possibility

of the samples' subtypes. It is also of some interest to note that the SIV reference samples largely maintained these motifs, and even the mutations that do occur in these regions can also be found in the HIV sequences (see reference sequences on NCBI CoreNucleotide: accession no. EF535994; AF057283; Y14359). This seems to suggest that even in the simian virus, these regions are quite critical in the function of the V3 loop and interaction with the chemokine coreceptors, and is therefore conserved.

A purely visual examination of the V3 sequences also revealed the presence of single amino acid deletions, which had occurred in two of the samples. With the subtype C V3 loop, mutation of this nature is very rare, but there seems to be minimal side effects from the loss of single residues (Batra et al., 2000; Engelbrecht et al., 2001). There are no specific positions where the deletion can occur, but it appears that they rarely occur within the motifs mentioned above. This is interesting as the SI variants have a tendency to gain amino acids, and typically creating a V3 loop that is 37 amino acids long. When such an event does occur, it is often a two amino acid insertion between position 13 and 14 (Cilliers et al., 2003).

4.2.3 Predicted chemokine coreceptor usage

Using the 11/25 rule, the samples sequenced were all of the NSI phenotype, as there were no mutations at position 11, while there were only 3 samples that possessed mutations at position 25. These included glutamic acid (E) and glycine (G). As neither of these was positively charged amino acids, they made no difference in the application of this positional rule. There have also been comments made concerning mutation occurring at

position 324 of the gp160 (position 29 of the V3 loop) as also being able to influence viral phenotype (De Jong et al., 1992). Three samples had mutations at this position, but there was no such indication in them to suggest that the phenotype had changed from NSI to SI. However, the mutations presented were all aspartate (D) to asparagines (N), where the previous observations saw changes to arginine (R), which were associated with the phenotypic change.

The samples were then subjected to both PSSM matrices. In the sinsi matrix, all the samples had low scores. This was indicative of the samples being non-syncytium inducing. Similarly in the x4r5 matrix, most of the samples also had low scores which leaned towards the use of CCR5 for the virus. There was one sample (HIV 15), however, that produced a score that tended toward an X4-using virus. But at the same time, the percentage of this likelihood indicated by the x4r5 PSSM was only 51% as being a X4-user, as opposed to the 98% likelihood of it being a R5-user. Together with the value from the x4r5 matrix, as well as results from other analyses, the data strongly suggest that HIV 15 is still a NSI/R5-using virus. The final process of calculating the total charges present for the sample V3 loops showed that there were two samples that possessed total positive charges above 5 – BMHIV04 and HIV19 both possessed a charge of 5.5, which placed them between the two typical numerical parameters that classified the R5- and X4-using viruses. But as with the previous discrepancies, the samples are more likely to be CCR5-users, as the charges are not highly positive, and other data points them towards being the original phenotype that has not mutated. The PSSM had also calculated the total positive charges of the V3 loop of the samples, where the histidine was considered

as a +1 charge. This difference was reflected in the nett charges of the samples, as the HIV19 total charge was rounded up, whereas the BMHIV04 was rounded down. The difference was influenced by the overall composition of the V3 loop.

Although phenotypic prediction via the V3 loop amino acid sequences is a viable, and a usually reliable method, it remains that there is no HIV-1 subtype C specific software to calculate the likelihood of the subtype C phenotype. The software used generate the SINSI and R5X4 matrixes uses sequence data from subtype B, which have been acknowledged to have quite different disease progression to subtype C. The one caveat is that even with a subtype C specific software, the predictions are still not guaranteed to match the phenotype. So while the result is a useful indicator, it is by no means correct in all aspects – it remains a phenotype *predictor* – useful for quick results, or where a culturing facility is not available.

The bioinformatics results are strong indicators that all the sequenced samples to be subtype C, NSI, CCR5-users. Only 11 out of the 21 samples produced V3 loop data after sequencing and the subsequent analysis. The remainder of the samples may be of the same phenotype, although it was not possible to determine this. This supposition is made because subtype C is the prevalent virus in South Africa, and previous publication has shown HIV-1 in the region to be largely non-syncytium inducing, and utilizing CCR5 for entry (Abebe et al., 1999; Bjornal et al., 1997; Morris et al., 2001). Even though there is a high rate of mutation within the V3 loop, the data presented here has shown the mutations

often do not contribute to the change in chemokine co-receptor usage. This is also in line with previously published research (Morris et al 2001; Treunicht et al., 2002).

4.3 Theories and hypotheses

Current and past data has shown that HIV-1 subtype C predominantly utilizes CCR5 as their choice of chemokine coreceptor to gain entry into the viable $CD4^+$ cells from the initiation of pathogenesis, and rarely changes to using CXCR4 as the disease progresses (Abebe et al., 1999; Peeters et al., 1999; Tien et al., 1999; Morris et al., 2001). This is in stark contrast to the HIV-1 subtype B infection, where around 50% of the cases do switch to utilize CXCR4, or both of these chemokine coreceptors (Deng et al., 1996; Feng et al., 1996; Connor, et al., 1997). There is nothing at the present to suggest that the disease progression in HIV-1 subtype C undergoes the same pattern with a change chemokine coreceptor usage. But there remains the fact that towards the end of the infection by HIV-1 subtype C, the $CD4^+$ T cells are depleted in the same manner as the subtype B infections. This is occurring even though the majority of the cells express low or no CCR5, which removes the primary route of direct interaction between the virus and cells. Here, FISH and V3 loop sequencing of the virus from HIV^+, ART-naïve patients have shown HIV-1 subtype C is able to infect monocytes/macrophages and T cells, in spite of the virus being CCR5-users. It therefore present a different scenario to those previously observed.

There are a number of possible hypothesis to explain the conundrum. The first theory deals with the status of the cells of the immune system in the infected individuals. The expression of chemokine coreceptors on these cells varies greatly in a natural, non-HIV infected system, as earlier described (Wu et al., 1997; Berkowitz et al., 1998; Mo et al., 1998). The presence of HIV may well alter this, especially if the virus was striving to modify the cellular environment to enhance its infection and transmission capability. This viral intervention may have led to CCR5 being upregulated. The chronic activation of the immune system may have also contributed to the spread of the virus into uninfected $CD4^+$ T cells. Under normal circumstances, only a very small percentage of the total T cell population are activated T cells, which express CCR5 in abundance (Bleul et al., 1997; Wu et al., 1997; Berkowitz, 1998). As the immune system is constantly challenged by the virus and other opportunistic infections, the majority of the activated cells are unlikely to survive the cytopathic effects of the virus. It has been shown that the rate of death for these cells does exceed the turnover rate of the population (Grossman et al., 2000). It would therefore be possible that, due to the inability of the T cells to shut down the infections, more and more CD4 T cells become activated. But instead of successfully curbing and ending the infections, the increased activation actually makes the greater portion of the population vulnerable to infection. Those cells that do survive and become memory cells, would still maintain the integrated HIV genome. When responses from these cells are required and they become activated, the viral genome would then be expressed and new virions produced. This may explain the large number of infected cells, which occurred by clonal expansion of the previously infected memory T cells rather than individual *de novo* infections.

It may be possible that the virus was utilizing other chemokine coreceptors present on the T cells to gain entry into these cells, which cannot be totally dismissed as the prediction for chemokine coreceptor usage was only for the CCR5 and CXCR4. Having said that, this *is* a less likely scenario compared to use of other corecepors, as the use of alternative chemokine coreceptors has usually been rarities, only demonstrated in cultures and subtype B viruses (Alkhatib, et al., 1997; Alkhatib et al., 1997a; Deng et al., 1997; Farzan et al., 1997; Pohlmann, Krumbiegel & Kirchhoff, 1999). Even in such circumstances, the efficiency of the alternative chemokine coreceptors usage was weak compared to CCR5 or CXCR4 (Zhang et al., 1998).

Secondly, variants of the viral envelope protein have been demonstrated to display stronger binding affinity for CD4 and CCR5. This allows the virus to enter cells with limited numbers of these receptors (Gorry et al., 2004; Miller et al., 2001; Peeters et al., 2003). Although this only in reference to a mechanism of CNS invasion by the virus where brain macrophages and microglia expresses low levels of these receptors, it is a possible avenue for the virus to exploit to gain entry into cells elsewhere in the body, where small numbers of these receptors exist, and may impose a similar limiting factor.

Thirdly, HIV-1 is not as haphazard in its approach of infecting new cells as one might expect, purely relying on the chance that when a virion encounters a cell, that there are the necessary receptors that will allow the virus access. *In vitro* studies had indicated the exploitation of cell-to-cell interfaces – known as synapses -that can occur naturally between cells (Jolly & Sattentau, 2004). Although cells of various lineages use these

contacts to communicate with each other, only neuronal and immune system cells posses the ability to form synapses, where specific arrangements of molecules and structures are created to allow the cross talking (Taner et al., 2004). For HIV, the immunological synapses are of certain importance, because the cells involved are ones that the virus is able to infect, including dendritic cells and T cells. The synapses can form between the mucosal dendritic cells, which capture the virus, and transfers these still intact viruses through the use of multivesicular bodies (MVB) to T helper cells. This is a much more efficient method of transference when compared with the infection initiated by cell free viruses. Secondly, the T cells also create synapse between one another, which is usually used to facilitate transcriptional activation, or secretion of activating cytokines like Interleukin-5 across the synapse (Jolly & Sattentau, 2004). When this is subverted to allow the transmission of HIV-1, the synapse allows the rapid recruitment of CD4 and CXCR4 along with other components necessary for the formation of the synapse (which includes talin, actin, and LFA-1) at the point of contact. This promotes the budding of virions into the synapse cleft – spaces that form between points of contact that the cells have created – and then followed by the fusion of the virus to the uninfected cell membrane.

Although such synapses have not been demonstrated in macrophages or between macrophages and T cells, such an event may be possible, with respect to the cell lineages that have been shown *in vitro* to form synapses (i.e. the immune system cells). It would thus allow the transference of viable virions from infected macrophages to uninfected T cells. In this scenario (transference from T cell to T cell), as well as the synapse

formation that occurs between mucosal dendritic cells and T cells, it would be an ideal situation for the virus to enter the usually invulnerable portion of the T cell population. Thus, the presence of the virological synapse is of certain detriment to the control of the viral transmission. HIV could exploit a naturally occurring event in the immune system, which is even more adapted at the efficient spread of the virus to uninfected cells than the infection initiated by free-floating viruses. Chemical communication by resting T cells utilizing the synapse for reactivation, may even allow infection of previously naïve resting T cell. However, the occurrence of these events has yet to be shown to be a widely spread phenomenon and of importance in the pathogenesis of HIV-1 *in vivo*. Nonetheless, it should be seen as a possibility of virus infection and transmission that the virus could use if the conventional processes were severely restricted by the patient's immune system.

Lastly, it is possible that HIV-1 subtype C have other points of entry into $CD4^+CXCR4^+$ T cells aside from the mechanisms and scenarios proposed above. Many of its interactions and characteristics remains unclear, and it is possible that the virus is exploiting alternative pathways into $CD4^+$ T cells. HIV-1 has so far managed to thwart all attempts aimed at its full eradication from an infected system, so it would not be unexpected for the virus to possess a myriad of approaches, aside from hypothesized influence over chemokine coreceptor expression, subversion of cell synapses, and taking advantage of opportunistic infections.

There remains the possibility that the percentage of subtype C X4/SI isolates is not an accurate one. Most of the publications concerning subtype C had shown the existence of the SI isolate to be extremely rare, ranging from none to 6% of the sample population (Abebe et al., 1999; Bjornal et al., 1997; Morris et al., 2001). On the other hand, there is also data showing that this may be higher than expected, although only by approximately another 10% (Cilliers et al., 2003). In all these cases, the sample population size were no bigger than 50. Furthermore, the samples were also all PBMC samples, and did not provide any data concering the virus' chemokine coreceptor usage in other compartments. This may lead to a bias in the use of CXCR4 by the virus.

Therefore, the number and percentage produced may provide a good indication of the trends and pattern of infection, but they are no by means a comprehensive model for subtype C. So while it is acceptable to infer the rarity of X4/SI isolates in HIV-1 subtype at this point, the situation where the existence of X4/SI isolate in subtype C being the similar rate of occurrence as subtype B cannot be ruled out entirely. This would also explain the depletion of the $CD4^+$ T cells in HIV-1 subtype C infections. Nonetheless, this would only apply for X4-using viruses, and certainly cannot account for the virus detected in this study, which has been determined to be CCR5-users.

4.4 Other observations from the study

The use of FISH to determine the sites of replication in the late stage, ART-naïve system

was informative to some extent. It had been previously assumed that the determinant to the success of active replication by the virus was predominantly based on the integration of the viral cDNA into the host genome. This has been shown to be incorrect, as the replication is actually dependent on the activation of the infected cell (Stevenson et al., 1990, Korin & Zack, 1998; Korin & Zack, 1999). The factors that prevent active replication by the virus are particularly important when dealing with the issue of the latent reservoir, as they in part form the block that keeps the replication rate of the virus in check. The direct infection of resting $CD4^+$ T cells can occur, but it usually does not lead to the integration of the virus. However, isolation of pure populations of resting $CD4^+$ T cells in infected individuals have been found to contain low level of integrated DNA (Bukrinsky, et al., 1992; Chun et al., 1997). This would be possible through the infection of previously activated T cells, and have managed to revert to the resting state to be reactivated sometime in the future, as part of the immunologic memory response. Most of the infected T cells detected in peripheral blood would theoretically be of naïve phenotype, as the majority of T cells in whole blood are of that group. But as naïve T cells have no CCR5 expression, and only CXCR4, it means that these infected T cells could be memory T cells, as they do present CCR5 on the surface (Bleul et al., 1997; Berkowitz et al, 1998a). Alternatively, the chronic activation of the immune system may have led to a very low level of naïve T cells in circulation. The T cells would thus either be activated, or memory T cells. The latter would also become the dominant T cell populations in the peripheral blood (Bajaria et al., 2002).

Therefore, the factors that limit the latent reservoir are the same as the one that limit sites of active viral replication. The HIV-gag probe obviously cannot take these components into account, and so cannot provide a successful description of these cells. FISH is still able to reveal the existence of cells that has viral genetic materials within their nuclei. This was positively correlated with T cells of HIV^+ patients, both in PBMCs, as well as in the bone marrow extracts. This also included the monocytes/MDMs within the two compartments, but these are less controversial when the chemokine coreceptor usage is taken into account, with them possessing the appropriate entry coreceptors already.

The overwhelming presence of infected T cells suggests that the depletion may be a consequence of infected cell death. The main argument against direct killing mechanisms have been the small number of cells the virus infects, as the overwhelming majority of $CD4^+$ T cells are usually inactive, and so cannot be infected by the virus. In this study, the virus appears to have overcome the problem. This suggests the direct interaction of the virus in infecting the T cells are of significant contribution to $CD4^+$ T cell depletion.

The number of signals detected in $CD4^+$ cells in this study was also contradicted the experimental results obtained in early HIV studies. Viral DNA detection put the number of cells harbouring the virus at 1 in 100 cells (Psallidopolous et al., 1989; Schnittmann et al., 1989), while the expression of HIV-1 mRNA by infected cells were in one to two order of magnitude lower (Harper et al., 1986; Embretson et al., 1993; Embretson et al., 1993a). The results obtained here were of several magnitudes higher. Almost all the viable $CD4^+$ cells were infected, and also possessed signals in the teens. The possible

explanation for the vast difference could be due to the difference in the experimental parameters. These earlier studies were *in vitro*, cultured experiments, and also used sample that were from HAART patients. Although the variant of HIV-1 was not stated in these experiments, it is likely that they were all subtype B infection, as these studies were done in USA and/or Europe. These are quite different parameters compared to those for this study, and may contribute to the dissimilar results. It would be useful to resolve by performing the experiments in this study on early stage HIV infection samples obtain in South Africa.

4.5 Future work & prospects

Although FISH and the V3 loop sequence analyses provided data and information towards the HIV localization in late stage, ART-naïve, subtype C HIV-1^{+} patients, there are still a number of questions that need to be answered.

The first concern lies within the collection of bone marrow samples. Bone marrow specimens are removed from the body via spinal puncture, which is a procedure that taps into the subarachnoid space between the lumbar regions of the vertebrae to obtain marrow material. The procedure is not performed upon the patient unless it is absolutely necessary, and the samples obtained from this procedure are used for morphological tests, like those for lymphomas. Consequently, the samples obtained in these cases, while they are HIV^{+}, also possess other anomalies in the bone marrow, which can distort the picture

presented on the slide. Although these anomalies were not observed on the slides in this study, it is unknown what effects they have on the cells and cellular environment. A purely HIV^+ bone marrow biopsy without the occurrences of marrow-related abnormalities would be useful to obtain and to be examined. But such samples would be unethical to obtain. The same principle applies with obtaining any other types of biopsies, as such invasive procedures are not done lightly, and only goes ahead because the tissues in question are diseased and a direct examination is necessary. Although there was no follow up on the bone marrow samples obtained whether they are confirmed lymphoma cases, the samples were taken because there were causes to suspect cancer. This situation is an unfortunate one, as it would be of great use to observe what interaction occurs between the virus infection in a normal, healthy bone marrow, or any other tissues that can be affected by the HIV-1 infection. All the bone marrow samples in this study were obtained as there was a need for a biopsy of the tissue. Although they were all AIDS-related lymphoma cases, it does not detract from the possibility the cancers may have influenced the pattern of infection, transmission, and integration of HIV-1 in the bone marrow.

With respect to FISH itself, one of the problems in the analysis is the lack of delineation between the cell cytoplasm and the intercellular spaces. Nuclei of the cells examined were positively labeled through the use of DAPI, and was the only cellular component visible under fluorescent microscopy. The visualization of coherent nuclei is a reliable way of detecting intact cells. Detection of signals showed that they were predominantly located within, or on the periphery of the nuclei. But there were also stray signals that

were situated in the void between the cell nuclei. Some were sufficiently close to the cell nuclei to be considered as being within a cell. But this is obviously quite subjective. An additional stain to define the cell cytoplasm would thus be of some importance when deciding whether the extra-nucleic signals are associated with the cells or not. Secondly, while the nucleus-associated signals are clearly distinguishable, the fact that the nucleus is a three-dimensional structure, makes it sometimes difficult to differentiate and count the individual signals. The analysis should therefore benefit from the use of confocal microscopy that can easily produce layered images of the nucleus, and make it easier to account for the signals. Thirdly, there is no specific association of the signals with the nucleic components. That is, although it is visible and fairly obvious that the signals are inside the nucleus, the cellular compartment remains one amorphous mass of genetic material. It would be highly interesting to see if the signals are integrated into specific chromosomes, or area of the chromosomes with which the virus is predisposed to associate. This could be achieved by culturing infected PBMC and induce the cells into mitosis, and block them in the metaphase stage where the individual chromosomes can be visualized. HIV-gag probes could then be hybridized to the metaphase spreads.

These possible improvements do not actually subtract from the fact that the HIV-gag probe and the accompanying protocol was a success in detecting the presence of the viral genetic material. The hybridization can be adapted to be used on the other HIV-infected tissues, such as HIV-related lymphoma, or forms of biopsies, that still possess fairly intact nucleic acid species.

Conclusion

We utilized FISH and sequencing to determine the cells that host HIV-1 in late infection, i.e. $CD4^+$ monocytes/MDMs, and/or $CD4^+$ T cells. Using FISH, we were able to observe the location of the viral genetic material in a minimally disrupted cellular environment. The V3 loop sequences and their input into a number of computational programs predicted chemokine coreceptor usage, virus subtype, and several other sequence based observations. The combined results revealed that in late stage, ART-naïve and OI-free infections, HIV-1 subtype C remains a R5/NSI using virus. The virus was found in $CD4^+$ cells in both whole blood and bone marrow. This was regardless of the assumed chemokine coreceptor expression, where monoocytes/macrophages preferentially express CCR5, as opposed to CXCR4 on the majority of T cells. Therefore, there is a discrepancy with the HIV-1 subtype C chemokine coreceptor usage, and the presence of viral genetic material in the nuclei of T cells.

We suggest several scenarios where HIV-1 subtype C may gain access into the T cells. Firstly, a change in CCR5 expression patterns, which may occur through viral pressure, or responses to suppression of opportunistic infections, is possible. The viral synapse, which is a subversion of the cell synapses used for cell communication and crosstalking, is also a possible route exploited by the virus. A sample population skewed due in the collection process would have contributed to the presence of virus in $CD4^+$ T cells without the appropriate chemokine coreceptor. However, this remains an unlikely explanation. The use of other chemokine receptors outside of CCR5 and CXCR4, may

also contribute to this. Additionally, there could be other, yet unknown pathways that HIV-1 is using to circumvent the limiting factor of chemokine coreceptor.

Regardless of how the virus enters the $CD4^+$ T cells, the fact remains that in HIV-1 subtype C infections, $CD4^+$ monocytes/MDMs and T cells exists with HIV-specific genomic material. This poses the same obstacles in the treatment of the disease as with HIV-1 subtype B. Even though the virus was present in $CD4^+$ cells, little else is known about their status and degree of interaction with the virus. For example, what chemokine receptors the infected cells were expressing in high concentrations; whether the infected T cells are naïve or memory cells; are the aforementioned cells in a resting or activated state; where in the chromosomes is the HIV genetic material integrated. Data on these questions would provide a better picture of what was occurring in these infected cells, and at present, work has begun on some of these aspects in order to produce data.

To summarize: FISH can be a useful tool in HIV/AIDS research, and while it is not a method that can provide fully satisfactory answer on its own, it can be highly informative in conjunction with other molecular techniques. V3 loop sequencing and subsequent data analysis using multiple programs is also reliable, and can provide additional description of the region. With respect to the results, the viral presence within all $CD4^+$ cells with confirmed usage of CCR5 and not CXCR4 poses an interesting dilemma. This requires a more detailed look into the virus-cell interaction; although this should probably take place *in vivo*, as to our knowledge, there are no hypothesis that has come out of *in vitro* experiments that can explain the contradicting phenomena.

We have shown the HIV-1 subtype C virus to be a persistent entity, able to infect cells without a direct route of infection. However, parts of the viral interaction remain unclear, and in-depth experimentations will be necessary to understanding these aspects. Knowledge of these issues will be important for the ultimate result of dealing with the disease, which has become a worldwide endemic.

References

Abebe, A., Demissie, D., Goudsmit, J., Brouwer, M., Kuiken, C. L., Pollakis, G., Schuitemaker, H., Fontanet, A. L., Rinke de Wit, T. F. (1999). **HIV-1 subtype C syncytium- and non-syncytium-inducing phenotypes and coreceptor usage among Ethiopian patients with AIDS**. AIDS 13(11):1305-11.

Adams, M., Sharmeen, L., Kimpton, J., Romeo, J. M., Garcia, J. V., Peterlin, B. M., Groudine, M., Emerman, M. (1994). **Cellular latency in human immunodeficiency virus-infected individuals with high CD4 levels can be detected by the presence of promoter-proximal transcripts**. PNAS 91(9):3862-6.

Alkhatib, G., Combadiere, C., Broder, C. C., Feng, Y., Kennedy, P. E., Murphy, P. M., Berger, E. A. (1996). **CC-CKR5: A RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage tropic HIV-1**. Science 272:1955-8.

Alkhatib, G., Ahuja, S. S., Light, D., Mummidi, S., Berger, E. A., Ahuja, S. K. (1997). **CC chemokine receptor 5-mediated signaling and HIV-1 Co-receptor activity share common structural determinants. Critical residues in the third extracellular loop support HIV-1 fusion**. J Biol Chem 272(32):19771-6.

Alkhatib, G., Berger, E. A., Murphy, P. M., Pease, J. E. (1997a). **Determinants of HIV-1 coreceptor function on CC chemokine receptor 3. Importance of both extracellular and transmembrane/cytoplasmic regions**. J Biol Chem 272(33):20420-6.

Ameisen, J. C. (1992). **Programmed cell death and AIDS: from hypothesis to experiment**. Immunol Today 13(10):388-91.

Asjo, B., Morfeldt-Manson, L., Albert, J., Biberfeld, G., Karlsson, A., Lidman, K., Fenyo, E. M. (1986). **Replicative capacity of human immunodeficiency virus from patients with varying severity of HIV infection**. Lancet 2(8508):660-2.

Barjaria, S. H., Webb, G., Cloyd, M., Kirschner, D. (2002). **Dynamics of naïve and memory CD4$^+$ T lymphocytes in HIV-1 disease progression**. JAIDS 30:41-58.

Barre-sinoussi, F., Chermann, J. C., Roy, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-brun, F., Rouzioux, C., Rozenbaum, W., Montagnier, L. (1983). **Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS)**. Science 220:868-71.

Barmak, K., Harhaj, E., Grant, C., Alefantis, T., Wigdahl, B. (2003). **Human T cell leukemia virus type I-induced disease: pathways to cancer and neurodegeneration**. Virology 308(1):1-12

Batra, M., Tien, P. C., Shafer, R. W., Contag, C. H., Katzenstein, D. A. (2000). **HIV type 1 envelope subtype C sequences from recent seroconverters in Zimbabwe**. AIDS Res Hum Retroviruses 16(10):973-9.

Berkowitz, R. D., Beckerman K. P., Schall T. J., McCune, J. M. (1998, Oct). **CXCR4 and CCR5 expression delineates targets for HIV-1 disruption of T cell differentiation**. J Immunol 161(7):3702-10.

Berkowitz, R. D., Alexander, S., Bare, C., Linquist-Stepps, V., Bogan, M., Moreno, M. E., Gibson, L., Wieder, E. D., Kosek, J., Stoddart, C. A., McCune, J. M. (1998a, Dec). **CCR5- and CXCR4-utilizing strains of human immunodeficiency virus type 1 exhibit differential tropism and pathogenesis in vivo**. J Virol 72(12):10108-17.

Betts, M. R., Ambrozak, D. R., Douek, D. C., Bonhoeffer, S., Brenchley, J. M., Casazza, J. P., Koup, R. A., Picker, L. J. (2001). **Analysis of total human immunodeficiency virus (HIV)-specific CD4(+) and CD8(+) T-cell responses: relationship to viral load in untreated HIV infection**. J Virol 75(24):11983-91.

Berkelman, R. L., Heyward W. L., Stehr-Green J. K., Curran, W J. (1989). **Epidemiology of human immunodeficiency virus infection and acquired immunodeficiency syndrome**. Amer J Med 86:761-70.

Bjorndal, A., Deng, H., Jansson, M., Fiore, J. R., Colognesi, C., Karlsson, A., Albert, J., Scarlatti, G., Littman, D. R., Fenyo, E. M. (1997). **Coreceptor usage of primary human immunodeficiency virus type 1 isolates varies according to biological phenotype**. J Virol 71(10):7478-87.

Blankson, J. N., Persaud, D., Siliciano, R. F. (2002). **The challenge of viral reservoirs in HIV-1 infection**. Annu Rev Med 53:557-93.

Bleul, C. C., Wu, L., Hoxie, J. A., Springer, T. A., Mackay, C. R. (1997). **The HIV coreceptors CXCR4 and CCR5 are differentially expressed and regulated on human T lymphocytes**. PNAS 94(5):1925-30.

Boldt-Houle, D. M., Rinaldo, C. R. Jr, Ehrlich, G. D. (1993). **Random depletion of T cells that bear specific T cell receptor V beta sequences in AIDS patients**. J Leuko Biol 54(5):486-91.

Borzi, R. M., Piacentini, A., Monaco, M. C. G., Lisignoli, G., Degrassi, A., Cattini, L., Santi, S., Facchini, A. (1996). **A fluorescent *in situ* hybridization method in flow cytometry to detect HIV-1 specific RNA**. J immunol methods 193:167-76.

Bukrinsky, M. I., Stanwick, T. L., Dempsey, M. P., Stevenson, M. (1991). **Quiescent T lymphocytes as an inducible virus reservoir in HIV-1 infection**. Science 254(5030):423-7.

Bukrinsky, M. I., Sharova, N., Dempsey, M. P., Stanwick, T. L., Bukrinskaya, A. G., Haggerty, S., Stevenson, M. (1992). **Active nuclear import of human immunodeficiency virus type 1 preintegration complexes**. PNAS 89(14):6580-4.

Caffrey, M., Cai, M., Kaufman, J., Stahl, S. J., Wingfield, P. T., Covell, D. G., Gronenborn, A. M., Clore, G. M. (1998). **Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41**. EMBO J 17(16):4572-84.

Chams, V., Idziorek, T. & Klatzmann, D. (1991). **Biological properties of anti-CD4 autoantibodies purified from HIV-infected patients**. AIDS 5(5):565-9.

Chen, J. J-Y. & Cloyd M. W. (1999). **The potential importance of HIV-induction of lymphocyte homing to lymph nodes**. Int Immunol 11(10):1591-4.

Chun, T-W., Finzi, D., Margolick, J., Chadwick, K., Schwartz, D., Siliciano, R. F. (1995). ***In vivo*** **fate of HIV-1 infected T cells: quantitative analysis of the transition to stable latency**. Nat Med 1(12):1284-90.

Chun, T. W., Carruth, L., Finzi, D., Shen, X., DiGiuseppe, J. A., Taylor, H., Hermankova, M., Chadwick, K., Margolick, J., Quinn, T. C., Kuo, Y. H., Brookmeyer, R., Zeiger, M. A., Barditch-Crovo, P., Siliciano, R. F. (1997). **Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection**. Nature 387(6629):183-8.

Chun, T-W. & Fauci, A. S. (1999). **Latent reservoirs of HIV: obstacles to the eraditcaiton of virus**.PNAS 96:10958-61.

Cilliers, T. & Morris, L. (2002), **coreceptor usage and biological phenotypes of HIV-1 isolates**. Clin chem lab med 40(9):911-7.

Cilliers, T., Nhlapo, J., Coetzer, M., Orlovic, D., Ketas, T., Olson, W. C., Moore, J. P., Trkola, A., Morris, L. (2003). **The CCR5 and CXCR4 coreceptors are both used by human immudeficiency virus type 1 primary isolates from subtype C**. J Virol 77(7):4449-56.

Clavel, F., Guetarel, D., Brun-Vezinet, F., Chamaret, S., Rey, M. A., Santos-Ferreira M. O., Laurent, A. G., Dauget, C., Katlana, C., Rouzioux, C., Klatzman, D., Champalimaud, J. L., Montagnier, L. (1986). **Isolation of a new human retrovirus from West African patients with AIDS**. Science 233:343-6.

Clavel, F., Mansinho, K., Chamaret, S., Guetard, D., Favier, V., Nina, J., Santos-Ferreira, M. O., Champalimaud, J. L., Montagnier, L. (1987). **Human immunodeficiency virus type 2 infection associated with AIDS in West Africa**. NEJM 316(19):1180-5.

Clerici, M., Stocks, N. I., Zajac, R. A., Boswell, R. N., Lucey, D. R., Via, C. S., Shearer, G. M.(1989). **Detection of three distinct patterns of T helper cell dysfunction in asymptomatic, human immunodeficiency virus-seropositive patients. Independence of CD4+ cell numbers and clinical staging**. J Clin Invest 84(6):1892-9.

Cloyd, M. W., Chen, J. J., Adeqboyega, P., Wang, L. (2001). **How does HIV cause depletion of CD4 lymphocytes? A mechanism involving virus signaling through its cellular receptors**. Curr Mol Med 1(5):545-50.

Cocchi, F., Devico, A. L., Garzino-Demo, A., Arya, S. K., Gallo, R. C., Lusso, P. (1995). **identification of RANTES, MIP-1α, and MIP-1β as the major HIV-suppressive factors produced by $CD8^+$ T cells**. Science 270:1811-5.

Cone, R. W. & Schlaepfer, E. (1997). **Improved *in situ* hybridization to HIV with RNA probes derived from PCR products**. J Histochem Cytochem 45(5):721-7.

Connor, R. I., Sheridan, K. E., Ceradini, D., Choe, S., Landau, N. R. (1997). **Change in coreceptor use coreceptor use correlates with disease progression in HIV-1--infected individuals**. J Exp Med 185(4):621-8.

Crowe S. (1995). **Role of macrophages in the pathogenesis of human immunodeficiency virus (HIV) infection**. Aust N Z J Med 25:777-83.

De Jong, J. J., De Ronde, A., Keulen, W., Tersmette, M., Goudsmit, J. (1992). **Minimal requirements for the human immunodeficiency virus type 1 V3 domain to support the syncytium-inducing phenotype: analysis by single amino acid substitution**. J Virol 66(11):6777-80.

De Maria, A., Pantaleo, G., Schnittman, S. M., Greenhouse, J. J., Baseler, M., Orenstein, J. M., Fauci, A.S. (1991). **Infection of CD8+ T lymphocytes with HIV. Requirement for interaction with infected CD4+ cells and induction of infectious virus from chronically infected CD8+ cells**. J Immunol 146(7):2220-6.

De Wolf, F., Hogervorst, E., Goudsmit, J., Fenyo, E. M., Rubsamen-Waigmann, H., Holmes, H., Galvao-Castro, B., Karita, E., Wasi, C., Sempala, S.D., et al. (1994). **Syncytium-inducing and non-syncytium-inducing capacity of human immunodeficiency virus type 1 subtypes other than B: phenotypic and genotypic characteristics. WHO Network for HIV Isolation and Characterization**. AIDS Res Hum Retroviruses 10(11):1387-400.

Deichmann, M., Bentz, M. & Haas, R. (1996). **Ultra-sensitive FISH is a useful tool for studying chronic HIV-1 infection**. J virol methods 65:19-25.

Deng, H., Liu, R., Ellmeier, W., Choe, S., Unutmaz, D., Burkhart, M., Di Amrzio, P., Marmon, S., Sutton, R. E., Hill C. M., Davis, C. B., Peiper, S. C., Schall, T. J., Littman, D. R., Landau, N. R. (1996). **Identification of a major co-receptor for primary isolate of HIV-1**. Nature 381: 661-7.

Douglas, S. D., Durako, S., Sullivan, K. E., Camarca, M., Moscicki, A. B., Wilson, C. M. **TH1 and TH2 cytokine mRNA and protein levels in human immunodeficiency virus (HIV)-seropositive and HIV-seronegative youths**. Clin Diagn Lab Immunol 10(3):399-404.

Dragic, T., Litwin, V., Allaway, G. P., Martin, S. R., Huang, Y., Nagashima, K. A., Cayanan, C., Maddon, P. J., Koup, R. A., Moore, J. P., Paxton, W. A. (1996). **HIV-1 entry into CD4$^+$ T cells is mediated by the chemokine receptor CC-CCKR5**. Nature 381:667-73.

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

Emerman, M. & Malim, M. H. (1998). **HIV-1 regulatory/accessory gene: keys to unraveling viral and host cell biology**. Science 280:1880-4.

Engelbrecht, S., de Villiers, T., Sampson, C. C., zur Megede, J., Barnett, S. W., van Rensburg, E. J. (2001). **Genetic analysis of the complete gag and env genes of HIV type 1 subtype C primary isolates from South Africa**. AIDS Res Hum Retroviruses 17(16):1533-47.

Farzan, M., Choe, H., Martin, K., Marcon, L., Hofmann, W., Karlsson, G., Sun, Y., Barrett, P., Marchand, N., Sullivan, N., Gerard, N., Gerard, C., Sodroski, J. (1997). **Two**

orphan seven-transmembrane segment receptors which are expressed in CD4-positive cells support simian immunodeficiency virus infection. J Exp Med 186(3):405-11.

Fauci, A. S. (1988). **The human immunodeficiency virus: infectivity and mechanism of pathogenesis**. Science 239:617-22.

Feinberg, M. B., Jarrett, R. F., Aldovini, A., Gallo, R. C., Wong-Staal, F. (1986). **HTLV-III expression and production involve complex regulation at the levels of splicing and translation of viral RNA**. Cell 46(6):807-17.

Feng, Y., Broder, C. C., Kennedy, P. E., Berger, E. A. (1996). **HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G-protein-coupled receptor**. Science 272:872-7.

Finkel, T. H., Tudor-Williams, G., Banda, N. K., Cotton, M. F., Curiel, T., Monks, C., Baba, T. W., Ruprecht, R. M., Kupfer, A. (1995). **Apoptosis occurs predominantly in bystander cells and not in productively infected cells of HIV- and SIV-infected lymph nodes**. Nat Med 1(2):129-34.

Finzi, D., Blankson, J., Siliciano, J. D., Margolick, J. B., Chadwick, K., Pierson, T., Smith, K., Lisziewicz, J., Lori, F., Flexner, C., Quinn, T. C., Chaisson, R. E., Rosenberg, E., Walter, B., Gange, S., Gallant, J, Siliciano, R. F. (1999). **Latent infection of $CD4^+$ T cells provides a mechanism for lifelong persistence of HIV-1 infection, even in patients on effective combination therapy**. Nat Med 5(5):512-7.

Fouchier, R. A., Groenink, M., Kootstra, N. A., Tersmette, M., Huisman, H. G., Miedema, F., Schuitemaker, H. (1992). **Phenotype-associated sequence variation in the third**

variable domain of the human immunodeficiency virus type 1 gp120 molecule. J Virol 66(5):3183-7.

Fox, C. H., Tenner-Racz, K., Racz, P., Firpo, A., Pizzo, P. A., Fauci, A. S. (1991). **Lymphoid germinal centers are reservoirs of human immunodeficiency virus type 1 RNA**. J Infect Dis 164:1051-7.

Gallo, R. C., Salahuddin, S. Z., Popovic, M., Shearer, G. M., Kaplan, M., Haynes, B. F., Palker, T. J., Redfield, R., Oleske, J., Safai, B., White, G., Foster, P., Markham, P. D. (1984). **Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS**. Science 224:500-3.

Gibellini, D., Re, M. C., Vitone, F., Rizzo, N., Maldini, C., La Placa, M., Zauli, G. (2003). **Selective up-regulation of functional CXCR4 expression in erythroid cells by HIV-1 Tat protein**. Clin Exp Immunol131:428-35.

Golding, H., Robey, F. A., Gates, F.T. 3rd, Linder, W., Beining, P. R., Hoffman, T., Golding, B. (1988). **Identification of homologous regions in human immunodeficiency virus I gp41 and human MHC class II beta 1 domain. I. Monoclonal antibodies against the gp41-derived peptide and patients' sera react with native HLA class II antigens, suggesting a role for autoimmunity in the pathogenesis of acquired immune deficiency syndrome**. J Exp Med 167:914-23.

Gorry, P. R., Sterjovski, J., Churchill, M., Witlox, K,, Gray, L., Cunningham, A., Wesselingh, S. (2004). **The role of viral coreceptors and enhanced macrophage tropism in human immunodeficiency virus type 1 disease progression**. Sex Health 1(1):23-34.

Grossman, Z. & Paul, W. E. (2000). **The impact of HIV on naive T-cell homeostasis**. Nat Med 6(9):976-7.

Grossman, Z., Meier-Schellersheim, M., Sousa, A. E., Victorino, R. M. M., Paul, W. E. (2002). **$CD4^+$ T-cell depletion in HIV-infection: are we closer to understanding the cause**? Nat Med 8(4):319-23.

Groux, H., Tropier, G., Monte, D., Mouton, Y., Carpon, A., Ameison, J. C. (1992). **Activation-induced death by apoptosis in $CD4^+$ T cells from human immunodeficiency virus-infected asymptomatic individuals**. J Exp Med 175:331-9.

Han, Y., Lassen, K., Monie, D., Sedaghat, A. R., Shimoji, S., Liu, X., Pierson, T. C., Margolick, J. B., Siliciano, R. F., Siliciano, J. D. (2004). **Resting CD4+ T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes**. J Virol 78(12):6122-33.

Harper, M. E., Marselle, L. M., Gallo, R. C., Wong-Staal, F. (1986). **Detection of lymphocytes expressing human T-lymphotropic virus type III in lymph nodes and peripheral blood from infected individuals by *in situ* hybridization**. PNAS 83:772-6.

Hellerstein, M., Hanley, M. B., Cesar, D., Siler, S., Papageorgopoulous, C., Wieder, E., Schmidt, D., Hoh, R., Neese, R., Macallan, D., Deeks, S., McCune, J. M. (1999). **Directly measured kinetics of circulating T lymphocytes in normal and HIV-1 infected humans**. Nat Med 5(1):83-9.

Herbein, G., Van Lint, C., Lovett, J. L., Verdin, E. (1998). **Distinct mechanisms trigger apoptosis in human immunodeficiency virus type 1-infected and in uninfected bystander T lymphocytes**. J Virol 72(1):660-70.

Ho, D. D., Rota, T. R., Hirsch, M. S. (1986). **Infection of monocyte/macrophages by human T lymphotropic virus type III**. J Clin Invest 77(5):1712-5.

Ho, D. D., Pomerantz, R. J., Kaplan, J. C. (1987). **Pathogenesis of infection with human immunodeficiency virus**. NEJM 317(5):278-86.

Ho, D. D., Neumann, A. U., Perelson, A. S., Chen, W., Leonard, J. M., Markowitz, M. (1995). **Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection**. Nature 373:123-6.

Hoffman, N. G., Seillier-Moiseiwitsch, F., Ahn, J., Walker, J. M., Swanstrom, R. (2002). **Variability in the human immunodeficiency virus type 1 gp120 Env protein linked to phenotype-associated changes in the V3 loop**. J Virol 76(8):3852-64.

Horuk, R., Hesselgesser, J., Zhou, Y., Faulds, D., Halks-Miller, M., Harvey, S., Taub, D., Samson, M., Parmentier, M., Rucker, J., Doranz, B. J., Doms, R. W. (1998). **The CC chemokine I-309 inhibits CCR8-dependent infection by diverse HIV-1 strains**. J Biol Chem 273(1):386-91.

Imami, N., Pires, A., Hardy, G., Wilson, J., Gazzard, B., Gotch, F. (2002). **A balanced type 1/type 2 response is associated with long-term nonprogressive human immunodeficiency virus type 1 infection**. J Virol. 76(18):9011-23.

Imlach, S., McBreen, S., Shirafuji, T., Leen, C., Bell, J. E., Simmonds, P. (2001). **Activated peripheral CD8 lymphocytes express CD4 in vivo and are targets for infection by human immunodeficiency virus type 1**. J Virol 75(23):11555-64.

Izmavola, E., Bertley, F. M. N., Huang, Q., Makori, N., Miller, C. J., Young, R. A., Aldovini, A. (2003). **HIV-1 tat reprograms immature dendritic cells to express chemoattractants for activeated T cells and macrophages**. Nat Med 9(2):191-7.

Jardetzsky, T. S., & Lamb, R. A. (2004). **A class act**. Nature 427: 307-8.

Jensen, M. A., Li, F. S., van 't Wout, A. B., Nickle, D. C., Shriner, D., He, H. X., McLaughlin, S., Shankarappa, R., Margolick, J. B., Mullins, J. I. (2003). **Improved coreceptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of human immunodeficiency virus type 1 env V3 loop sequences**. J Virol 77(24):13376-88.

John, H. A., Birnstiel, M. L. & Jones, K. W. (1969). **RNA-DNA hybrids at the cytological level**. Nature 223:583-7.

Jolly, C. & Sattentau, Q. J. (2004). **Retroviral spread by induction of virological synapses**. Traffic 5(9):643-50.

Kalyanaraman, V. S, Sarngadharan, M. G., Robert-Guroff, M., Miyoshi, I., Golde, D., Gallo, R. C. (1982). **A new subtype of human T-cell leukemia virus (HTLV-II) associated with a T-cell variant of hairy cell leukemia**. Science 218(4572):571-3.

Kilby, J. M. & Eron J. J. (2003). **Novel therapies based on mechanisms of HIV-1 cell entry**. NEJM 348:2228-38.

Knuchel, M. C., Graf, B., Schlaepfer, E., Koster, H., Fischer, M., Weber, R. Cone, R. W. (2000). **PCR-derived ssDNA probes for fluorescent *in situ* hybridization to HIV-1 RNA**. J Histochem Cytochem 48(2):285-93.

Koot, M., Keet, I. P., Vos, A. H., de Goede, R. E., Roos, M. T., Coutinho, R. A., Miedema, F., Schellekens, P. T., Tersmette, M. (1993). **Prognostic value of HIV-1 syncytium-inducing phenotype for rate of CD4+ cell depletion and progression to AIDS**. Ann Intern Med 118(9):681-8.

Korin, Y. D. & Zack, J. A. (1998). **Progression to the G1b phase of the cell cycle is required for completion of human immunodeficiency virus type 1 reverse transcription in T cells**. J Virol 72(9):3161-8.

Korin, Y. D. & Zack, J. A. (1999). **Nonproductive human immunodefiency virus type 1 infection in nucleoside-treated G0 lymphocytes**. J Virol 1999, 73(8):6526-32

Kostrikis L. G., Huang, Y., Moore J. P., Wolinsky S. M., Zhang, L., Guo, Y., Deutsch, L., Phair, J., Neumann, A. U., Ho, D. D. (1998). **A chemokine receptor CCR2 allele delays HIV-1 disease progression and is associate with a CCR5 promoter mutation**. Nat Med 4(3):350-3.

Kostrikis, L. G., Neumann, A. V., Thomson, B., Korber, McHardy, P., Karanicolas, R., Deutsch, L., Huang, Y., Lew, J. F., McIntosh, K., Pollack, H., Borkowsky, W., Spiegel, H. M. L., Palumbo, P., Oleske, J., Bardeguez, A., Luzuriaga, K., Sullivan, J., Wolinsky, S. M., Koup, R. A., Ho, D. D., Moore, J. P. (1999). **A polymorphism in the regulatory region of the CC-chemokine receptor 5 gene influences perinatal transmission of human immunodeficiency virus type 1 to African-American infants**. J Virol 73(12):10264-71.

Kumar, S., Tamura, K. & Nei, M. (2004). **MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment**. Brief Bioinform 5(2):150-63.

Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., Hendrickson, W. A. (1998). **Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody**. Nature 393(6686):648-59.

Levsky, J. M. & Singer, R. H. (2003). **Fluorescent in situ hybridization: past, present and future**. J Cell Sci 116(14): 2833-2838.

Li, C. J., Friedman, D. J., Wang, C., Metelev, V., Pardee, A. B. (1995). **Induction of apoptosis in uninfected lymphocytes by HIV-1 Tat protein**. Science 268(5209):429-31.

Lifson, J. D., Reyes, G. R., McGrath, M. S., Stein, B. S., Engleman, E. G. (1986). **AIDS retrovirus induced cytopathology: giant cell formation and involvement of CD4 antigen**. Science 232(4754):1123-7.

Lin, C-L., Sewell, A. K., Gao, G. F., Whelan, K. T., Phillips, R. E., Austyn J. M. (2000). **Macrophage-tropic HIV induces and exploits dendritic cell chemotaxis**. J Exp Med 192:587-593.

Lin, Y. L., Mettling C., Portales, P., Reynes, J., Clot, J., Corbeau, P. (2002). **Cell surface CCR5 density determine the postentry efficiency of R5 HIV-1 infection**. PNAS 99(24):1590-5.

Liu, R., Paxton, W. A., Choe, S., Ceradini, D., Martin, S. R., Horuk, R., MacDonald, M. E., Stuhlmann, H., Koup, R. A., Landau, N.R. (1996) **Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection**. Cell 86(3):367-77.

Mariani, R., Chen, D., Schrofelbauer, B., Navarro, F., Konig, R., Bollman, B., Munk, C., Nymark-McMahon, H., Landau, N. R. (2003). **Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif**. Cell 114(1):21-31.

Martin, M. P., Dean, M., Smith, M. W., Winkler, C., Gerrard, B., Michael, N. L., Lee, B., Doms, R. W., Margolick, J., Buchbinder, S., Goedert, J. J., O'Brien, T. R., Hilgartner, M. W., Vlahov, D., O'Brien, S. J., Carrington, M. (1998). **Genetic acceleration of AIDS progression by a promoter variant of CCR5**. Science 282(5395):1907-11.

Martin-Garcia, J., Kolson, D. L. & Gonzalez-Scarano, F. (2002). **Chemokines receptors in the brain: their role in HIV infection and pathogenesis**. AIDS 16:1709-30.

McCune, J. M. (2001). **The dynamics of $CD4^+$ T-cell depletion in HIV disease**. Nature 410:974-9.

Meng, G., Wei, X., Wu, X., Sellers, M. T., Decker, J. M., Moldoveanu, Z., Orenstein J. M., Graham, M. F., Kappes, J. C., Mestecky, J., Shaw, G. M., Smith, P. D. (2002). **Primary intestinal epithelial cells selectively transfer R5 HIV-1 to $CCR5^+$ cells**. Nat Med 8(2):150-6.

Mercure L., Phaneuf, D. & Wainberg M. A. (1993). **detection of unintegrated human immunodeficiency virus type 1 DNA in persistently infected $CD8^+$ cells**. J Gen Virol 74:2077-83.

The MeSH database. www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=**mesh**

Miller, E. D., Duus, K. M., Townsend, M., Yi, Y., Collman, R., Reitz, M., Su, L. (2001). **Human immunodeficiency virus type 1 IIIB selected for replication in vivo exhibits increased envelope glycoproteins in virions without alteration in coreceptor usage: separation of in vivo replication from macrophage tropism**. J Virol 75(18):8498-506.

Morris, L., Cilliers, T., Bredell, H., Phoswa, M., Martin, D. J. (2001). **CCR5 is the major coreceptor used by HIV-1 subtype C isolates from patients with active tuberculosis**. AIDS Res Hum Retroviruses 17(8):697-701.

Morrow, C. D., Park, J. & Wakefield, J. K. (1994). **Viral gene produces and replication of the human immunodeficiency type 1 virus**. Am J Physiol 266:C1135-56.

Mulherin, S. A., O'Brien T. R., Ioannidis, J. P., Goedert, J. J., Buchbinder, S. P., Coutinho, R. A., Jamieson, B. D., Meyer, L., Michael, N. L., Pantaleo, G., Rizzardi, G. P.,

Schuitemaker, H., Sheppard, H. W., Theodorou, I. D., Vlahov, D., Rosenberg, P. S.; international meta-analysis of HIV host genetics (2003). **Effects of CCR5-Delta32 and CCR2-64I allele on HIV-1 disease progression: the protection varies with duration of infection**. AIDS 17(3):377-87.

NDI Foundation. http://www.ndif.org/t-a.html

O'Brien, T. R., Winkler, C., Dean, M., Nelson, J. A., Carrington, M., Michael, N. L., White, G. C. 2nd (1997). **HIV-1 infection in a man homozygous for CCR5 delta 32**. Lancet 349(9060):1219.

Pantaleo, G., Grazlosi, C. & Fauci, A. S. (1993). **The immunopathogenesis of human immunodeficiency virus infection**. NEJM 328(5):327-35.

Patterson, B. K., Till, M., Otto, P., Goolsby, C., Furtado, M. R., McBride, L. J., Wolinsky, S. M. (1993). **Detection of HIV-1 DNA and messenger RNA in individual cells by PCR-driven *in situ* hybridization and flow cytometry**. Science 260:976-9.

Peeters, M., Vincent, R., Perret, J. L., Lasky, M., Patrel, D., Liegeois, F., Courgnaud, V., Seng, R., Matton, T., Molinier, S., Delaporte, E. (1999). **Evidence for differences in MT2 cell tropism according to genetic subtypes of HIV-1: syncytium-inducing variants seem rare among subtype C HIV-1 viruses**. J Acquir Immune Defic Syndr Hum Retrovirol 20(2):115-21.

Pierson, T. C., Zhou, Y., Kieffer, T. L., Ruff, C. T., Buck, C. , Siliciano, R. F. (2002). **Molecular characterization of preintegration latency in human immunodeficiency virus type 1 infection**. J Virol 76(17):8518-31.

Pohlmann, S., Krumbiegel, M. & Kirchhoff, F. (1999). **Coreceptor usage of BOB/GPR15 and Bonzo/STRL33 by primary isolates of human immunodeficiency virus type 1**. J Gen Virol 80(Pt 5):1241-51.

Polzer, S., Dittmar, M. T., Schmitz, H., Schreiber, M. (2002). **The N-linked glycan g15 within the V3 loop of the HIV-1 external glycoprotein gp120 affects coreceptor usage, cellular tropism, and neutralization**. Virology 304(1):70-80.

Pomerantz, R. J. (2001). **Residual HIV-1 infection during antiretroviral therapy: the challenge of viral persistence**. AIDS 15:1201-11.

Poulsen, A. G., Aaby, P., Larsen, O., Jensen, H., Naucler, A., Lisse, I. M., Christiansen, C. B., Dias, F., Melbye, M. (1997). **9-year HIV-2-associated mortality in an urban community in Bissau, West Africa**. Lancet 349(9056):911-4.

Pu, H., Tian, J., Flora, G., Lee, Y. W., Nath, A., Henning, B., Toborek, M. (2003). **HIV-1 Tat protein upregulate inflammatory mediators and induces monocytes invasion into the brain**. Mol Cell Neurosci 24:224-37.

Quinn, T. C. (1990). **The epidemiology of the human immunodeficiency virus**. Ann Emer Med19:225-32.

Resch, W., Hoffman, N., Swanstrom, R. (2001). **Improved success of phenotype prediction of the human immunodeficiency virus type 1 from envelope variable loop 3 sequence using neural networks**. Virology 288(1):51-62.

Reynes, J., Portales, P., Segondy, M., Baillat, V., Andre, P., Avinens, O., Picot, M. C., Clot, J., Eliaou, J. F., Corbeau, P. (2001). **CD4 T cellsurface CCR5 density as a host factor in HIV-1 disease progression**. AIDS 15(13):1627-34.

Ribeiro, R. M., Mohri, H., Ho, D. D., Perelson, A. S. (2002). ***In vivo* dynamics of T cell activation, proliferation and death in HIV-1 infection: why are $CD4^+$ but not $CD8^+$ T cells depleted**? PNAS 99(24):15572-7.

Richman, D. D. & Bozette, S. A. (1994). **The impact of the syncytium-inducing phenotype of human immunodeficiency virus on disease progression**. J Infect Dis 169:968-74.

Rowland-Jones, S. L. (1999). **HIV: the deadly passenger in dendritic cells**. Curr Biol 9:R248-50.

Saitou, N. & Nei, M. (1987). **The neighbor-joining method: a new method for reconstructing phylogenetic trees**. Mol Biol Evol 4(4):406-25.

Saksena N. K. & Potter, S. J. (2003). **Reservoirs of HIV-1 *in vivo*: implications for antiretroviral therapy**. AIDS Rev 5:3-18.

Samson, M., Libert, F., Doranz, B. J., Rucker, J., Liesnard, C., Farber, C. M., Saragosti, S., Lapoumeroulie, C., Cognaux, J., Forceille, C., Muyldermans, G., Verhofstede, C., Burtonboy, G., Georges, M., Imai, T., Rana, S., Yi, Y., Smyth, R. J., Collman, R. G., Doms, R. W., Vassart, G., Parmentier, M. (1996). **Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene**. Nature 382(6593):722-5.

Sarih, M., Maataoui, W. E. & Benslimane, A. (1996). **Analysis of apoptosis and a Th1/Th2 phenotype in HIV-infected patients**. Immunol Lett 50(3):125-9.

Sarngadharan, M. G., Popovic, M., Bruch, L., Schupbach, J., Gallo, R. C. (1984). **Antibody reactive with human T-lymphotropic retroviruses (HTLV-III) in the serum of patients with AIDS**. Science 224:506-8.

Schim van der Loeff, M. F., Aaby, P., Aryioshi, K., Vincent, T., Awasana, A. A., Da Costa, C., Pembrey, L., Dias, F., Harding, E., Weiss, H.A., Whittle, H. C. (2001). **HIV-2 does not protect against HIV-1 infection in a rural community in Guinea-Bissau**. AIDS 15(17):2303-10.

Schuitemaker, H., Koot, M., Kootstra, N. A., Dercksen, M. W., de Goede, R. E., van Steenwijk, R. P., Lange, J. M., Schattenkerk, J. K., Miedema, F., Tersmette, M. (1992). **Biological phenotype of human immunodeficiency virus type 1 clones at different stages of infection: progression of disease is associated with a shift from monocytotropic to T-cell-tropic virus population**. J Virol 66(3):1354-60.

Schupbach, J., Popovic, M., Gilden, R. V., Gonda, M. A., Sarngadharan, M. G., Gallo, R. C. (1984). **Serological analysis of a subgroup of human T-lymphotropic retroviruses (HTLV-III) associated with** AIDS. Science. 224(4648):503-5.

Shafer, R. W., Aguiniga, E., Merigan, T. C. (1995). **Quantitative analysis of syncytium-inducing and non-syncytium-inducing virus in patients infected with human immunodeficiency virus type 1**. J Clin Microbiol 33(1):212-4.

Siliciano, J. D., Kajdas, J., Finzi, D. Quinn, T. C., Chadwick, K., Margolick, J. B., Kovacs, C., Gange, S. J., Siliciano, R. F. (2003). **Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4$^+$ T cells**. Nat Med 9(6):727-8.

Simmons, R. & Siliciano, R. F. (2004). **Can antiretroviral therapy ever be stopped? An update**. AIDS Read 14(8):435-8, 441-2.

Smith, M. W., Dean, M., Carrington, M., Winkler, C., Huttley, G. A., Lomb, D. A., Goedert, J. J., O'Brien, T. R., Jacobson, L. P., Kaslow, R., Buchbinder, S., Vittinghoff, E.,

Vlahov, D., Hoots, K., Hilgartner, M. W., O'Brien, S. J. (1997). **Contrasting genetic influences of CCR2 and CCR5 variants on HIV-1 infection and disease progression. Hemophilia growth and development study (HGDS), multicentre AIDS cohort study (MACS), multicentre hemophilia cohort study (MHCS), San Francisco city cohort (SFCC), ALIVE study**. Science 277:959-65.

Somasundaran, M. & Robinson, H. L. (1987). **A major mechanism of human immunodeficiency virus-induced cell killing does not involve cell fusion**. J Virol 61(10):3114-9.

Spina, C. A., Guatelli, J. C., Richman, D. D. (1995). **Establishment of a stable, inducible form of human immunodeficiency virus type 1 DNA in quiescent CD4 lymphocytes in vitro**. J Virol 69(5):2977-88.

Stevenson, M., Stanwick, T. L., Dempsey, M. P., Lamonica, C. A. (1990). **HIV-1 replication is controlled at the level of T cel activation and proviral integration**. EMBO J 9(5):1551-60.

Stoddart, C. A., Geleziunas, R., Ferrell, S., Linquist-Stepps, V., Moreno, M. E., Bare, C. , Xu, W., Yonemoto, W., Bresnahan, P. A., McCune, J. M., Greene, W. C. (2003). **Human immunodeficiency virus type 1 Nef-mediated downregulation of CD4 correlates with Nef enhancement of viral pathogenesis**. J Virol. 77(3):2124-33.

Taner, S. B., Onfelt, B., Pirinen, N. J., McCann, F. E., Magee, A. I., Davis, D. M. (2004). **Control of immune responses by trafficking cell surface proteins, vesicles and lipid rafts to and from the immunological synapse**. Traffic 5(9):651-61.

Theodorou, I., Meyer, L., Magierowska, M., Katlama, C., Rouzioux, C. (1997). **HIV-1 infection in an individual homozygous for CCR5 delta 32. Seroco Study Group**. Lancet 349(9060):1219-20.

Tien, P. C., Chiu, T., Latif, A., Ray, S., Batra, M., Contag, C. H., Zejena, L., Mbizvo, M., Delwart, E. L., Mullins, J. I., Katzenstein, D. A. (1999). **Primary subtype C HIV-1 infection in Harare, Zimbabwe**. J Acquir Immune Defic Syndr Hum Retrovirol 20(2):147-53.

Treurnicht, F. K., Smith, T., Engelbrecht, S., Claassen, M., Robson, B. A., Zeier, M., van Rensberg, E. J. (2002). **Genotypic and phenotypic analysis of the *env* gene from South Africa HIV-1 subtype B and C isolates**. J Med Virol. 68:141-6.

Tscherning, C., Alaeus, A., Fredriksson, R., Bjorndal, A., Deng, H., Littman, D. R., Fenyo, E. M., Albert, J. (1998). **Differences in chemokine coreceptor usage between genetic subtypes of HIV-1**. Virology 241(2):181-8.

Turner, B. G., & Summers, M. F. (1999). **Structural biology of HIV**. J Mol Biol 285:1-32.

The United Nations AIDS organization. http://www.unaids.org

van't Wout, A. B., Kootstra, N. A., Mulder-Kampinga, G. A., Albrecht-van Lent, N., Scherpbier, H. J., Veenstra, J., Boer, K., Coutinho, R. A., Miedema, F., Schuitemaker, H. (1994). **Macrophage-tropic variants initiate human immunodeficiency virus type 1 infection after sexual, parenteral, and vertical transmission**. J Clin Invest. 94(5):2060-7.

Wang, Z. Q., Orlikowsky, T., Dudhane, A., Mittler, R., Blum, M., Lacy, E., Riethmuller, G., Hoffmann, M. K. (1994). **Deletion of T lymphocytes in human CD4 transgenic**

mice induced by HIV-gp120 and gp120-specific antibodies from AIDS patients. Eur J Immunol 24(7):1553-7.

Wang, L., Robb, C. W., Cloyd, M. W. (1997). **HIV induces homing of resting T lymphocytes to lymph nodes**. Virology 228(2):141-52.

Wei, X., Ghush, S. K., Taylor, M. E., Johnson, V. A., Emini, E. A., Deutsch, P., Lifson, J. D., Bonhoeffer, S., Nowak, M. A., Hahn, B. H., Saag, M. S., Shaw, G. M. (1995). **Viral dynamics in human immunodeficiency virus type I infection**. Nature 373:117-22.

Wu, L., Paxton, W. A., Kassam, N., Ruffing, N., Rottman, J. B., Sullivan, N., Choe, H., Sodroski, J., Newman, W., Koup, R. A., Mackay, C. R. (1997). **CCR5 levels and expression pattern correlate with infectability by macrophage-tropic HIV-1, in vitro**. J Exp Med 185(9):1681-91.

Zack, J. A., Arrigo, S. J., Weitsman, S. R., Go, A. S., Haislip, A., Chen, I. S. (1990). **HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure**. Cell 61(2):213-22.

Zagury, D., Bernard, J., Leonard, R., Cheynier, R., Feldman, M., Sarin, P.S., Gallo, R. C. (1986). **Long-term cultures of HTLV-III-infected T cells: a model of cytopathology of T-cell depletion in AIDS**. Science 231:850-3.

Zambruno, G., Giannetti, A., Bertazzoni, U., Girolomoni, G. (1995). **Langerhans cells and HIV infection**. Immunol Today, vol. 16, 11:520-24.

Zhang, Y. J., Dragic, T., Cao, Y., Kostrikis, L., Kwon, D. S., Littman, D. R., KewalRamani, V. N., Moore, J. P. (1998). **Use of coreceptors other than CCR5 by non-syncytium-inducing adult and pediatric isolates of human immunodeficiency virus type 1 is rare in vitro**. J Virol 72(11):9337-44.

Zhang, L., Ramratnam, B., Tenner-Racz, K., He, Y., Vesanen, M., Lewin, S., Talal, A., Racz, P., Perelson, A. S., Korber, B. T., Markowitz, M., Ho, D. D. (1999). **Quantifying residual HIV-1 replication in patients receiving combinational antiretroviral therapy**. NEJM 340(1):1605-13.

Zhang, .M, Gaschen, B., Blay, W., Foley, B., Haigwood, N., Kuiken, C., Korber, B. (2004). **Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin**. Glycobiology 14(12):1229-46.

Appendices

Appendix I: Experimental protocols

I.i. PCR/RT-PCR

I.i.i. HIV-1 Gag amplification

H1G777 (1st round) 5' – tca cct aga act ttg aat gca tgg g – 3'

H1P202 (1st round) 5' – cta ata ctg tat cat ctg ctc ctg t – 3'

H1Gag1584 (nested) 5' – aaa gat gga taa tcc tgg g – 3'

g17 (nested) 5' – tcc aca ttt cca aca gcc ctt ttt – 3'

1st round

	1X reaction
H1G777 (10 pmol/µl)	1.00 µl
H1P202 (10 pmol/µl)	1.00 µl
dNTP (10 mM)	10.0 µl
$MgCl_2$ (25 mM)	3.00 µl
10X buffer	5.00 µl
Taq polymerase	0.25 µl
AMV RT	0.20 µl
Sabax Water	27.5 µl
RNA	2.00 µl
Total	50.0µl

94 ºC 5 min

31 cycles : 94 ºC 1 min

55 ºC 45 sec

	72 ºC	90 sec
72 ºC	10 min	
4 ºC	∞	

Alternatively:

1st round (Using Qiagen 1-step kit)

	1X reaction
H1G777 (10 pmol/µl)	2.00 µl
H1P202 (10 pmol/µl)	2.00 µl
dNTP (10 mM)	0.80 µl
5X buffer	4.00 µl
RT	0.80 µl
Sabax Water	8.40 µl
RNA	2.00 µl
Total	20.0 µl

50 ºC	30 min	
94 ºC	15 min	
31 cycles :	94 ºC	1 min
	55 ºC	45 sec
	72 ºC	90 sec
72 ºC	10 min	
4 ºC	∞	

Do not place in to thermocycler until the temperature has reached 50 ºC.

2nd round

	1X
H1Gag1584 (10 pmol/µl)	1.00 µl
G17 (10 pmol/µl)	1.00 µl
dNTP (10 mM)	1.00 µl
$MgCl_2$ (25 mM)	2.50 µl
10X buffer	2.50 µl
Taq polymerase	0.25 µl
Sabax Water	13.75 µl
DNA (undiluted)	3.00 µl
Total	25.0 µl

94 ºC	5 min	
31 cycles :	94 ºC	1 min
	55 ºC	45 sec
	72 ºC	90 sec
72 ºC	10 min	
4 ºC	∞	

3-5 µl of the product are run on 2% agarose gel at 100 V for 40 minutes, and examined under UV light for the product. The product is ~460 bp.

I.i.i.a HIV-1 Gag amplification, 2nd primer set

Primer sequences:

gag A – IP protected

gag B – IP protected

gag C – IP protected

gag D - IP protected

1st round (Using Qiagen 1-step kit)

	1X reaction
Gag A (25 poml/µl)	2.00 µl
Gag B (25 poml/µl)	2.00 µl
dNTP (10 mM)	0.80 µl
5X buffer	4.00 µl
RT	0.80 µl
Sabax Water	8.40 µl
RNA	2.00 µl
Total	20.0 µl

50 ºC	30 min	
94 ºC	15 min	
30 cycles :	94 ºC	15sec
	55 ºC	30 sec
	72 ºC	3 min
72 ºC	10 min	
4 ºC	∞	

2nd round

	1X
Gag C (25 pmol/µl)	2.50 µl
Gag D (25 pmol/µl)	2.50 µl
dNTP (1 mM)	8.00 µl
$MgCl_2$ (25 mM)	4.00 µl
10X buffer	5.00 µl
Taq polymerase	0.30 µl
Sabax Water	24.7 µl
DNA (undiluted)	3.00 µl
Total	50.0 µl

94 ºC	3 min	
30 cycles :	94 ºC	30sec
	55 ºC	45 sec
	72 ºC	3 min

72 ºC 10 min

4 ºC ∞

3-5 µl of the product are run on 2% agarose gel at 100V for 40 minutes, and examined under UV light for the product. The fragment is ~1.2 kb in size.

I.i.ii. HIV-1 Env/V3 loop amplification

Primer sequences:

ED5 (1st round) 5' – atg gga tca aag cct aaa gcc atg tg – 3'

ED12 (1st round) 5' – agt gct tcc tgc tgc tcc caa gaa ccc aag – 3'

ES7 (nested) 5' – ccc gcc tgt tra atg gya gyc tag c – 3'

ES8 (nested) 5' – gcc gcc ata att cay ttc tcc aat tg – 3'

1st round (DNA)

	1X reaction
ED5 (10 pmol/µl)	1.00 µl
ED12 (10 pmol/µl)	1.00 µl
dNTP (10 mM)	0.80 µl
$MgCl_2$ (25 mM)	2.80 µl
10X buffer	5.00 µl
Taq polymerase	0.125 µl
Sabax Water	38.875 µl
DNA	5.00 µl
Total	50.0 µl

94 ºC 5 min

31 cycles : 94 ºC 1 min

55 ºC 45 sec

72 ºC 90 sec

72 ºC	10 min	
4 ºC	∞	

2nd round

	1X reaction
ES7 (10 pmol/µl)	1.00 µl
ES8 (10 pmol/µl)	1.00 µl
dNTP (10 mM)	0.80 µl
$MgCl_2$ (25 mM)	2.80 µl
10X buffer	5.00 µl
Taq polymerase	0.25 µl
Sabax Water	36.15 µl
DNA	3.00 µl
Total	50.0 µl

94 ºC	5 min	
31 cycles :	94 ºC	1 min
	55 ºC	45 sec
	72 ºC	90 sec
72 ºC	10 min	
4 ºC	∞	

3-5µl of the product are run on 2% agarose gel at 100V for 40 minutes, and examined under UV light for the product. The fragment is ~600 bp.

I.i.iii. PCR labelling

The 1st round amplification occurs as usual, and the nested following is almost the same, save the addition of fluorochromes to the master mix. The volume of water was adjusted accordingly.

2nd round

	1X
H1Gag1584 (10 pmol/µl)	1.00 µl
G17 (10 pmol/µl)	1.00 µl
dNTP (10 mM)	1.00 µl
$MgCl_2$ (25 mM)	2.50 µl
10X buffer	2.50 µl
Taq polymerase	0.25 µl
Sabax Water	12.15-11.95µl
FITC	0.6-0.8µl
DNA	4.00 µl
Total	25.0 µl

94 ºC	5 min	
31 cycles :	94 ºC	1 min
	55 ºC	45 sec
	72 ºC	90 sec
72 ºC	10 min	
4 ºC	∞	

I.i.iv. PCR for V3 loop sequencing

2nd round product from the env/V3 loop nested PCR was used. Both strands of the area to be sequenced were amplified separately before sequencing.

	1X
ES7 (1.5 pmol/µl) OR	2.00 µl
ES8 (1.5 pmol/µl)	
5X sequencing buffer	2.00 µl
Sabax water	10.0 µl
DNA	2.00 µl
Ready reaction mix	4.00 µl

25 cycles :	96 ºC	10 sec
	50 ºC	5 sec

	60 ºC	4 min
4 ºC	∞	

I.i.v. Pre-sequencing clean up

Add 90 µl 100% ethanol and 10 µl sodium acetate to the amplified products.

Spin at 2000x g for 30 minutes at 4 ºC.

Remove the supernatant by spinning upside down at 150x g for 1 minute at 4 ºC.

Add 100 µl 70% ethanol and spin again at 150x g for 5 minutes at 4 ºC.

Place at 65 ºC for 5 minutes to remove any residual liquid.

Store at 4 ºC until sequencing.

I.i.vi. Nick translation

	1X
DNA (from 2nd round, undiluted)	20.0 µl
Nick translation buffer	10.0 µl
B-mercaptoethanol	10.0 µl
Nucleotide/spectrum mix (see below)	8.00 µl
DNA polymerase	3.00 µl
Dnase I (see below)	1.00 µl
Distilled water	48.0 µl
Total	100 µl

dNTP mix

	1X
dATP (10 mM)	1.50 µl
dCTP (10 mM)	1.50 µl
dGTP (10 mM)	1.50 µl
dTTP (10 mM)	0.75 µl
Spectrum green-dUTP	7.50 µl
Distilled Water	1.00 µl
Total	30.0 µl

15 ºC 30 sec

15 ºC 75min

4 ºC ∞

Dnase I

Dilute 0.4 µl of stock Dnase I with 999.6 µl of sabax water. Use 1µl for every 20 µl of product to be digested.

I.i.vii. Probe precipitation/cleanup

For FITC PCR labelling

Add 1/10th volume of sodium acetate and 2.5x volume of ice cold 100% ethanol to the PCR product. Leave at -70 ºC for 20 minutes.

Spin at 13 000x g for 30 minutes at 4 ºC.

Remove supernatant and add 100 µl of ice cold 70% ethanol.

Spin at 15 000x g for 10 minutes at 4 ºC.

Dry the pellet and resuspend in the same volume of hybridization buffer as the PCR product before probe precipitation began.

There's also a variation of precipitating the probes for the spectrum green nick translation which requires the inactivation of enzymes to stop further activity.

For spectrum green nick translation:

After nick translation, 10 µl (denatured at 96 ºC for 3 minutes) was allowed to run on a 2% agarose gel to check the degree of digestion. The remainder of the digested/labelled product was kept at -20 ºC. If the digestion was acceptable, the probes can be inactivated completely and precipitated. If more time for digestion is required, the product can be taken from -20 ºC and further digested.

Incubate the 90 µl product at 65 ºC for 10 minutes after adding 3 µl of 0.5M EDTA, and 1 µl 10% SDS.

Add 90 µl C_0t 1 DNA (undiluted), $1/10^{th}$ volume of sodium acetate, and 2.5x volume of ice cold 100% ethanol. Leave at -70 ºC for 30 minutes.

Spin at 13 000x g for 30 minutes at 4 ºC. Pour off the supernatant and add 100 µl of ice cold 70% ethanol.

Spin at 15 000x g for 10 minutes. Remove supernatant and air dry.

Appendix II: Reagents

II.i. Eletrophoresis reagents

II.i.i. Ethylenediamine tetra-acetic acid (0.5M EDTA)

Dissolve 9.3 g Na_2EDTA in 40ml of distilled water. pH to 8 with NaOH. When it has completely dissolved, make up to 50 ml with water and autoclave. Store at room temperature.

II.i.ii. Ethidium bromide (EtBr, 10mg/ml)

Dissolve 100 mg ethidium bromide in 10 ml distilled water. Leave on stirrer overnight. Filter and store at 4 °C in the dark.

II.i.iii. 1X TAE buffer

Dissolve 40 mM Tris base (hydroxymethylaminomethane), 20 mM $CH_3COONa.3H_2O$, and 1 mM EDTA. pH to 7.2 and make up to 1 L with distilled water. Autoclave and store at room temperature.

II.i.iii. 2% agarose gel

Dissolve 0.6 g of agarose in 30 ml of 1X TAE buffer. Heat until agarose has dissolved.

Add 1.5 µl EtBr. Pour into template and allowed to set.

II.ii. FISH reagents

II.ii.i. Phosphate buffer

Solution A: KH_2PO_4. 4.54g/500 ml (pH 4.51)

Solution B: $Na_2PO_4.2H_2O$. 5.94/500 ml (pH 8.97)

41.3 ml solution A + 58.7 ml solution B for 100ml phosphate buffer stock solution.

II.ii.ii. 20X SSC (stock solution)

175 g sodium chloride (NaCl)

88.23 g tri-sodium citrate

Dissolve in 900 ml of distilled water. pH to 7, and make up to 1000 ml with distilled water.

Autoclave and store at room temperature.

II.ii.iii. Deionized filtered formamide

Add 1 full spatula of analytical grade mixed bed resin for every 100 ml of formamide.

Place on magnetic stirrer for 2 hours. Filter with Whatman No. 1 filter paper.

Store at 4 °C.

II.ii.iv. Fixative

Add 1 part acetic acid to 3 parts methanol. Use on same day.

II.ii.v. Ethanol series

70% ethanol: 7 parts ethanol plus 3 parts distilled water.

90% ethanol: 9 parts ethanol plus 1 part distilled water.

100% ethanol: use straight from stock bottle.

Store at room temperature, or -20 °C.

II.ii.vi. Hybridization buffer

50% deionized formamide

2X SSC

10% dextran sulphate

50mM sodium dihydrogen orthophosphate

Adjust pH to 7 with dihydrogen orthophosphate

For 25 ml of hybridization buffer:

12 ml 50% deionized formamide

2.5 ml 20X SSC

2.5 g detran sulphate

0.195 g sodium dihydrogen orthophosphate

II.ii.vii. Denaturing solution

35 ml deionized filtered formamide

5 ml phosphate buffer

5 ml 20X SSC

5 ml dH_2O

Adjust pH to 7 with HCl.

Can be used for hybridization of 10 slides, or be kept for up to a week at 4 ˚C.

II.ii.viii. 50% formamide washing solution

15 ml 20X SSC

65 ml dH_2O

70 ml formamide

Adjust pH to 7 with HCl.

Place 50 ml into 3 different coplin jars for washing.

II.ii.ix. 2X SSC solution

5 ml of 20X SSC solution plus 45 ml distilled water.

II.ii.x. TWEEN solution

5 ml of 20X SSC solution plus 45 ml distilled water. Add 25 µl of TWEEN-20 and mix well.

II.ii.xi. DAPI solution

5 ml of 20X SSC solution plus 45 ml distilled water. Add 50 µl of DAPI and mix well. Keep solution out of light.

Appendix III: Miscellaneous

III.i. Conversion table: molar to moles

1 Molar (M) = 1 mole/1000 ml

1 mM = 0.001 mole/1000 ml

0.001 mole/1000 ml = 1 mmole/1000 ml

1 mM = 1 μmole/ml

1 μmole/ml = 0.001μmole/μl

1 mM = 1 nmole/μl

1 μM = 1 ρmole/μl

100 μM = 100 ρmole/μl

Appendix IV: Ethical clearance

Ethical clearance was covered under the University of Witwatersrand committee for research on human subject (medical) protocol number M961129.