The Influence of HIV-1 Subtype C LTR Genotype on Latency Potential

By

Deelan Sudhir Doolabh

DLBDEE001

Supervisor: Dr Melissa-Rose Abrahams

Co-Supervisor: Professor Carolyn Williamson

Thesis submitted to the University of Cape Town in fulfilment of the degree;
MSc (Med) in Medical Virology
Division of Medical Virology
Department of Pathology
Faculty of Health Sciences
University of Cape Town
Submission Date: 05 April 2018

Declaration

I, **Deelan Sudhir Doolabh**, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: ….Signed by candidate

Date: ……28.09.2018……..

Table of Contents

Acknowledgements

I would like to show my appreciation to all those who contributed to or assisted me in producing this thesis.

I owe my deepest gratitude to my supervisor, Dr. Melissa-Rose Abrahams, for her endless guidance, support, **patience** and contribution of her extensive knowledge throughout my project. Nothing was ever too much to ask and you always went out of your way to provide me with all the help I needed. You have helped me develop as a scientist beyond my belief and are a role model as both a scientist and person that I aspire to live up to.

To my co-supervisor, Professor Carolyn Williamson, for her contributions and the incisive suggestions she provided which went a long way in the development of this study. I would also like to thank her for giving me the opportunity to be part of her research group and work in her laboratory.

To Dr. Denis Chopera and Dr. Phillipe Selhorst who contributed to the design of this study and provided preliminary data that was incorporated into the study.

I am grateful to the HIV Diversity Group for all their support and always being available to assist me with any obstacles I came across during my lab work.

I would like to thank the brave women of the CAPRISA 004 cohort and all members of the CAPRISA study team.

Special thanks to Dr. Rubina Bunjun, Dr. Ramla Tanko and Fidilia Omondi and other members of the Wendy Burgers group for their assistance with flow cytometry training.

I greatly value Sherazaan Ismail's, Smritee Dabee's, Dr. Ramla Tanko's and Fidilia Omondi's assistance with statistical and data analysis, thesis formatting and proofreading and continuous encouragement.

Thanks to my friends, with special mention to Jonty, Danica, Leah, Brendan, Donovan, Nadia, Mikhail, Michael, Heinrich and Simon your love, laughs, support and motivation is invaluable and has helped maintain my sanity during difficult times.

To the National Research Foundation (NRF), the Poliomyelitis Research Foundation (PRF), the DST-NRF Centre Of Excellence in HIV Prevention, and the University of Cape Town for providing funding for research and student support.

To Sia Furler, you will probably never read this, but music plays a huge role in my life and yours inspires me to never give up and has pulled me through some very stressful moments over the last couple of months☺.

Finally, I am indebted to my parents and family for their unconditional love, encouragement and financial support all of which have made this thesis possible. You have sacrificed the world to provide the best life for me.

Abbreviations

β-ME	beta-mercaptoethanol
˚C	Degrees Celsius
AHRI	Africa Health Research Institute
AIDS	Acquired Immunodeficiency Syndrome
ANOVA	Analysis of variance
AP-1	Activator protein 1
AP-4	Activator protein 4
ART	Antiretroviral therapy
ARV	Antiretroviral
BD	Becton Dickinson
BLT	Bone marrow-liver-thymus mice
bp	Base pair(s)
BRD4	BET bromodomain protein 4
c-Myb	c-myeloblastosis
C/EBP	CCAAT/enhancer binding protein
C/EBP U1	C/EBP upstream site 1
C/EBP U2	C/EBP upstream site 2
CA	Viral capsid protein
Ca^{2+}	Calcium ions
CAPRISA	Centre for the AIDS Programme of Research in South Africa
CBF-1	C-promoter binding factor-1
CCR5	C-C chemokine receptor 5
CD28	Cluster of differentiation 28
CD3	Cluster of differentiation 3
CD317	Cluster of differentiation 317
CD4	Cluster of division 4
CDK9	Cyclin dependent kinase 9
CI	Confidence interval
CMV	Cytomegalovirus
COUP-TF	Chicken ovalbumin upstream promoter transcription factor
CPB	CREB-binding protein
CPE	Cytopathic effect
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CXCR4	C-X-C chemokine receptor 4
CycT1	Cyclin T1
DEAE	Diethylaminoethyl
dH_2O	Distilled Water
DMEM	Dulbecco's modified eagle's medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferases
DNMTi	DNA methyltransferase inhibitor
DPBS	Dulbecco’s phosphate-buffered saline

DPP	12-deoxyphorbol 13-phenylacetate
dsDNA	Double stranded deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
EF1α	Elongation factor-1 alpha
eGFP	Enhanced green fluorescent protein
EHMT2/G9a	Euchromatic histone-lysine N-methyltransferase 2
Ets-1	E-twenty-six 1 transcription factor
EZH2	Enhancer of Zeste 2
FACS	Fluorescence-activated cell sorting
FBS	Foetal bovine serum
FDA	Food and Drug Administration
FLL	Firefly luciferase luminescence
FSC-A	Forward scatter area
FSC-H	Forward scatter height
g	gram(s)
G_2 phase	Gap phase 2
GFP	Green fluorescent protein
gp120	120kDa envelope glycoprotein
gp41	41kDa envelope glycoprotein
GR	Glucocorticoid receptor
h	hour(s)
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HDAC-1	Histone deacetylase 1
HDACi	Histone Deacetylase inhibitor
HEK	Human embryonic kidney
HIV	Human immunodeficiency virus
HIV-1	Human immunodeficiency virus type 1
HIV-2	Human immunodeficiency virus type 2
HMBA	Hexamethylene bisacetamide
HMT	Histone methyltransferases
HMTi	Histone methyltransferase inhibitor
HSC	Hematopoietic stem cell
IL-2	Interleukin 2
IN	Viral integrase protein
Iono	Ionomycin
Kb	kilobase pair(s)
kDa	Kilodaltons
l	litre(s)
LB	Luria-Bertani
LEF-1	Lymphoid enhancer-binding factor 1
LSF	Late SV40 factor
LTR	Long terminal repeat
m	Milli
MA	Viral matrix protein
MAPK	Mitogen activated protein kinase
Mg^{2+}	Magnesium ions
$MgCl_2$	Magnesium chloride
MHCI	Major histocompatibility class I
min	minute(s)

mRNA	Messenger ribonucleic acid
mTOR	Mammalian target of rapamycin
n	Nano
N-TEF	Negative transcription elongation factor
NC	Viral nucleocapsid protein
Nef	Viral negative factor
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NFAT	Nuclear factor of activated T cells
NIH	National Institute of Health
NRE	Negative regulatory element of the LTR
nt	Nucleotide(s)
P-TEFb	Positive transcription elongation factor b
PBMC	Peripheral blood mononuclear cell
PBS	Primer binding site
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PIC	Pre-initiation complex
PID	Participant Identification Number
PKC	Protein Kinase C
PMA	Phorbol 12-myristate 13-acetate
PPT	Polypurine tract
PR	Viral protease protein
PrEP	Pre-exposure prophylaxis
R	Repeat region of the LTR
RBF-2	Ras response element binding factor 2
RF	Restriction-free
RGH	Red-Green-HIV-1
RLL	*Renilla* luciferase luminescence
RLU	Relative light units
RNA	Ribonucleic acid
RNAPII	RNA polymerase II
RRE	Rev response element
RT	Reverse transcriptase
s	second(s)
SAHA	Vorinostat
SFFV	Spleen focus forming virus
siRNA	Small interfering ribonucleic acid
SIV	Simian immunodeficiency virus
snRNP	Small nuclear ribonucleoprotein
SOC	Super Optimal broth with Catabolite repression
Sp1	Specificity protein 1
SSC-A	Side scatter area
ssRNA	Single stranded ribonucleic acid
Suv39H1	Suppressor of variegation 3–9 homologue 1
TAR	Trans-activation response loop/element
Tat	Viral trans-activator of transcription protein
TBE	Tris-Borate-EDTA
TBP	TATA binding protein
$TCID_{50}$	50% tissue culture infectious dose

TDF	Tenofovir disoproxil fumarate
TFII-I	General Transcription Factor II-I
TFIID	Transcription factor II D
TFIIH	Transcription factor II Human
TNF-α	Tumour Necrosis Factor alpha
ToM	T cell only mice
tRNA	Transfer ribonucleic acid
TsA	Trichostatin
U	unit(s)
U3	Unique 3' region of the LTR
U5	Unique 5' region of the LTR
UCT	University of Cape Town
USF-1	Upstream stimulatory factor 1
UV	Ultra-violet
v/v	volume/volume
Vif	Viral infectivity factor
VISCONTI	Virological and Immunological Studies in Controllers after Treatment Interruption
VPA	Valproic acid
Vpr	Viral protein R
Vpu	Viral protein U
VSV	Vesticular stomatis virus
w/v	weight/volume
WT	Wild-type
x g	times gravity (centrifugation speed)
YY1	Yin Yang 1 factor
µ	Micro

List of Tables

List of Figures

Abstract

The persistence of latent viral reservoirs, that are insensitive to antiretroviral therapy (ART), remains the greatest barrier to HIV-1 eradication. The role that viral factors play in HIV-1 latency establishment and maintenance is poorly understood, and characterisation of these factors is imperative for the development of curative strategies or interventions that could lead to HIV-1 remission in infected individuals. Subtype level genotypic variation within regulatory elements of the HIV-1 promoter, the long terminal repeat (LTR), has been shown to influence latency establishment in *in vitro* models. We investigated the influence of inter-participant subtype C LTR genotypic variation on the establishment of latency in a dual reporter HIV-1 plasmid model and evaluated potential correlates of this latency potential.

Long terminal repeats from 11 ART-naïve, acutely subtype-C infected women in the CAPRISA 004 cohort from Durban, South Africa were cloned into an HIV-1-expressing vector (pRGH) used to generate pseudovirions following HEK293T transfection. Pseudoviruses harboured a *gag*-eGFP gene under the control of the participant LTR, allowing measurement of active replication, and an mCherry gene under the control of a constitutive CMV promoter allowing measurement of viral integration. Latency potential was expressed as the ratio of mCherry only (latent) to eGFP and mCherry (active replication), as measured by flow cytometry after infection of Jurkat E6-1 and CEM.NK^{R} CCR5+ cell lines before and after T cell activation with PMA/Ionomycin. A panel of LTRs cloned into a pGL4.10 luciferase expression vector were used to measure basal LTR expression and Tat-induced LTR expression.

All LTR sequences were classified as subtype C, with an average inter-participant pairwise DNA distance of 7.6%. The median basal LTR activity was approximately two times higher than that of the BaL isolate (interquartile range: 1.38-2.14), and Tat-induced activity approximately nine times higher than that of BaL (interquartile range: 6.16-10.33). We observed consistently greater proportions of latently infected cells than actively infected cells. In Jurkat E6-1 cells, the median latent:active infection ratio was 1.97 (range 0.86-2.83; three experiments). Latency was reversible in a proportion of cells as the median latent:active infection ratio decreased to 0.55 (range 0.46-0.78).

The latent:active ratio was unchanged, post-stimulation, in CEM.NK^{R} CCR5+ cells and was therefore found not to be a suitable cell-line for the model. Latency potential did not correlate with basal or Tat-induced activity (Spearman correlation tests, basal p=0.25, r=-0.38, Tat-induced p=0.42, r=-0.27). The DNA distance in characterised functional sites from consensus did not correlate with latency potential (Spearman correlation test p=0.67, r=0.14).

Our data suggest that HIV-1 LTRs have intrinsic properties which influence latency potential and the proportion of latently infected cells early post-infection. However, since differences were independent of basal and Tat-induced LTR activity, other factors such as regulatory element interaction and the efficiency of recruitment of molecules responsible for establishing latency, such as histone modifiers, may play a role.

Chapter 1: Literature Review

1.1 Introduction

The Human Immunodeficiency Virus type 1 (HIV-1) is the pathogen responsible for Acquired Immunodeficiency Syndrome (AIDS)(1, 2). To date, HIV-1/AIDS has been responsible for up to 35 million deaths worldwide. In 2016, approximately 36.7 million people were living with HIV-1, 1.8 million new infections were recorded, and 1 million people died from AIDS-related illnesses. South Africa remains one of the countries with the highest HIV-1 prevalence with 7.1 million people (19% of global infections) living with HIV-1 at the end of 2016(3).

In many areas of the world, the number of new HIV-1 infections are declining as a result of increasing coverage of antiretroviral therapy (ART), as patients with suppressed viral loads are less likely to transmit the virus(4). However in 2016, only 56% of infected individuals in South Africa had access to treatment(3). Antiretrovirals (ARVs) are also used in pre-exposure prophylaxis (PrEP), however, there is as yet no evidence that this is contributing to controlling the epidemic. With ARVs used for both treatment and prevention, the emergence of drug resistance, which would render ART ineffective, is a concern(5). While ART can suppress viral loads to below detectable levels and prolong life, these drugs are unable to eradicate the virus from the body due to persistence of latent viral reservoirs and anatomical sanctuaries where drugs are unable to penetrate(6, 7). Furthermore, sustaining life-long treatment of large populations of infected individuals is challenging and economically unviable(8). Therefore, urgent efforts are needed to either improve the current treatment strategies or work towards development of an HIV-1 cure or state of remission (control in absence of therapy).

HIV-1 latency allows the virus to remain within infected cells for many years unaffected by treatment and undetected by host immune responses. However, when ART is discontinued, viral replication is reinitiated, allowing the rapid re-emergence of viremia(9, 10). Most strategies aimed at eliminating the latent reservoir are geared towards reactivating latent viruses while patients are still on ART(11). Unfortunately, the factors that govern HIV-1 latency are still poorly understood, particularly the

contribution of the viral factors such as genotypic variability, which is the focus of this thesis.

This literature review will cover aspects of the structure of the HIV-1 genome focussing on the organisation of the HIV-1 promoter element, the long terminal repeat (LTR), and cellular transcription factors that play a role in regulation of its function. Furthermore, the molecular mechanisms of viral latency establishment and maintenance will be described with particular attention given to those mechanisms that act directly on the LTR. Finally, the current models available for the study of viral latency and strategies to reverse latency will be described.

1.2 Hope for an HIV-1 Cure

A number of recent case studies(12–16) providing evidence of the possibility of a cure have driven a strong interest in HIV-1 cure research. Most noteworthy is the report of Timothy Ray Brown, the “Berlin patient” who is the only individual reported to be cured of HIV-1 infection(12). The Berlin patient was on ART when he developed myeloid leukaemia. After chemotherapy and irradiation therapy, he received a hematopoietic stem cell (HSC) transplant from a donor who was homozygous for a 32 base-pair deletion in the CCR5 gene, a gene encoding an essential co-receptor for HIV-1 infection. These cells are therefore resistant to HIV-1 infection. This patient has been off therapy for approximately nine years and has shown no evidence of viral rebound. This is an example of a sterilizing cure where the virus was completely eradicated from the infected individual.

In two similar cases, the “Boston patients”, the HSC transplant was likewise attempted, however the stem cell donors were heterozygous for the 32 base-pair deletion in the CCR5 gene and the patients remained on ART post-transplant in the hopes of preventing infection of donor cells(13). Disappointingly, only a delay in viral rebound was observed in both patients after ART was discontinued. It is possible that a subset of remaining latently infected cells were reactivated and were ultimately responsible for viral rebound.

Further evidence of the possibility of a cure is provided by the cases of the Mississippi baby and the Virological and Immunological Studies in Controllers after Treatment Interruption (VISCONTI) cohort where ART was initiated in infected individuals very early post-infection and was interrupted after 18 months and an average of three years respectively(14, 15). The Mississippi baby experienced remission but eventually succumbed to viral rebound. The majority of the men in the VISCONTI cohort did not experience viral rebound for at least ten years off therapy.

In addition, two other cases have been reported where a French teenage girl and South African child have experienced remission for 12 and eight years respectively(17, 18). Both individuals were infected at birth and treated within three months after birth and treated for six years and 40 weeks respectively.

These incidences demonstrate the possibility of a functional cure where an infected individual may live disease-free without treatment, while still harbouring the virus. Additionally, they highlight the need to understand the dynamics of the inducible latent reservoir and to develop a strategy to purge it.

1.3 HIV classification

HIV is a lentivirus that is part of the larger retrovirus group. It can be classified into two types/lineages, namely HIV type 1 and 2 (HIV-1 and HIV-2) which are a result of separate zoonotic transmission events(19). HIV-1 is the dominant type, accounts for the vast majority of global infections, and is responsible for the AIDS pandemic. A small percentage of HIV-2 infections occur chiefly in West African regions, however this type is less pathogenic(20).

HIV-1 is further classified based on origin and genotype. Firstly, HIV-1 is comprised of four groups that are assigned as follows: group M (major), which is responsible for over 95% of infections, group O (outlier), group N (non-M/non-O) and the most recently discovered group P(21–23). HIV-1 group M has diversified over time into 9 discrete subtypes: A, B, C, D, F, G, H, J and K. Subtype C, the dominant virus in South Africa, is reported to be responsible for the majority of infections globally, with subtypes A, B,

G and D contributing significantly (Figure 1.1)(24). In addition, HIV-1 group M has further diversified into 96 circulating recombinant forms (CRFs) (25). These CRFs are formed when recombination occurs between subtypes within an individual infected with two or more subtypes. These recombinants are subsequently passed on to other individuals(26). The recombinant forms AG and AE collectively contribute to approximately 13% of global infections.

The HIV-1 M subtypes are differentially represented in terms of geographical location. Of importance, the dominant subtype C viruses are responsible for the majority of infections in the greater part of Southern Africa, India and Ethiopia. Subtype A is overrepresented in areas of Central Asia, Eastern Europe and Central and West Africa. In North and Latin America, Western Europe, and East Asia infections are predominantly caused by Subtype B viruses. Subtype D is mainly isolated in patients in Northern, Western and East Africa. Most Subtype G infections are located in certain regions of Western Africa. There is no evidence for geographical overrepresentation of subtypes, F, H, J and K(24).

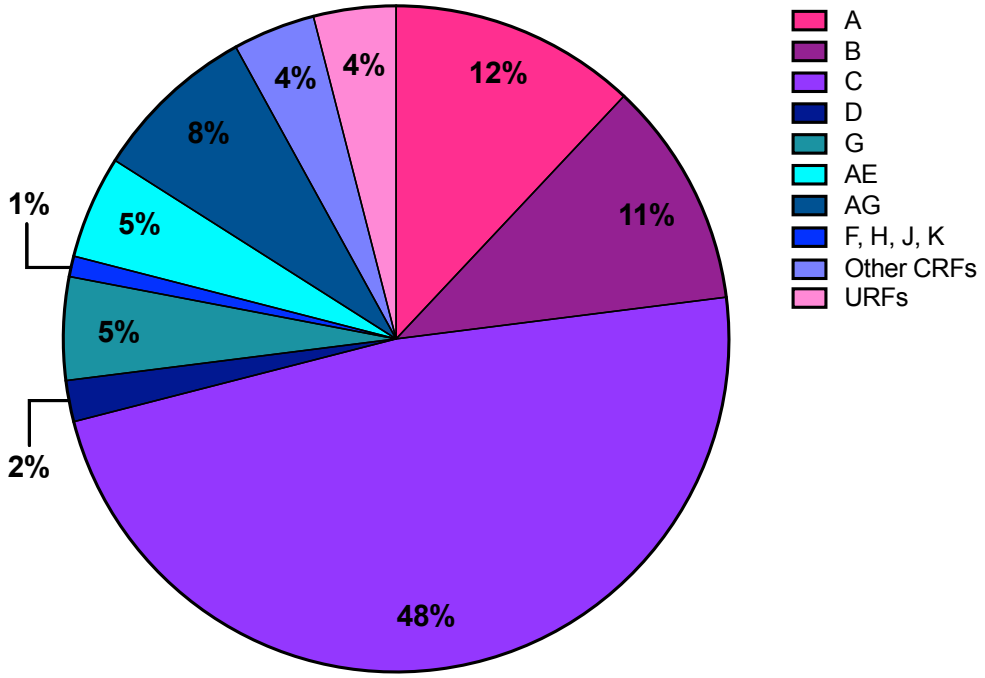


Figure 1.1 Global contribution of HIV-1 group M subtypes. The percentage of total infections between 2004 and 2007 caused by each HIV-1 group M subtype and recombinant(24).

1.3.1 HIV-1 subtype pathogenicity

In addition to differences in genotype and geographic distribution, the different HIV-1 group M subtypes display different degrees of pathogenicity. Of importance, Amornkul *et al.*(27) reported that subtype C infected individuals experience a faster decline in CD4+ T cells and therefore more rapid disease progression to AIDS compared to individuals infected with subtype A and D. A more rapid increase in viral load was also observed compared to subtype A infected individuals. Conversely, at least two studies reported that subtype C viruses were less fit, that is having reduced ability to adapt and replicate in a given environment, in comparison to other subtypes(28, 29). Furthermore, some studies have demonstrated subtype D as being the most virulent(30–32). However, the fact that subtype C viruses are globally dominant highlights the necessity to further characterise these viruses. Overall these studies show that genotypic features of HIV-1 subtypes contribute to phenotypic characteristics of the virus that determine disease progression and clinical outcomes of infection.

1.4 The HIV-1 genome

The entire HIV-1 genome comprises of approximately 9700 nucleotides. It is partitioned into nine viral genes (*gag*, *pol*, *env*, *tat*, *rev*, *vif*, *vpr*, *vpu* and *nef*) and is flanked by two identical LTRs (reviewed in detail in section 1.6). To minimise on genome length, the virus makes use of all three reading frames to encode the various protein sequences (Figure 1.2). The nine viral genes are divided into three classes: structural gens (*gag*; *pol* and *env*); regulatory gens (*rev* and *tat*) and accessory genes (*vif*, *vpr*, *vpu* and *nef*). Short descriptions of each of the nine proteins are provided below.

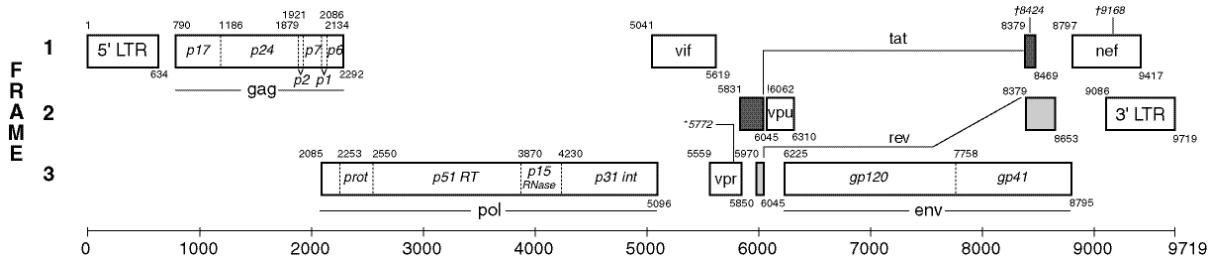


Figure 1.2 Gene map of the HXB2 HIV-1 strain. The three open reading frames are depicted. Positions of each of the nine genes are shown and gene subunits are indicated. Nucleotide start and end positions for genes are indicated at the top left and bottom right of each gene respectively. Figure obtained at: https://www.hiv.lanl.gov/content/sequence/HIV/MAP/landmark.html.

1.4.1 Gag

The *gag* gene encodes the major structural proteins of the viral core. There are four separate proteins encoded by *gag*, including the inner core membrane or matrix (MA, p17), the capsid protein (CA, p24), the nucleocapsid protein (NC, p7), which forms the nucleoprotein/RNA complex, and p6, which is involved in particle release(33). p7 and p6 are separated by p1, which is responsible for frameshifting into *pol* during translation(34). p2 is a spacer peptide that has been implicated in correct protein folding(33, 35).

1.4.2 Pol

The viral catalytic enzymes are encoded by *pol*. Protease (PR, p12) is responsible for proteolytic cleavage of precursor proteins; reverse transcriptase (RT, p51) is responsible for conversion of the viral RNA to DNA; RNase H (p15) for degradation of the viral RNA during reverse transcription, and Integrase (IN, p32), which facilitates insertion of viral nucleic acid into the host cell genome(36).

1.4.3 Env

Env consists of two glycoproteins that make up the viral envelope; gp120 (SU, surface glycoprotein), which contains the binding sites for the CD4 and CCR5/CXCR4 cell receptors, and gp41 (TM, transmembrane glycoprotein), which is responsible for viral/cell membrane fusion. These proteins associate in a covalently-bound trimeric form on the viral surface(37).

1.4.4 Tat

The viral trans-activator of transcription protein, Tat, is the most essential viral protein for efficient HIV-1 gene expression. Tat is encoded on two separate exons in the viral genome and is one of the first genes expressed during infection. Tat acts by binding to the trans-activation response loop/element (TAR) located within the LTR and powerfully enhances viral expression by preventing termination of transcription(38–40).

1.4.5 Rev

Rev is a highly conserved regulatory factor which is responsible for HIV-1 partially spliced and unspliced mRNA export and stabilisation from the nucleus to the cytoplasm. Rev binds the mRNA via the Rev response element (RRE)(41, 42).

1.4.6 Vif

The viral infectivity factor (Vif) protein enhances infectivity but does not directly influence replication. Vif binds the human cytidine deaminases APOBEC3G/F inhibiting their incorporation into the viral particle. In the absence of Vif, these host factors deaminate the (-)strand of viral transcripts resulting in incorporation of multiple G to A mutations. These mutations lead to the production of non-functional mRNA(43).

1.4.7 Vpr

Viral protein R (Vpr) is another viral accessory factor incorporated into the viral particle. Vpr acts in the nucleus of infected cells and prevents cell division by arresting the cell in the G_2 phase of the cell cycle in which the HIV-1 LTR was found to be most active. This protein is also thought to affect the early stages of the viral life cycle i.e. entry, uncoating, reverse transcription and nuclear transport(44, 45).

1.4.8 Vpu

Viral protein U (Vpu) is an accessory integral membrane protein involved in antagonising the action of human cell CD317 tetherin, a host restriction factor that tethers viral particles to the cell membrane. Vpu is therefore essential for the release of newly produced viral particles out of infected cells(46). Furthermore, it has been shown that Vpu is responsible for targeted degradation of newly synthesised CD4 molecules at the endoplasmic reticulum of infected cells(47).

1.4.9 Nef

The negative factor (Nef) protein is involved in multiple functions of HIV-1 pathogenesis and is required for successful spread of viral infection, and also contributes to disease progression. Specifically, Nef contributes to downregulation and degradation of CD4 and MHCI complexes on the cell surface membrane preventing detection of infected cells by the host immune system(48). Additionally, Nef interacts with a number of signalling molecules within an infected cell that are collectively capable of inducing T cell activation in a CD4 independent manner, thus producing a cellular environment optimal for viral production. Nef has also been implicated in contribution to viral spreading by upregulation of receptors on dendritic cells which capture viral particles and facilitate transfer to uninfected CD4 cells(49, 50).

1.5 The HIV-1 life cycle

Each HIV-1 particle is enveloped with two copies of single-stranded RNA (ssRNA) and a number of viral enzymes including PR, RT and IN, amongst other viral proteins such as NC, all encapsulated within the core region of the virus(51). HIV-1 virions primarily target CD4+ T cells for infection(52, 53). The virions gain access to the cells by initial attachment of the viral gp120 Env protein to the CD4 cell surface receptor(54). Attachment of gp120 induces a change in conformation allowing further binding of the virus to either the CCR5 or CXCR4 chemokine receptors on the cells(55, 56). A secondary viral Env protein, gp41, facilitates fusion of the cell membrane and viral

envelope, which in turn allows the entry of the capsid(57). As the capsid enters CD4+ T cells, it undergoes uncoating to release its contents.

The (+)strand ssRNA is converted into double-stranded DNA (dsDNA) by the packaged HIV-1 enzyme, RT. Reverse transcription takes place in the cytoplasm within the reverse transcription complex (RTC), a multiprotein complex which protects the viral genome and proteins from host anti-viral responses(58). Reverse transcriptase is error prone and therefore incorporates mutations during this process(59). Upon completion of viral DNA synthesis, the RTC matures into the preintegration complex (PIC) via protein re-arrangement(58). The PIC is then actively imported into the nucleus where viral DNA is integrated into the host chromosome by the viral enzyme, IN (60, 61).

Once integrated, the HIV-1 DNA (genome) is referred to as a provirus and is able to replicate and initiate transcription from the 5' LTR to produce viral transcripts. Initial transcription relies on the use of the host T cell machinery, however in the absence of Tat, transcription is hindered and elongation of transcripts is inefficient(40). The viral Tat protein is expressed off of initial transcripts and binds to the TAR loop of the LTR, potently upregulates the synthesis of full length viral transcripts stabilising elongation by recruitment of other cellular factors(62, 63). Hence a positive feedback loop is established.

Subsequently, viral mRNA (sliced and unspliced) is exported from the nucleus to the cytoplasm. Here the spliced mRNA is translated into the polyproteins and processed into the proteins necessary for production of a new viral particle by the viral PR. These proteins together with unspliced viral RNA are transported to the cell plasma membrane and assemble into an immature viral particle with a partial envelope where the matrix and capsid structures are not completely cleaved. This immature particle pushes its way out of the CD4 T cell in a process termed budding, where envelope assembly is completed. The viral protein PR is then activated to complete processing of remaining polyproteins and rearrangement of the Matrix and Capsid structural proteins takes place thereby producing a mature infectious HIV-1 particle(51). This pathway of the HIV-1 life cycle is termed the lytic cycle.

Alternatively, HIV-1 may enter the lysogenic cycle upon infection of a CD4+ T cell. Integration takes place as described as above, however the provirus simply lies transcriptionally silent, or latent, within the host cell and no new viral particles are produced despite the provirus being capable of doing so(64).

The establishment of viral latency requires the suppression of HIV-1 transcription(65). As HIV-1 transcription is dependent on the LTR, and LTR activity varies based on its genotype, it is possible that activity of the LTR can have a direct effect on the establishment of latency of a virus(66). Latency will be discussed in detail in section 1.7.

1.6 The HIV-1 Long Terminal Repeat (LTR)

1.6.1 Reproduction, Structure & Function

The integrated proviral LTR consists of three discreet segments, the unique 3' (U3), repeat (R) and unique 5' (U5) regions(67, 68). When in RNA form, the HIV-1 genome is flanked by two partial LTR sequences, an R and U5 region on the 5' end and a U3 and R region on the 3' end (Figure 1.3A). During the process of reverse transcription, prior to integration, two identical complete LTRs are generated on either end of the resulting dsDNA sequence. The steps of this process are illustrated in Figure 1.3 A-H.

This complex process was described by Huber *et al*. and Esposito *et al*. (69, 70). Briefly, HIV-1 replication is initiated by tRNA hybridisation to the primer binding site (PBS) on the (+)strand RNA at the end of the 5' U5 region. The (-)strand is extended from the bound tRNA in the 3' direction to the end of the 5' R region by the RT-associated DNA polymerase(Figure 1.3A). The R and U5 regions of the (+)strand are subsequently degraded by viral RNAse H (Figure 1.3B). The newly synthesised (-)strand fragment is therefore free to move to and hybridise with the complimentary R region at the 3'end of the RNA (+)stand in a process known as the first strand transfer. This hybridised fragment behaves as a primer and synthesis of the (-)strand by DNA polymerase is carried out as far as the PBS at the 3' end of the RNA (Figure 1.3C).

Following synthesis, the entire template RNA strand is RNAse H digested with the exception of two purine rich regions upstream of the U3 region, known as polypurine tracts (PPTs), which allows segments of the (+)strand to remain bound to the (-)strand (Figure 1.3D). DNA synthesis takes place using these (+)strand segments as a primer and the fragment is extended in the 3' direction thus producing two fragments, one near the centre region of the HIV-1 genome and the other including the U3, R and U5 (including the PBS) regions (Figure 1.3E). RNAse H degrades the RNA portion of these RNA:DNA hybrids releasing the terminal (+)strand fragment (Figure 1.3F) and a second strand transfer takes place. The newly synthesised fragment hybridises with the PBS on the 3' end of the (-)strand. DNA synthesis of the (-)strand is completed towards the 3' end of the RNA using this (+)strand fragment as a template. Finally, the (+)strand is extended to the 3' end of the genome making use of the hybridised fragment and the central segment as primers(Figure 1.3G).

Ultimately, the product of these processes is an integration competent, dsDNA molecule encoding the entire HIV-1 genome, flanked by two identical 634bp LTR sequences each comprising of a U3, R and U5 region (Figure 1.3H). Of importance, due to the strand transfers, genetic mutations within the 3'LTR are therefore ultimately introduced into the 5'LTR sequence.

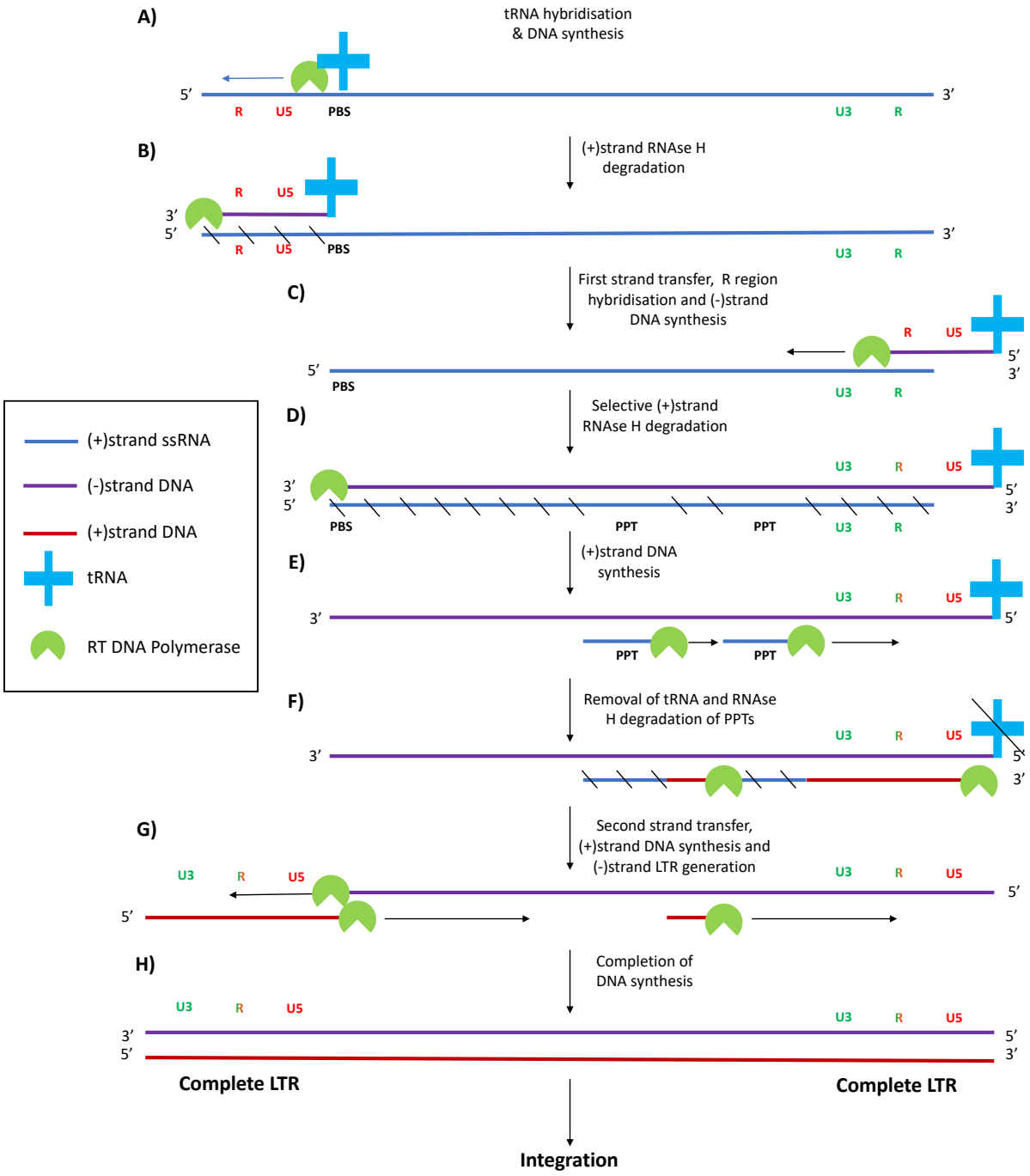


Figure 1.3 Schematic representation of LTR regeneration during viral reverse transcription.

Once the complete HIV-1 proviral genome has integrated into the host cell genome, the two LTRs, despite having identical nucleotide sequences, perform different functions. The 5'LTR acts as the HIV-1 promoter and harbours the transcriptional start site during viral replication whereas, the 3'LTR encodes a large portion of the *nef* gene and contains the necessary nucleotide sequence responsible for termination of

transcription and post transcriptional modification such as polyadenylation(67, 71). We will focus here on the function of the 5'LTR as a promoter and its role in HIV-1 latency.

The 5'LTR sequence provides a binding/interaction region for a plethora of transcription factors and other proteins that influence transcription initiation and maintenance and promoter activity. The TAR element is contained within the R region of the LTR. This is the region where transcription is initiated and where the viral Tat protein binds in concert with the positive transcription elongation factor b (P-TEFb) complex composed of Cyclin T1 (CycT1) and cyclin dependent kinase 9 (CDK9)(72) (section 1.7.1.4) (Figure 1.4).

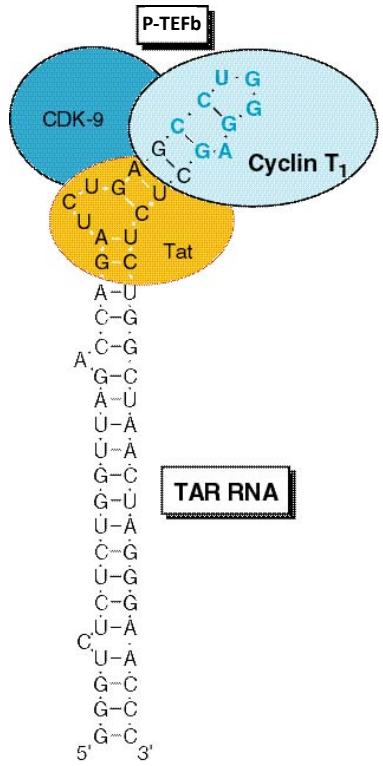


Figure 1.4 Interaction of Tat and the P-TEFb complex with the HIV-1 TAR loop. Tat binds to a U-C-U bulge region and the Cyclin T1 subunit of the P-TEFb complex binds to top of the loop. Figure adapted from Karn, 2000(40)

The complete LTR is 634bp in length and is assigned nucleotide (nt) positions relative to the transcriptional start site hence, the TAR element spans from nt +1 to +60(73). Three other functional regions have been designated within the U3 region, namely the modulatory (nt -454 to -104), enhancer (-105 to -79) and core promoter (-78 to -1) regions (Figure 1.4). These regions are classified based on their combinatory effect

on promoter activity and viral genome expression(67, 74, 75). A fragment of the modulatory region was identified, by deletion and mutational analysis, to reduce expression of the LTR and was therefore designated as the negative regulatory element (NRE) extending between nt -340 and -184(76). These elements of the LTR are discussed in greater detail in the next section.

1.6.2 Organisation of transcription factor binding/functional sites

For the most part, LTR functional sites are conserved across HIV-1 group M subtypes. However, minor differences between subtypes exist which can impact the replicative capacity and fitness of each virus(77, 78). Most studies to date have focused on characterising the functional binding sites of subtype B LTR sequences (75, 79). The functions and positions of *cis*-regulatory binding sites within the HIV-1 LTR are illustrated below (Figure 1.5). An overview of each well-characterised site and its function will be reviewed below, and subtype C-specific features will be highlighted.

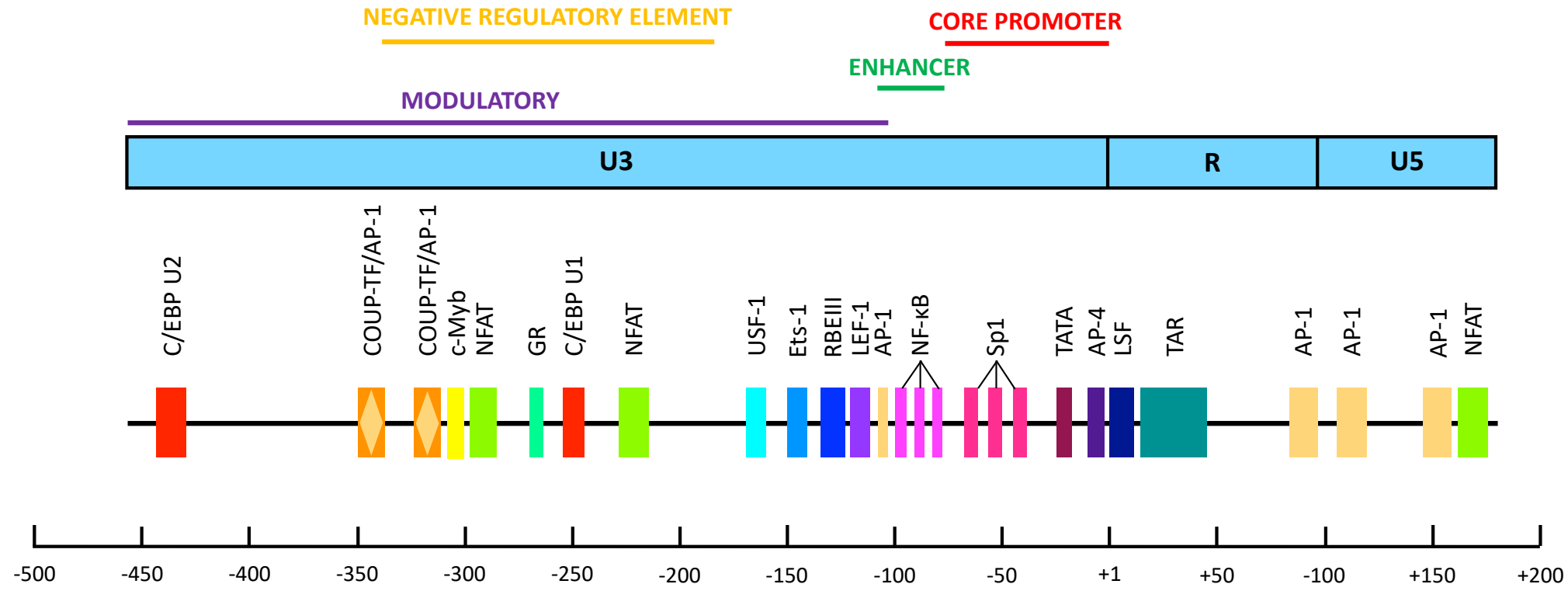


Figure 1.5 Schematic representation of approximate putative binding sites of well characterised *cis*-regulatory/transcription factor binding sites in the 5'LTR of HIV-1. The U3, R and U5 regions and regions of functional sites are indicated by coloured boxes. The transcriptional start site is indicated at nucleotide +1 at the U3 and R region junction.

1.6.2.1 The core promoter element

The core promoter element comprises of the initiator element, which overlaps with the start of the TAR region, and the canonical TATA box promoter element to which the TATA binding protein (TBP)/transcription factor II D (TFIID) binds (Figure 1.5), initiating formation of the viral pre-initiation complex (PIC)(80). Subunits of the PIC bind directly to the core promoter region upstream of the TAR element. Three tandem specificity protein 1 (Sp1) sites also lie within the core promoter region upstream of the TATA box(81). Binding of the Sp1 protein to these sites has a positive effect on viral transcription, via recruitment of histone deacetylases (HDACs), and has been found to be essential for both basal and Tat-induced viral gene expression(82–84). An activator protein 4 (AP-4) binding site was also found within the core promoter, directly downstream of the of the TATA box(85). Interestingly, Imai *et al*.(86) demonstrated that binding of AP-4 to this site represses HIV-1 expression by competitively blocking TFIID binding to the TATA box and in addition, recruiting histone deacetylase 1 (HDAC-1).

1.6.2.2 The enhancer element

Immediately 5' of the core promoter, is the LTR enhancer region (Figure 1.5). This region primarily comprises of one, two or three adjacent, conserved nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) protein binding sites depending on the HIV-1 group M subtype. Of note, the majority of HIV-1 subtype C LTRs contain three NF-κB binding sites and have been found to have the highest promoter activity compared to other subtypes(87, 88). In addition, the acquisition of a fourth NF-κB site within the enhancer region has been reported in viruses from Southern Indian infected patients and was found to confer higher transcriptional activity within subtype C viruses(89).

NF-κB binding sites are the most well characterised of the HIV-1 LTR elements, and the binding of NF-κB is considered one of the most essential steps in HIV-1 expression(74, 88, 90). NF-κB, in its active form, binds to the LTR as a p65/p50 heterodimer protein complex(91). Protein interactions between the p65/p50

heterodimer and molecules of Sp1 form as NF-κB binds to the LTR. These coordinated interactions are essential for enhancer functionality(92). Importantly, the binding of NF-κB is involved in recruitment of elements of RNA Polymerase II (RNAPII) complex to the transcription initiation site(93). Specifically, NF-κB was shown to recruit Transcription factor II Human (TFIIH), a transcription factor that phosphorylates RNAPII leading to promoter clearance(94). Furthermore, NF-κB contributes to chromatin remodelling where the p65 subunit of the heterodimer recruits the histone acetyltransferases (HATs), CREB-binding protein (CPB) and p300. HAT-driven acetylation of specific histones results in the enhancement of viral gene expression(91, 95, 96). Sequestering of CBP and p300 by NF-κB is also important in Tat-driven HIV-1 transcription. These HATs directly acetylate Tat in a process that has been shown to be important in Tat trans-activation of the LTR(97).

An additional cellular transcription factor, nuclear factor of activated T cells (NFAT), is capable of binding to the NF-κB binding site within the LTR enhancer element as the two transcription factor recognition sequences overlap(98). NFAT is also capable of recruiting CBP and p300 to induce transcriptional activation and enhance HIV-1 gene expression(99, 100). Due to the overlapping recognition sites, binding of these two factors cannot take place simultaneously.

Finally, an activator protein 1 (AP-1) binding site has been reported upstream of the NF-κB site within the enhancer element. Conservation of this binding site has been shown to be important in the ability of the virus to establish latent infection(101).

1.6.2.3 The modulatory region

The modulatory region lies at the 5’ end of the U3 region in the LTR (Figure 1.5) and contains recognition sequences for a wide range of regulatory factors. Binding of transcription factors to some of these sites results in positive regulation of transcription whereas others have an inhibitory effect. Most positive regulatory sites within the modulatory region are required for HIV-1 transcription but are not as essential as those mentioned in sections 1.6.2.1 and 1.6.2.2(79).

In subtype B LTRs, three CCAAT/enhancer binding protein (C/EBP) sites lie within the modulatory region, one at the extreme 3' end of the modulatory region and the other two further upstream(102). However, in subtype C viruses, the extreme 3' site does not exist, but instead, has been replaced with the extra NF-κB site described in section 1.6.2.2, and the other two C/EBP sites lie further upstream in the LTR. These sites have been assigned as C/EBP upstream site 1 and 2 (C/EBP U1 and C/EBP U2) and lie between nt -259 to -247 and -446 to -433 relative to the transcriptional start site respectively(103, 104). In subtype B viruses, C/EBP has been shown to interact with the p65 subunit of NF-κB and contributes to both basal and Tat-induced transcriptional enhancement in macrophages and monocytes but this has not been demonstrated in CD4+ T cells(105–108). The specific role of the C/EBP sites within subtype C viruses has not been well characterised.

A lymphoid enhancer-binding factor 1 (LEF-1) binding site has been described between nt -142 and -122 in the modulatory region(75, 109). Binding of the LEF-1 protein to this site is thought to induce disruption of the chromatin nucleosome and interacts with HATs recruited by NF-κB. Furthermore, LEF-1 promotes bending of DNA and interacts with Sp1 and also facilitates interactions and correct positioning between other regulatory factors(109–112). Therefore LEF-1 has been classified as a positive regulator of HIV-1 expression.

Continuing upstream, a binding site for the E-twenty six 1 (Ets-1) transcription factor lies between nt -149 and -142 in the LTR(109, 113). Additionally, a binding site for the upstream stimulatory factor 1 (USF-1) binds to an E-Box consensus sequence found between nt -166 and -161. It has been demonstrated, that interaction between these two proteins is necessary for them to bind to the LTR, and that their cooperative binding enhances viral gene expression. Ets-1 and USF-1 also contribute to bending of the LTR DNA, with Ets-1 interacting with LEF-1, to promote interactions of proteins within the PIC, such as Sp1 to aid in recruiting basal transcriptional machinery(109, 111, 114, 115). Interestingly, it was more recently demonstrated that overexpression of an alternatively spliced form of the Ets-1 protein was capable of inducing HIV-1 transcription from latently infected resting CD4+ T cells from patients on therapy independently, without causing T cell activation(116). This demonstrates the importance of the Ets-1 factor in positive regulation of the LTR.

Interestingly, the USF-1 protein has also been shown to form a heterodimer with USF-2 and is collectively known as Ras response element binding factor 2 (RBF-2)(117). This heterodimer binds to a region of the LTR termed RBEIII which overlaps with the LEF-1 recognition site and is found between nt -131 to -121(118). RBF-2 binds to the E-box described above more efficiently than RBEIII, however a co-factor General Transcription Factor II-I (TFII-I), initially identified as an initiator binding protein, is essential for promoting this interaction. Chen *et al*.(117) demonstrated that binding of RBF-2 to RBEIII induces latent integrated viral expression in a T cell activation dependent manner, specifically in response to activation by the Ras/mitogen activated protein kinase (MAPK) cascade(119, 120). No published reports have detailed any interaction between LEF-1 and RBF-2.

Two more NFAT DNA recognition sites were found in the modulatory region of the U3, one between nt -254 and -216 and another between nt -303 and -288, via DNA foot-printing experiments. However, NFAT associated more weakly to these upstream sites compared to those in the enhancer region(121). The effect of NFAT binding to these upstream sites has not been well characterised as increased levels of and higher affinity binding of NFAT has been demonstrated in activated T cells(122). However, there is evidence to show that HIV-1 transcription is not significantly affected by NFAT binding(123).

A glucocorticoid receptor (GR) binding site was originally identified, via mobility shift assays, between the two NFAT recognition sites between nt -264 and -269(124). It was subsequently demonstrated that addition of the glucocorticoid, dexamethasone, had a negative regulatory effect on viral expression in T and B lymphocytic lineages. Furthermore, this suppression was dependent on the LTR having an intact GR(125). Despite this discovery, the specific interactions of this region and the PIC are yet to be elucidated.

A binding site for the proto-oncogene encoded c-Myb (c-myeloblastosis) protein lies between nt -304 and -299. Binding of this protein was found to stimulate viral expression via trans-activation of the LTR(126, 127). Interestingly, the c-Myb protein is upregulated during mitogen dependent T cell activation, the state in which viral expression is at its highest, suggesting that this protein may form a direct link between

the cell activation and degree of viral expression(128). The c-Myb consensus binding sequence has additionally been implicated in the formation of other protein complexes associated with transcription(129).

Finally, two putative binding sites for both AP-1 and chicken ovalbumin upstream promoter transcription factor (COUP-TF) are found at nt -352 to -324 and nt -306 and -285. However, studies have demonstrated that the AP-1 protein is incapable of direct interaction with these two sites, and the effect of AP-1 is unknown here(130, 131). The COUP-TF protein reportedly binds directly to the nt -352 to-324 region and induces transcriptional activity. However, COUP-TF additionally is capable of positively or negatively mediating viral expression, in a cell type-specific manner, via interaction with Sp1 in the core promoter element(132, 133).

1.6.2.4 Downstream elements in the R and U5 region

A number of *cis*-regulatory binding sites are also present downstream of the transcriptional start site and cellular factors binding here interact with the transcription initiation complex and regulate its activity.

Firstly, a late SV40 factor (LSF) binding site overlaps the transcriptional start site spanning from nt -10 to +27. Binding of LSF is thought to play an important negative regulatory role in HIV-1 expression as it recruits a secondary factor, Yin Yang 1 (YY1), which in turn recruits HDAC-1 which has been shown to contribute to Tat activation inhibition(134–136).

Further downstream, an additional three AP-1 binding sites were identified at nt +85 to +93 (site I), nt +121 to +138 (site II), and nt +153 to +161 (site III) via footprinting analysis(137, 138). Mutation of these sites, especially site I and II, was shown to decrease HIV-1 basal expression and viral infectivity(139, 140).

Finally, another NFAT binding site was discovered in the U5 region, between nt +158 and +175(137). In addition to recruitment of stimulatory co-factors (section 1.6.2.2),

prevention of NFAT binding in this region reduced Tat-induced LTR activity, suggesting interplay with Tat(100, 141).

This is not an extensive list of all regulatory transcription factor binding sites that have been reported within the LTR. However, sites not listed here are poorly characterised, particularly in subtype C viruses, and their effect on viral transcription is unclear. Furthermore, the role of the LTR in HIV-1 latency is under-characterised. We will first review the literature on HIV-1 latency below before reviewing the contribution of the LTR to this latent state.

1.7 HIV-1 latency

HIV-1 has the ability to stably integrate its genetic material into a host cells genome (section 1.5). This proviral DNA can lie dormant, in a state known as latency until it is reactivated to produce new viruses later on during infection. Primary evidence for HIV-1 latency was first demonstrated by Folks *et al.*(142), where viral gene expression, in cells that survived initial infection and were unable to produce virus under normal conditions, was induced after exposure to a number of T cell activation compounds.

HIV-1 infects CD4+ T cells which are in an immune activated state, which is conducive to novel HIV-1 production, often resulting in cell death(143–145). Therefore, it is presumed that the phenomenon of latency occurs when a target activated CD4+ T cell is infected, but does not experience cell death despite production of virions within the cell, and instead reverts to a resting memory state(146, 147). However, a number of studies have provided evidence to demonstrate that these latently infected cells can arise from direct infection of resting CD4+ T cells(148–151). This population of infected resting memory cells constitutes the majority of the cellular latent reservoir(152). Seeding of this latent reservoir has been shown to take place early during the acute phase of infection, as early as three days post-infection(153, 154). Productive viral gene expression cannot occur in these resting CD4+ memory T cells, as the necessary transcription factors and other host factors essential for HIV-1 replication are not expressed(155). Conversely, studies have also demonstrated establishment of latency in actively proliferating CD4+ T cells(151, 156, 157), therefore, in addition to

the immune activation state of the infected host cell, a number of other molecular mechanisms contribute to establishing and maintaining HIV-1 latency, these will be reviewed in section 1.7.1.

Latent proviral DNA does not produce new virus or viral proteins, and the immune system and current therapies are unable to detect, target or eliminate it (158, 159). In addition, resting CD4+ memory T cells persist in the infected individual for many years due to an extended cell half-life of over 3.5 years on average(160). Therefore, latent infection allows for HIV-1 to remain present for the entire lifespan of infected patients who are on therapy(161). The persistence of replication competent proviral DNA for long periods of time allows for rapid viral rebound from within a period of days up to approximately 12 weeks of ART interruption(10, 162, 163). Therefore, latency poses an immense barrier to a cure and also necessitates lifelong ART.

1.7.1 Molecular mechanisms of latency

A comprehensive understanding of how latency is established and maintained has not been demonstrated to date. This is largely due to the difficulty of studying latently infected cells as they are rare, occurring at an estimated rate of only 10-100 cells per million CD4+ T cells(164). Nevertheless, there is evidence to show that a number of molecular mechanisms and factors contribute to the establishment and maintenance of HIV-1 latency and these are reviewed in the sections below.

1.7.1.1 Nature of integration

It would be expected that, due to the diversity of the human genome, HIV-1 expression would be influenced by the specific locus into which the virus integrated. However, it is widely accepted that HIV-1 does not integrate in a random manner, but rather the virus is reported to integrate in expressed regions of the genome(165, 166). Furthermore, Han *et al*.(167) demonstrated that 93% of the integration sites sampled within resting memory CD4+ T cells resided in intronic regions of actively transcribed genes. This observation would seem counterintuitive considering the lack of viral gene expression from these resting memory cells. A number of mechanisms have been

proposed where the active transcription of host genes may interfere with viral gene expression resulting in latency, including steric hindrance, promoter occlusion and enhancer trapping. This interference can be influenced by the proximity of the viral genome to a host cell promoter or the orientation of integration of the virus(168, 169).

Steric hindrance occurs when the HIV-1 genome integrates downstream of a host gene promoter in the same orientation. In this scenario, an RNAPII transcription complex from the host gene promoter reaches the HIV-1 5'LTR and dislodges a viral transcription pre-initiation complex thereby preventing viral expression. This displacement was demonstrated in two studies wherein specifically a reduction in binding of the Sp1 transcription factor to the HIV-1 LTR was observed(170). Contrary to these findings, a study by Han *et al*.(168) found that upstream transcription of host genes was capable of increasing HIV-1 expression during RNAPII read-through, if the virus was integrated in the same orientation.

Alternatively, HIV-1 may integrate in the opposite orientation compared to a nearby host gene. In this instance, during elongation, an RNAPII complex from a host gene may collide with the RNAPII complex which initiated at the HIV-1 LTR resulting in early termination of one or both of the transcription complexes(171). This phenomenon is termed promoter occlusion(79). There is evidence to show that convergent RNAPII complexes do not actually displace each other, but rather simply stall at the site of collision and remain stable, suggesting that complete elongation from either RNAPII is not possible(172). Furthermore, integration in this manner can result in both DNA strands of the provirus being transcribed, leading to the production of double-stranded RNA molecules. These dsRNA molecules can be processed into small interfering RNA (siRNA) capable of preventing transcription by RNA interference, targeted degradation of mRNA transcripts via antisense RNA guides or siRNA associated DNA methylation(173–177).

The final mechanism of suppression influenced by HIV-1 is known as enhancer trapping. This trapping effect takes place when the virus integrates in close proximity to a host cell gene. The upregulatory effect of the 5'LTR enhancer region is hijacked by the host gene preventing enhancement of viral expression, leading to a reduction or total repression of HIV-1 transcription(178, 179).

Overall, the collective influence of the nature of integration on HIV-1 latency is not well characterised and can be highly variable(180). It is likely that a range of other molecular mechanisms also contribute to viral latency.

1.7.1.2 Epigenetic control and chromatin environment

Chromatin is the molecular complex of DNA and proteins that forms chromosomes within cells and can be found in two general forms, euchromatin and heterochromatin(181–183). Briefly, euchromatin is typically gene-rich and less condensed allowing for active transcription of gene regions. Conversely, heterochromatin does not harbour many genes and is highly condensed. In this condensed state, gene expression is prevented as access to transcription factors essential for transcription initiation and elongation is hindered.

Chromatin is composed of nucleosomes, where DNA is wrapped around histone octamers. In latently infected cells, two nucleosomes, nuc-0 and nuc-1, lie within the LTR upstream of the modulatory region and immediately downstream of the transcriptional start site respectively. These nucleosomes impose a transcriptional block by preventing RNAPII movement(184). Initiation of transcription necessitates the remodelling of nuc-1 to alleviate this block(185).

A range of cellular proteins are capable of targeting histones and altering the state of this chromatin in a reversible epigenetic manner, which in turn alter gene expression. Histones associated with transcriptionally silent proviral DNA, within heterochromatin, are characterised by deacetylation and high levels of methylation of lysine residues in histones. These modifications aid in maintenance of the repressive state(186). Primarily, HDACs are recruited by *cis*-acting regulatory factors bound to the LTR (section 1.6) such as YY1, LSF, AP-4, C-promoter binding factor-1 (CBF-1) and the NF-κB p50/p50 homodimer(136, 187–189). Histone methyltransferases (HMTs) including, Enhancer of Zeste 2 (EZH2), suppressor of variegation 3–9 homologue 1 (Suv39H1) and Euchromatic histone-lysine N-methyltransferase 2 (EHMT2/G9a)

methylate lysine residues in histones, which in turn recruits heterochromatin remodelling proteins(187, 190–192).

Finally, direct methylation of DNA in close proximity of nuc-1 by DNA methyltransferases (DNMTs) has also been reported to contribute to HIV-1 latency establishment(193, 194). It was proposed that this methylation may either result in recruitment of HDACs or alternatively, directly inhibit binding of necessary transcription factors(195). Conversely, low levels of LTR methylation were observed in on-therapy patients suggesting that this mechanism may not play an important role in transcriptional silencing(196, 197).

1.7.1.3 Availability of cellular transcription factors

The HIV-1 LTR sequence contains binding sequences for several inducible regulatory cellular transcription factors (section 1.6.2). Many of these factors are essential for initial transcription of HIV-1, most notably NF-κB, NFAT and Sp1. NF-κB and NFAT are sequestered to the host cell cytoplasm in resting CD4+ T cells limiting initiation of transcription(198, 199). Furthermore Sp1 was downregulated in resting T cells compared to activated cells(200).

1.7.1.4 Transcriptional control by Tat/P-TEFb

During the second phase of HIV-1 transcription i.e. the Tat-dependent phase, much higher levels of transcription are observed(201). RNAPII and other complexes including TFIID, collectively the PIC, initiate RNA synthesis at the TAR region of the LTR(80). Soon after initiation this complex experiences promoter-proximal pausing to allow for 5’ capping of the nascent RNA strand before resumption of transcription(202). This continuation is hindered in HIV-1 by the negative transcription elongation factor (N-TEF) complex bound at the nuc-1 region (section 1.7.1.2), leading to the premature abortion of viral transcripts(203–205). Productive transcription/elongation of transcription takes place after expression and interaction of viral Tat with the TAR loop which recruits the cellular P-TEFb complex (CycT1 and CDK9)(72).

P-TEFb promotes transcription via phosphorylation of RNAPII and promoting dissociation of the N-TEF complex from the TAR loop(206, 207). Tat/TAR interaction additionally recruits other cellular cofactors such as the histone modifiers p300 and CBP presumably to maintain a conducive chromatin environment for transcription and may also contribute to assembly of the PIC(208, 209). Therefore, the action of Tat and P-TEFb cumulatively result in the promotion of efficient elongation and clearance of inhibitory factors.

As described in section 1.5, a Tat positive feedback loop is established in HIV-1 infection. Entry into latency is thought to be a consequence of disruption of this feedback loop by down-regulation of Tat expression(210). Interestingly, HIV-1 does not independently encode any genes that actively repress viral expression; instead this process relies mostly on cellular factors.

P-TEFb activity is regulated in both resting and activated cells. In resting T cells, expression of the essential CycT1 subunit is heavily repressed via microRNA-mediated mechanisms(211). CDK9 is constitutively expressed in resting cells, however it exists in an inactive form due to dephosphorylation of its regulatory activation loop(212). Therefore, P-TEFb cannot form or function efficiently in resting cells and transcription of HIV-1 is inhibited.

In the absence of Tat in activated CD4+ T cells, P-TEFb subunits are present at high levels in their active form. However, the majority of P-TEFb is sequestered into a repressive complex known as 7SK small nuclear ribonucleoprotein (snRNP), rendering it inactive as it cannot be accessed by the LTR(213). When Tat is expressed in an activated cell, it is capable of disrupting the 7SK snRNP complex, liberating the P-TEFb complex allowing it to act on HIV-1 expression(214). Furthermore, another cellular protein BET bromodomain protein 4 (BRD4) is capable of recruiting P-TEFb from the 7SK snRNP complex, thus competing with Tat and resulting in viral repression(215).

Therefore, P-TEFb exists in an active free form and an inaccessible form sequestered by 7SK snRNP, and the equilibrium of these forms is responsible for control of viral repression/expression(216).

Taken together, is it clear that HIV-1 latency is a highly complex phenomenon with a plethora of mechanisms acting together to induce the latent state.

1.7.1.5 Role of the LTR genotype in latency

A role for the HIV-1 LTR, in the establishment of latency has been proposed. Two studies have directly implicated the genotype of the LTR in latency establishment(101, 156).

Firstly, Duverger *et al*.(101) identified an AP-1 binding sequence slightly upstream of the transcriptional start site within the LTR that varied across subtype isolates. They found that this sequence independently altered the proportion of cells that became latently infected by a particular isolate using mutational analysis and assessing the ability of these viral isolates to establish latency in an inducible reporter cell line. The 4 nt motif in subtype B viruses showed reduced latency potential compared to the 7 nt counterpart found in subtype A or C viruses. Extension of this 4 nt motif to the 7 nt increased the latency potential of the infecting virus, whereas complete deletion of the motif impaired the latency potential of the virus.

Secondly, Dahabieh *et al*.(156), confirmed that subtype level genotypic variance of the entire LTR influenced viral sensitivities to latency. Specifically, this group used a reporter-based model (described in detail in section 1.8.2.3) in which a CD4+ T cell line was infected with reporter viruses harbouring a subtype-specific LTR. It was demonstrated that the majority of cells were latently infected shortly after infection. Subtype D and F had the highest latency potential, with subtype AE displaying the lowest latency potential. A specific mechanism of LTR silencing or features of the LTR associated with higher latency potential were not defined in this study.

Conversely, van der Sluis *et al*.(217), reported no major disparities in initial levels of latency across HIV-1 subtype LTRs, with the exception of subtype AE where lower levels of latent cells were observed upon infection. However, this study utilised a model which measured levels of p24 antigen i.e. products of HIV-1 particles, whereas

the aforementioned models directly measured the transcriptional activity of integrated provirus.

1.8 Models for assessing HIV-1 latency

A number of *in vitro* and *in vivo* models have been developed to mimic HIV-1 latency such that latently infected cells can be evaluated under a more controlled set of conditions. Examples of these models are discussed below.

1.8.1 *In vivo* models

It is possible to replicate/model HIV-1 latency in Simian Immunodeficiency Virus (SIV)-infected non-human primates following administration of ART. Four studies have used different combinations of ART to lower the viral loads in infected macaques to similar levels of those in virally suppressed HIV-1 infected patients(218–221). In these animals, researchers were able to interrogate different anatomical regions to attempt to isolate replication competent virus, measure the effect of ART on frequencies of resting cells with replication competent virus and measure levels of viral DNA and RNA within resting infected cells. However, such experiments are expensive, time consuming and the viral latency phase is much shorter compared to that of humans. Another caveat is that SIV, despite having high similarity to HIV, is still genetically distinct and properties of the two viruses may differ(222, 223).

A number of *in vivo* murine models have also been developed, where immuno-deficient mice are humanized via human stem cell, tissue transplant or a combination of the two. These include the bone marrow-liver-thymus (BLT) SCID-hu (Thy/Liv) and T cell only mice (ToM)(224–226). These murine models allow for measurement of different properties of HIV-1 infection. All models produce functional human immune cells and are therefore capable of supporting HIV-1 infection and replication and establishment of latency. Furthermore, human cells in these mice have been shown to react similarly to drugs as they do in humans, and latently infected HIV-1 cells can be isolated from the models after viral suppression after ART treatment(227–229). Of

note, one model, the T cell only mice (ToM) was developed where mice have high levels of human T cells in all tissues and additionally, lack other antigen presenting cells including B cells, dendritic cells macrophages and monocytes(225). Therefore, it is possible to evaluate HIV-1 latency exclusively in T cells without interplay of other immune cells.

Limitations of murine models include the small size and short lifespan of the mice meaning that minimal amounts of blood and tissue can be obtained from infected mice for study(230). Memory cell infection is also difficult to study in these models as the majority of human cells are naïve(222). There are also fundamental anatomical differences between human and mice tissue structure that may influence infection dynamics. Finally, in some cases animals can develop graft-versus-host disease(231).

1.8.2 *In vitro* models

A variety of *in vitro* latency models have also been developed which have number of advantages over the *in vivo* models described above. In general, these models are less expensive to maintain, and assays are faster and less tedious. Three such models are discussed below.

1.8.2.1 Cell-line models

A number of groups have developed cell-line based models of HIV-1 latency(232–234). In these models, cells are chronically infected and little constitutive HIV-1 gene expression is observed. However, significant HIV-1 gene expression can be induced upon treatment of these cells with a variety of stimulating cytokines, mitogens and T cells activation compounds. These include the ACH-2, U1 and JΔK cell lines. These models have provided a system to study viral integration sites, mechanisms of viral latency and host gene expression post viral induction(204, 235–237).

Despite the usefulness of these cell lines, their relevance and ability to truly represent viral latency within an infected patient has come under question. It was subsequently found that mutations within the *tat* gene and TAR region of the viral genome (section 1.3) were responsible for the low levels of constitutive viral expression in ACH-2 and

U1 respectively(238, 239). Similarly, the latent phenotype of JΔK cells is due to the deletion of an NF-κB site within the LTR of the integrated infecting virus(234).

Another cell-line based latency model, J-Lat, was later developed(240). In this model a T lymphocyte cell line, Jurkat, was infected with a full-length HIV-1 vector derived from a replication incompetent pseudovirus expressing green fluorescent protein (GFP) and Tat under the control of an LTR. Latently infected cells were enriched for by selection of cells with integrated virus that did not express GFP in the absence of activation, using fluorescence-activated cell sorting (FACS). Clones of these cells were expanded and were shown to express viral proteins and GFP after cell activation. A number of different stimuli were tested with this model to assess their reactivation potential on latency. Other studies have used the J-Lat cell-line to study aspects of transcription and interactions of cellular factors with the LTR in latent cells(188, 241). More recently, this cell line was used to show the efficacy of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) mediated latency activation and the identification of the mammalian target of rapamycin (mTOR) complex and its influence on viral latency(242–244).

Once again there are limitations to this model, due to it being established using a replication incompetent virus. J-Lat clones have single sites of integration that are not representative of infection in patients. Furthermore, for all cell-line models, cells constantly undergo replication and are in a state of cellular activation, whereas latently infected resting memory cells are long-lived and only experience low levels of proliferation(245, 246). Hence, verification of findings in cell-line latency models in primary cells is therefore necessary.

1.8.2.2 Primary cell models

A number of primary cell models have been developed wherein infection of activated CD4+ T cells is simulated *in vitro*. Of note, five different groups followed similar procedures to generate a population of latently infected cells(247–251). Firstly, CD4+ T cells were isolated from donors and activated via a number of mechanisms e.g. α-CD3/α-CD28 antibodies in the presence of interleukin 2 (IL-2). Thereafter activated

cells were infected with replication competent viruses, HIV-1 derived vectors or pseudotyped viruses. Infected cells were then returned to a resting state over an extended culture period. The final step of obtaining a resting state in these cells is the most challenging, as cells are prone to death soon after activation unless cultured under specific conditions or cytokine milieu. Therefore, obtaining large populations of latently infected cells is difficult(252). Furthermore, in some cases cells still expressed some markers of activation, indicating that they may not be completely representative of latently infected resting memory cells. One group, Yang et al.(251) was able to circumvent some of these challenges by transducing their activated cells with a Bcl-2 expressing vector. This acts as an anti-apoptotic protein allowing for long term culture of infected activated cells and establishment of latency in the absence of modified growth conditions.

Despite this, all groups were able to demonstrate increased HIV-1 expression in their "resting" infected cells after activation the addition of Tat or latency reversing agents, suggesting latency had been established in some cells. These models have proven useful in reactivation studies and determination of factors contributing to viral latency.

Other models have been developed where resting CD4+ T cells are directly infected(253, 254). Isolated memory CD4+ T cells were directly infected with replication competent HIV-1 by spinoculation. This eliminates the challenging process of allowing activated cells to transition to the resting state and latency is established within a few days of infection. Nonetheless, this model has limitations. Very few cells are infected as resting cells are not highly susceptible to infection due to unfavourable transcription conditions, and the pool of latently infected cells does not survive for extended periods of time without culturing cells in the presence of cytokines as mentioned above(155, 222).

Overall, both active and resting T cell HIV-1 infected latency models provide an *in vitro* method of assessing properties of latently infected cells which more accurately represents the in vivo conditions of infection than clonally derived cell-lines.

1.8.2.3 Plasmid reporter constructs

A number of HIV-1 reporter plasmids have been developed with a reporter gene inserted within the HIV-1 genome under the control of the HIV-1 LTR(255, 256). These plasmids are used to generate viruses that are, in turn, used to infect T cell lines or primary cells, and the activity of the infecting virus can be monitored by measurement of the reporter protein. Essentially, latently infected cells can be identified by the absence of the reporter protein. Burke *et al*.(255) developed a construct harbouring an enhanced-GFP(eGFP):Luciferase fusion gene. In this model, transcriptionally inactive immature CD4+ thymocytes were infected with a pseudovirus generated from the construct. eGFP expression was measured by flow cytometry after infection as these cells matured. A subset of cells was stimulated, and the proportion of latently inducible cells could be measured as eGFP expression increased. Furthermore, these inducible latently infected cells underwent purification by FACS and were used to study LTR regulation.

Sorting of cells to obtain the latently infected subset can be tedious and time consuming. Therefore, other HIV-1 based viral reporter constructs have been developed with two independent reporter genes, one under the control of a strong/constitutive promoter and another under the control of the viral LTR(256). The constitutive reporter acts as an indicator of successful infection and integration of the construct whereas the LTR driven reporter denotes viral activity. Calvanese *et al*.(256) developed two models where the HIV-1 *nef* gene was substituted with either an mApple or eGFP reporter and a spleen focus forming virus (SFFV)p-eGFP or Elongation factor-1 alpha (EF1α)-mCherry cassette inserted downstream, respectively. A population of both productively and latently infected cells were identified simultaneously. Chavez *et al*.(151) made use of the eGFP/EF1α-mCherry to demonstrate direct establishment of latency in both active and resting CD4+ T cells.

An additional HIV-1 based viral reporter was developed by Dahabieh *et al*.(156), the Red-Green-HIV-1 (RGH) model. This model will be referred to throughout this thesis. The original plasmid construct, pRGH (LAI-gagiGFP-CMVmCherry-deltaEnv),

contains a full-length replication competent, infectious HIV-1 subtype B LAI molecular clone(257) (Figure 1.6).

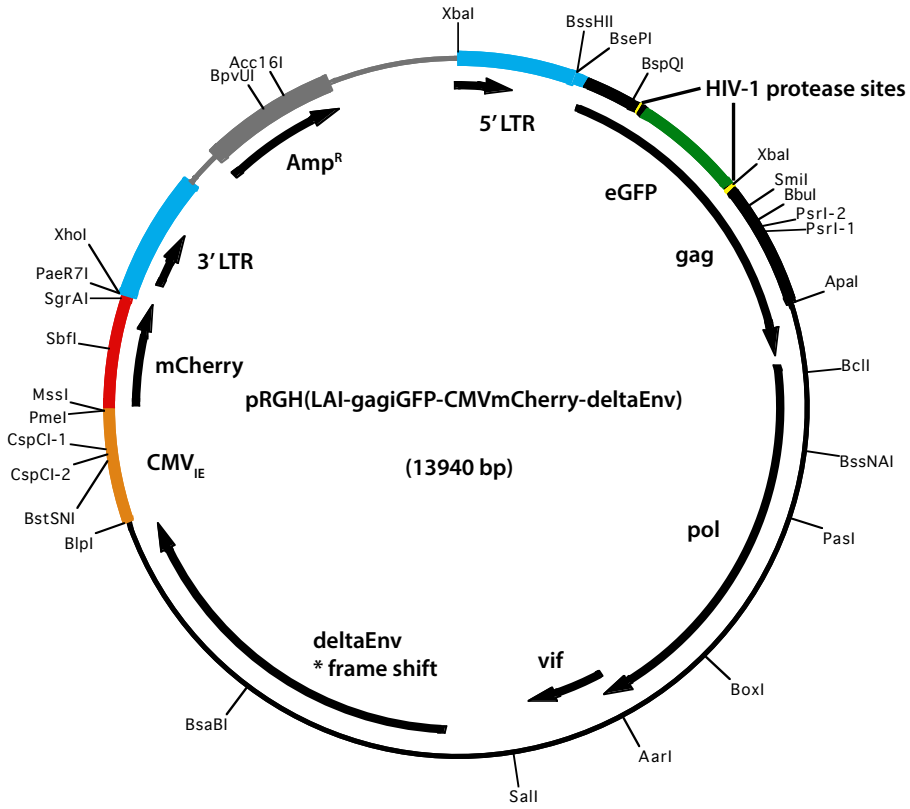


Figure 1.6 Schematic depiction of the RGH plasmid. pRGH is derived from pLAI full length replication competent, infectious HIV-1 subtype B LAI molecular clone. An eGFP gene has been inserted in-frame between the matrix and capsid regions of the *gag* gene. The *env* open reading frame was disrupted by a single base pair deletion at position 7136 resulting in a frameshift. The *nef* gene has been partially deleted and replaced with a CMV_{IE} driven mCherry gene construct. Figure obtained at: http://www.aidsreagent.org/support_docs/12427_map.pdf.

The infectious molecular clone was modified by the authors as follows for the purpose of investigating latency:

- Firstly, the LAI *gag* gene was replaced by a *gag*-eGFP fragment from the Gag-iGFP NL4-3 clone, where an eGFP gene was inserted in-frame between the matrix and capsid regions(258). Insertion of the reporter gene at this position leads to production of a Gag-eGFP-capsid fusion protein and results eGFP-labelled viruses. As the expression of Gag-eGFP is dependent on transcription from the HIV-1 LTR promoter, detection of this protein within infected cells is indicative of HIV-1 expression and can therefore be used to identify actively transcribing HIV-1 LTRs.

- Secondly, the *nef* gene was partially deleted and replaced with a cytomegalovirus (CMV) promoter-driven mCherry gene construct, amplified from the pcDNA3.1+-mCherry(259). CMV_{IE} is a strong promoter, and capable of constitutive expression(260, 261). Therefore, mCherry is expressed, in the host cell, post-integration independently of transcriptional activity of the HIV-1 LTR and can therefore be used to detect cells with successfully integrated virus.
- Finally, the *env* open reading frame was disrupted by a single base-pair deletion at position 7136 resulting in a frameshift. This deletion renders the *env* gene inactive, hence viruses generated from this construct are capable of infecting cells, integrating into the host cell genome and expressing viral proteins, but they are incapable of producing new infectious viral particles. Therefore, further rounds of infection are not possible.

1.9 Latency reversing agents

Currently, strategies to clear the latent reservoir focus on induction of viral gene expression while an infected individual is still on therapy such that the infected cells will be recognized and killed by the immune system, but infection of new cells will not occur. Chromatin structure, methylation state of DNA and the presence/availability of a number of transcriptional factors and activators are all mechanisms that play a role in maintenance of HIV-1 latency (section 1.7). A number of latency reversing agents (LRAs) have been developed to target these molecular mechanisms of latency in the hopes of reactivating latent provirus to purge the latent reservoir and ultimately cure an infected individual.

1.9.1 Histone Deacetylase inhibitors (HDACis)

HDACis are compounds that inhibit HDACs recruited to the HIV-1 LTR during latency and therefore result in relaxation of chromatin facilitating transcription and outgrowth of proviruses(185). This has been the major focus of latency reversal in clinical trials, a number of which are still in progress(262). Multiple studies have demonstrated initiation of transcription in both latently infected T cell models and primary cells

isolated from on-therapy patients. Examples of this class of LRAs include valproic acid (VPA), Vorinostat (SAHA) and Trichostatin (TsA)(185, 263, 264). These compounds have been well characterised in the context of anticancer therapies and some have already been approved by the Food and Drug Administration (FDA) for other treatments(265). However, an *in vitro* study using SAHA demonstrated that the majority of proviral latently infected cells was not reactivated upon HDACi treatment (266).

1.9.2 Histone methyltransferase inhibitors (HMTis)

These compounds also act to prevent the formation of heterochromatin via inhibition of H3K9 methylation (section 1.7.1). Two compounds BIX01294 and chaetocin have been shown to inhibit the action of G9a and Suv39H1 respectively *in vitro* and have been able to induce viral outgrowth from latently infected cells isolated from on-therapy patients. Furthermore, they are capable of enhancing the effect of the HDACis mentioned above(187, 267).

1.9.3 DNA methyltransferase inhibitors (DNMTis)

The degree that direct DNA methylation plays in latency is somewhat debated (section 1.7.1.2). However once again these compounds are hypothesised to prevent formation of transcriptional inhibitory heterochromatin and allow for efficient viral transcription. There is some evidence to demonstrate that the DMNTi Decitabine in addition to the T cell activation agent Tumour Necrosis Factor alpha (TNF-α) is capable of HIV-1 latency reactivation in certain latently infected cell lines(268).

1.9.4 Protein Kinase C activators

The protein kinase C (PKC) pathway is a cellular cascade induced upon T cell activation that stimulates production of a number of HIV-1 upregulatory transcription factors such as NF-κB, NFAT and AP-1(269–271). Therefore, stimulation of this cascade induces HIV-1 expression by binding of these factors to the LTR(272). Many PKC activators have been developed including phorbol 12-myristate 13-acetate

(PMA), prostratin, 12-deoxyphorbol 13-phenylacetate (DPP) and bryostatin. These compounds have been shown to be highly effective at reactivating latency in T cell models and PBMCs and astrocytes from on therapy HIV-1 infected individuals(273–275). However, some PKC activators induce a number of unwanted effects such as tumour promotion and upregulation of a range of other cytokines that can cause cellular toxicity(276–278). This is due to the non-specific nature of these compounds on T cell activation.

1.9.5 Positive transcription elongation factor b activator

The P-TEFb protein is an essential protein complex in viral expression (section 1.7.1.4). Hexamethylene bisacetamide (HMBA) activates HIV-1 transcription by facilitating release of P-TEFb from the 7SK snRNA allowing it to act on the LTR(279). HMBA has shown limited efficacy at reversing latency in a T cell model(251). Furthermore, this compound did not induce CD8+ T cell recognition after treatment of infected cells(280). Another compound JQ1s, capable of reactivating HIV-1 expression in latently infected cell lines, was developed which has an inhibitory effect on Brd4, the complex which competes for P-TEFb(281).

Despite some of these compounds showing some efficacy in reactivating latent proviral DNA in infected cells, to date there has not been any evidence to show that they are capable of depleting the latent reservoir. It has been demonstrated that combinations of some of these LRAs, used on latently infected cells isolated from on-therapy patients, induced higher levels of HIV-1 transcription than when used alone(282). This suggests that combination reactivation therapy may be necessary to effectively reverse latency. Interestingly, the size of the latent reservoir has been implicated in the reactivation capacity of some of these latency reversing agents(283). Therefore, understanding potential factors contributing to latency establishment may be important when making use of LRAs to induce durable remission.

1.10 Rationale

HIV-1 latency represents the greatest barrier to a cure. The HIV-1 LTR, serves as the viral promoter element that regulates HIV-1 gene transcription. Subtype level genotypic variation within the LTR, has been implicated in influencing the establishment of viral latency and has been shown to affect the proportion of infected cells in which viruses enter the latent state. Furthermore, variation of the U3 region of the LTR has been shown to influence the transcriptional activity of the virus(77, 284). The U3 region encodes regulatory functional sites to which host cell factors such as transcription factors and cytokines as well as the viral protein Tat, which upregulates LTR promoter activity, bind. There is limited information on the role of the LTR, and more specifically of subtype C viruses, which dominate globally, in viral latency.

Further understanding of the molecular mechanisms that influence latency is therefore vital in developing strategies to target latent reservoirs as part of cure and remission strategies.

This study provides insights into the influence that the subtype C LTR has on establishing latency during early HIV-1 infection. More specifically it characterises the influence of the LTR genotype on a participant level. In addition, this study investigates potential correlates of LTR-driven latency potential including Tat-induced transcriptional activity.

1.11 Aims and Objectives

1.11.1 Aim

To determine the influence of HIV-1 LTR genotype and transcriptional activity on the ability of viruses to establish latency within CD4+ T cells.

1.11.2 Objectives

1. Generate a panel of LTR-pseudotyped viruses from acutely-infected individuals from the CAPRISA 004 cohort
2. Determine the latency potential of each pseudovirus in a CD4+ T cell-line
3. Measure Tat-induced promoter activity of participant LTRs using luciferase reporter constructs
4. Evaluate potential correlates of LTR-driven latency potential

Chapter 2: Materials and Methods

2.1 Ethics Statement

This is a sub-study of the project titled “Viral set point and clinical progression in HIV-1 Subtype C infection: the role of immunological and viral factors during acute and early infection”. This study was approved by the University of Cape Town’s Faculty of Health Sciences Human Research Ethics Committee (UCT FHS HREC) (reference number: 025/2004) and the University of KwaZulu-Natal Biomedical Research Ethics Committee (reference number: E013/04). Ethical approval for this sub-study was obtained from the UCT FHS HREC (reference number 574/2016). Informed consent was obtained from all individuals enrolled.

2.2 Samples

The LTR (HXB2 positions 9086-9621) was amplified from viral isolates generated from 39 participants in the CAPRISA 002 cohort(285). Viruses were cultured in peripheral blood mononuclear cells (PBMCs) over a median of 25 days (range: 25-58) and the LTR was cloned into the pGL4.10 vector (viral culturing and clones were provided by, P Selhorst, UCT)(285). A 330bp 3’LTR sequence (HXB2 position 9086-9415), from plasma virus from acute infection and 12 months post-infection, was available for 26 of the participants (provided by D. Chopera, UCT/AHRI)(286). The CAPRISA 002 cohort recruited women who seroconverted during the CAPRISA 004 1% tenofovir microbicide gel trial. This is a cohort of sexually active South African women aged between 18 and 40 years recruited as HIV-1 negative in high risk urban and rural KwaZulu-Natal. Enrolled women were randomised in equal proportions into two study arms, the 1% Tenofovir disoproxil fumarate (TDF) microbicide gel or placebo gel. Women were followed up monthly and two HIV rapid tests (Determine HIV 1/2 (Abbott Laboratories, IL, USA) and Uni-Gold Recombigen® HIV test (Trinity Biotech, Wicklow, Ireland)) were performed at each visit. In cases of possible seroconversion, two independent RNA PCR assays were performed(287, 288). The date of infection was defined as the midpoint between the last HIV negative test and the first HIV positive test.

2.3 Sequence analysis

Sanger sequencing was performed at the Central Analytical Facilities at the University of Stellenbosch, South Africa, using the ABI3000 Genetic Analyzer and BigDye terminator reagents (Applied Biosystems, CA, USA). Sequences were compiled using Sequencher (v5.3) and visualised and aligned in AliView (v1.21). Maximum likelihood phylogenetic trees were compiled and pairwise DNA distances calculated in MEGA 7 and trees were edited in Dendroscope (v3.5.9). Nucleotide differences in cloned participant LTR sequences were visualised on a Highlighter plot (https://www.hiv.lanl.gov/content/sequence/HIGHLIGHT/highlighter_top.html).

The consensus subtype C LTR used for all alignments, pairwise distance calculations and phylogenetic analysis was generated from a curated web alignment of a total of 373 subtype C LTR sequences (HXB2 position 1-532, up to and including sequences from 2016), obtained from the Los Alamos National Laboratories Sequence Database (http://www.hiv.lanl.gov/), using Consensus Maker tool (https://www.hiv.lanl.gov/content/sequence/CONSENSUS/consensus.html).

To confirm the subtypes of our selected participant derived virus panel, a maximum likelihood phylogenetic tree was generated from a nucleotide alignment of the 26 participant derived LTR sequences, 50 subtype C South African and each subtype (A through G), consensus LTR sequences (obtained from the Los Alamos National Laboratories HIV Sequence Database). To confirm that pGL4.10 cloned LTR sequences were representative of plasma derived viruses, another maximum likelihood phylogenetic tree was generated from an alignment where the pGL4.10 cloned sequence was aligned with the acute plasma sequences and 12 months post-infection. Sequences were trimmed to the overlapping region (HXB2 position 9086-9415).

2.4 Gel electrophoresis

DNA fragments were visualised using agarose gel electrophoresis (1% w/v). Agarose gels were made up with 1X Tris-Borate-EDTA (TBE) buffer (Appendix 1). 6X Loading dye (Thermo Scientific, MA, USA) containing GelRed™ (Biotium, CA, USA) was added to all DNA samples to a 1X final concentration before electrophoresis. GelRed™ allows for the visualisation of DNA molecules on an agarose gel as it is an intercalating agent that interacts with DNA and fluoresces upon exposure to ultraviolet (UV) light. Electrophoresis was performed at 100V for between 60 and 75 minutes. Gels were visualised under UV light using a Gel Doc™ XR+ Gel Documentation System (Bio-Rad, CA, USA).

2.5 pRGH reporter vector

The following reagents were obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: pRGH-WT (#12427) and pRGH-ΔU3 (3') (#12430) from Ivan Sadowski and Viviana Simon(156, 257–259, 289–293). These constructs were used to determine latency potential and for the generation of flow cytometry compensation controls (sections 2.6, 2.9, 2.10 and 2.11).

2.6 Construction of Flow cytometry compensation controls

The emission spectra of mCherry and eGFP overlap. Flow cytometry compensation controls were constructed expressing only mCherry or eGFP to correct for false positive detection of either fluorescent protein during flow cytometric analysis (Figure A2.1, Appendix 2).

2.6.1 Construction of an eGFP only reporter plasmid

To generate the eGFP only expressing pseudovirus, the CMV promoter and mCherry gene were deleted from the parent plasmid [pRGH (LAI-gagiGFP-CMV-mCherry-

deltaEnv)] by restriction enzyme digestion. The Thermo Scientific FastDigest restriction enzyme kit (Thermo Scientific) was used following the manufacturer protocol with these modifications: a total of 3µg of the pRGH plasmid in a final volume of 20µl was double-digested with 30 units each of FastDigest BlpI and XhoI restriction enzymes (Thermo Scientific) for 2 hours at 37˚C. The total volume of digested products was visualised by agarose gel electrophoresis (section 2.4) and the larger DNA fragment was excised and purified using the QIAquick Gel Extraction Kit (QIAGEN, Hilden, Germany) following the manufacturer protocol.

A blunting reaction was performed on the isolated DNA to remove the incompatible extended sticky ends/overhangs, such that blunt end self-ligation of the plasmid was possible. The Thermo Scientific CloneJET PCR Cloning Kit (Thermo Scientific) was used following the manufacturer protocol with these modifications: 625ng of digested linear DNA product in a total volume of 20µl with 10 units of DNA blunting enzyme. The blunting reaction was set up on ice followed by incubation at 70˚C for 5 minutes. The DNA Blunting Enzyme is a thermostable DNA polymerase with proofreading activity, capable of removing 3' overhangs and filling in 5' overhangs to produce dsDNA ends.

The blunt-ended linear plasmid was then self-circularized using the Thermo Scientific T4 DNA Ligase Self Circularization of Linear DNA Kit (Thermo Scientific) following the manufacturer protocol with these modifications: 100ng of blunted linear DNA was added to the ligation reaction mixture, 50% (w/v) polyethylene glycol (PEG 4000) solution was added to the ligation reaction mixture at a final volume of 5% (w/v) and the final volume of the ligation reaction mixture was 20µl. PEG 4000 greatly increases the efficiency of the ligation reaction. The mixture was incubated overnight at 16˚C. A volume of 5µl of the ligation mixture was subsequently used to transform *E. coli* XL10-Gold® Ultracompetent Cells (Agilent Technologies, CA, USA). These are cells optimised for transformation of large DNA molecules with high efficiency. The XL10-Gold® Ultracompetent Cells Transformation manufacturer protocol was followed with these modifications: 50µl of competent cell mixture was used, 2µl of the β-mercaptoethanol (β-ME) mix was added to the cells prior to DNA addition, 200µl of Super Optimal broth with Catabolite repression (SOC) was added to the

transformation mix post heat-pulse. The total transformation mix was plated on Luria-Bertani (LB) agar plates containing carbenicillin at a final volume of 100µg/ml (Appendix 1) and incubated overnight at 32˚C (Appendix 2).

Bacterial colonies were then screened using Super-Therm *Taq* polymerase (JMR Holdings, London, UK) according to manufacturer instructions to identify clones lacking the CMV promoter and mCherry gene. Primers were designed to the *nef* gene and a region within the RGH plasmid downstream of the 3'LTR. Specifically, the Nef F forward and pRGH Colony RP reverse primers were used (Table A2.1, Appendix 2). The final concentrations of PCR reagents per reaction were as follows: 1X Super-Therm buffer, 0.015U/µl Super-Therm *Taq* Polymerase, 2mM magnesium chloride ($MgCl_2$), 160µM each dNTPs and 0.15µM of each primer. The following thermal cycling conditions were used; 94˚C for 3 min followed by 45 cycles of 94˚C for 30s, 60˚C for 45s and 72˚C for 2 min, followed by a final elongation step at 72˚C for 10 min. The PCR products were analysed by agarose gel electrophoresis (section 2.4). Colonies from which the correct DNA fragment was amplified were inoculated in 4ml of LB broth (Appendix 1) and incubated at 32˚C overnight (Appendix 2). Plasmid DNA was isolated from overnight cultures using the QIAprep® Spin Miniprep Kit (QIAGEN) following the manufacturer protocol.

To confirm deletion of the CMV and mCherry genes, 500ng of the plasmid DNA was restriction digested with 5 units of FastDigest KpnI (Thermo Scientific) for 1 hour at 37˚C following the Thermo Scientific FastDigest manufacturer restriction enzyme protocol. Digested products were analysed by agarose gel electrophoresis (section 2.4), clones from samples with correct sized bands of 7300bp, 4830bp and 330bp were sequenced. Plasmid DNA was Sanger sequenced (section 2.3) using the Nef F forward and pRGH Colony RP (Table A2.1, Appendix 2). Clones with successful deletion were designated as pRGH-ΔCMV/ΔmCherry.

2.6.2 Construction of an mCherry-only reporter plasmid

To generate the mCherry-only expressing pseudovirus, the eGFP gene was removed from the parent plasmid (pRGH (LAI-gagiGFP-CMV-mCherry-deltaEnv)) by inverse

PCR, using the Phusion High-Fidelity PCR Kit (New England Biolabs, MA, USA) following the manufacturer protocol. Phusion is an artificially constructed thermostable polymerase with the *Pfu* DNA polymerase fused to the Sso7d DNA binding domain, resulting in an enzyme with extremely high fidelity(294). This polymerase was therefore used to minimise the introduction of nucleotide mismatches during the PCR reaction. The ^eGFP-pRGH-FP forward and ^eGFP-pRGH-InvPCR(R) reverse primers were designed to the start of the p24 and end of p17 regions of the *gag* gene (Table A2.1, Appendix A2), respectively. The final concentrations of PCR reagents per reaction were as follows: 1X Phusion HF buffer, 0.02U/µl Phusion High-Fidelity Polymerase, 200µM each dNTPs and 0.2µM of each primer. The following thermal cycling conditions were used; 98˚C for 2 min followed by 35 cycles of 98˚C for 10s, 67˚C for 50s and 72˚C for 8 min, followed by a final elongation step at 72˚C for 10 min.

The Thermo Scientific FastDigest manufacturer restriction enzyme protocol was followed. The total volume of PCR product (50µl) was digested with 20 units of FastDigest DpnI (Thermo Scientific), a restriction enzyme that fragments methylated DNA resulting in removal of residual template DNA. Digestion reaction was incubated at 37˚C for 2 hours. The DpnI digested product was then self-circularized using the Thermo Scientific T4 DNA Ligase Self Circularization of Linear DNA Kit following the manufacturer protocol with the following modifications: 5µl of DpnI digested DNA was added to the ligation reaction mixture, PEG 4000 Solution was added to the ligation reaction mixture at a final volume of 5% (w/v) and the final volume of the ligation reaction mixture was 50µl. The ligation reaction mixture was incubated overnight at 16˚C. A volume of 5µl of the ligation mix was used to transform XL10-Gold® Ultracompetent Cells (section 2.6.1).

Bacterial colonies were screened for clones lacking the eGFP gene by PCR amplification of DNA in colonies following the Supertherm *Taq* polymerase protocol. Primers were designed to the p24 and p17 regions of *gag*. The GF-40 forward and Gag-BR reverse primers were used (Table A2.1, Appendix 2). Cycling conditions, colony PCR screening, plasmid DNA isolation protocols and gel electrophoresis used were as described (sections 2.4 and 2.6.1).

To confirm deletion of the eGFP gene, 500ng of the plasmid DNA was restriction digested with 5 units of FastDigest HindIII (Thermo Scientific) for 1 hour at 37˚C following the Thermo Scientific FastDigest manufacturer restriction enzyme protocol. Digested products were analysed by agarose gel electrophoresis (section 2.4), clones from samples with correct sized bands of 6470bp, 2620bp, 1450bp and 1210bp were sequenced. Plasmid DNA were Sanger sequenced (section 2.3) using the gagD-F, GF-40 and GF-80 forward and the gagB-R and gagD-R reverse primers (Table A2.1, Appendix 2). Clones with successful deletion were designated as pRGH-ΔeGFP.

2.7 Cloning of participant-specific LTRs

Previously, a panel of isolate-derived patient-specific LTR sequences were each cloned into the pGL4.10 vector (Figure A2.2, Appendix 2) (provided by, P. Selhorst, UCT). pGL4.10 is a promoterless vector with a multiple cloning site upstream of a luciferase reporter gene to allow for cloning and measurement of activity of a promoter of choice.

Participant LTR sequences were amplified from the pGL4.10 plasmid panel following the Phusion High-Fidelity PCR Kit protocol. The pRGH_LTR_F forward and pRGH_LTR_R reverse primers (Table A2.1, Appendix A2) were designed to the *nef* gene region 50bp upstream of the start of the U3 region of the LTR and the U5 region of the LTR respectively, such that the amplicon included the entire U3 and R region of the LTR. The resulting ~650bp PCR product included HIV-1 viral sequence from HXB2 position 9014 to 9621. Final concentrations of reagents were as described in section 2.6.2. The following thermal cycling conditions were used; 98˚C for 30s followed by 35 cycles of 98˚C for 10s, 60˚C for 30s and 72˚C for 5 min, followed by a final elongation step at 72˚C for 5 min.

The PCR products were purified following the QIAquick PCR Purification Kit Protocol (QIAGEN). To remove the unnecessary flanking sequence, 500ng of purified DNA samples were simultaneously digested with 5 units each of the FastDigest XhoI and HindIII restriction enzymes, following the Thermo Scientific FastDigest manufacturer restriction enzyme protocol. Digested products were purified, following the QIAquick PCR Purification Kit Protocol, in preparation for cloning into the RGH-ΔU3 vector

(Figure A2.3, Appendix 2). This version of the plasmid was chosen for cloning as the participant LTR sequences to be cloned mostly comprised of this U3 region. Therefore, this deletion prevented interference or potential recombination between the existing U3 and the sequence to be cloned.

Due to the lack of unique restriction sites at the cloning site in the pRGH-ΔU3 plasmid, traditional restriction enzyme (sticky-end) cloning was not possible. Therefore, an adaption of restriction-free (RF) cloning was performed (Figure 2.1). RF-cloning is a PCR-based method designed to introduce a DNA region of interest into a circular DNA vector without restriction digestion or ligation. Canonically, primers are designed to add additional base pairs to the DNA sequence to be cloned that are complimentary to regions of the target plasmid. This is accomplished by a standard PCR reaction. This sequence with added base pairs is then used in a secondary PCR reaction as a "mega-primer" which binds to the target plasmid "template" via complimentary sequences. The double stranded DNA sequence to be cloned becomes single stranded during denaturation and the complimentary regions anneal to the target plasmid on both ends with the sequence to be cloned behaving as an overhanging loop. Double stranded DNA is generated from both annealed ends and the sequence to be cloned is incorporated as the PCR is completed(295, 296). However, when cloning participant LTRs into the pRGH-ΔU3 plasmid complimentary regions to approximately 90bp of the *nef* region and 80bp of R region of the LTR were present in the 650bp amplicon described above, therefore addition of excess complimentary bases on either end was not necessary. A schematic of the RF-cloning procedure is shown below.

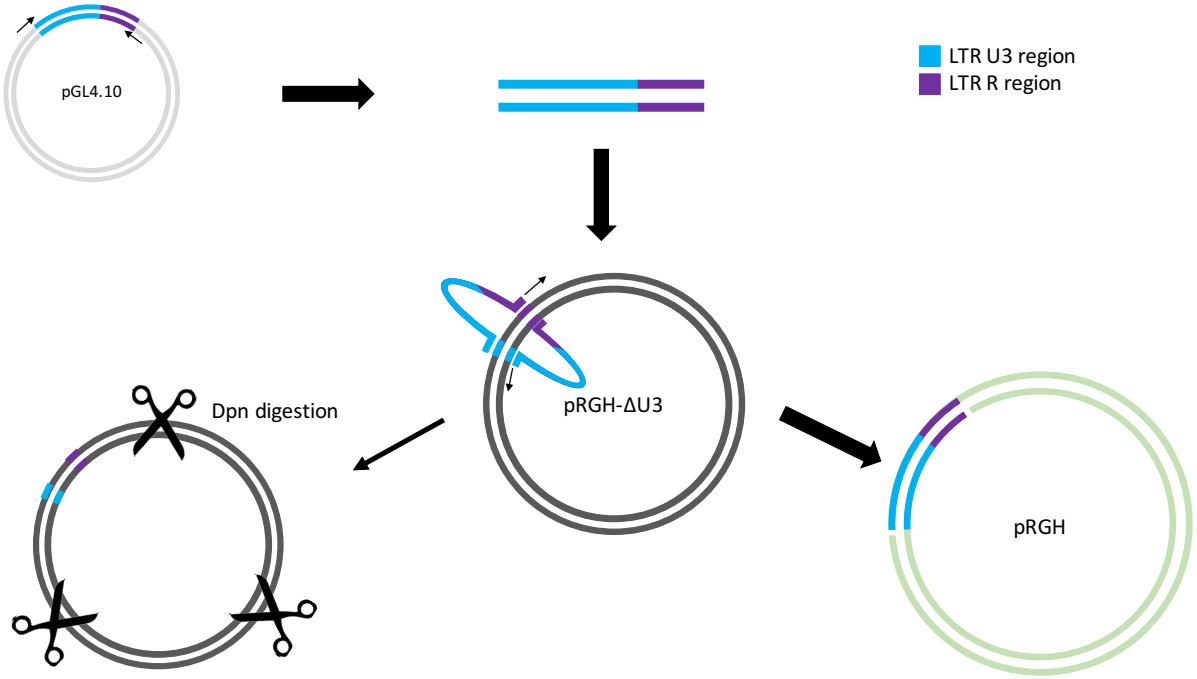


Figure 2.1 Schematic representation of the RF-cloning reaction. Primers were designed to regions flanking the U3 and R regions of the LTR which had previously been cloned into the pGL4.10 vector. During a first round PCR the LTR is amplified out of the pGL4.10 plasmid to create a ”mega-primer”. This “mega-primer” is used in a secondary PCR reaction where the 5’ and 3’ ends of the amplicon hybridise to the corresponding regions flanking the U3 and R regions of the LTR, which are also present in the pRGH-ΔU3 plasmid acting as the DNA template. Replication is initiated from each end of the primer and the entire pRGH-ΔU3 plasmid is replicated incorporating the mega-primer into the sequence. PCR amplification results in production of a double nicked circular double-stranded DNA molecule. DpnI digestion is performed on the secondary PCR to selectively remove parental template pRGH-ΔU3 DNA leaving the nicked dsDNA molecule with the LTR of interest incorporated intact.

A total of three condition sets were used to attempt the RF-cloning procedure with the Phusion High-Fidelity PCR Kit. Initially, a total of 125ng of patient-specific LTR “mega-primer” and 10ng of the pRGH-ΔU3 “template” was added to the reaction in a final volume of 50µl. The final concentrations of PCR reagents per reaction were as follows: 1X Phusion HF buffer, 0.02U/µl Phusion High-Fidelity Polymerase, 200µM each dNTPs. The following thermal cycling conditions were used; 98˚C for 2 min followed by 18 cycles of 98˚C for 10s, 60˚C for 30s and 72˚C for 8 min, followed by a final elongation step at 72˚C for 5 min. The total volume of PCR product (50µl) was digested with DpnI as described in section 2.6.2, however 10 units of enzyme were used and the reaction was incubated for 1 hour.

Secondly, we reduced the recommended amount of plasmid template DNA to 50ng and PCR denaturation and annealing cycling steps increased to 30s and 50s

respectively, and finally, the total volume of PCR product (50µl) was digested with DpnI (section 2.6.2) with 20 units of enzyme per reaction for 2 hours. Other reaction conditions remained the same as described above.

A volume of 5µl of the RF-cloning product was used to transform XL-10 Gold Ultra-Competent Cells (section 2.6.1). Bacterial colonies were screened for inserts by colony PCR. PCR amplification of DNA in colonies was performed following the Supertherm *Taq* polymerase protocol. The pRGH_LTR_FP forward and pRGH Colony RP reverse primers (Table 2A.1, Appendix 2) were designed to the *nef* gene and a region within the plasmid downstream of the 3'LTR, respectively. Cycling conditions, colony PCR screening, plasmid DNA isolation and gel electrophoresis protocols used were as described (section 2.4 and 2.6.1). To confirm insertion of the participant LTR 500ng of the plasmid DNA was restriction digested with 5 units of FastDigest HindIII for 1 hour at 37˚C following the Thermo Scientific FastDigest manufacturer restriction enzyme protocol. Digested products were analysed by agarose gel electrophoresis (section 2.4). A schematic of the bacterial transformation and clone screening procedures are shown below (Figure 2.2).

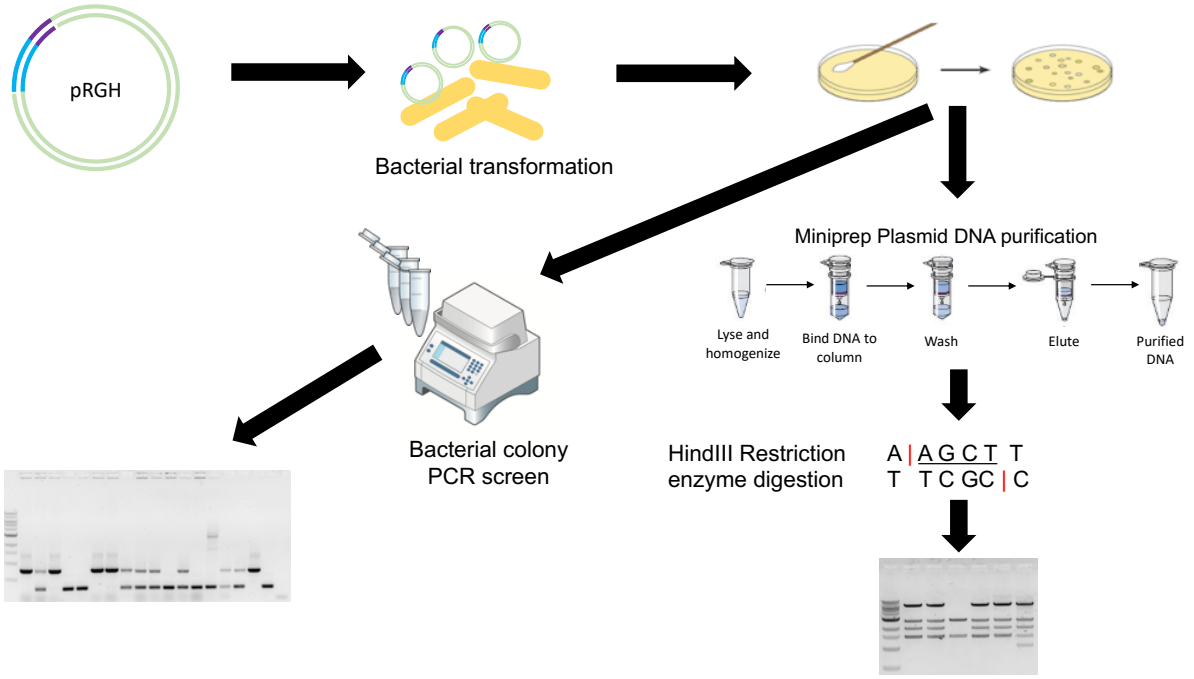


Figure 2.2 Screening for clones with participant LTRs incorporated into the pRGH-ΔU3 plasmid after RF-cloning. The intact nicked dsDNA molecule with the LTR of interest incorporated is subsequently transformed into competent bacteria and plated on LB agar containing carbenicillin and incubated at 32˚C overnight. Bacterial colonies were inoculated directly into PCR master mixes and also into LB broth containing carbenicillin and incubated at 32˚C overnight. PCR reactions were run with primers flanking the target site of interest and products run on an agarose gel to screen clones for the presence of a correctly sized DNA insert (LTR). Plasmid DNA was extracted the overnight cultures, digested with HindIII and run on an agarose gel to screen for clones. Figure elements adapted from the QIAGEN QIAprep® Miniprep Handbook obtained at: https://www.qiagen.com/us/resources/resourcedetail.

Successful clones were confirmed by Sanger sequencing (section 2.3) using the pRGH_LTR_F forward and pRGH Colony RP reverse primers (Table 2A.1, Appendix 2).

2.8 Cell Culture

2.8.1 Adherent cell lines

Two adherent cells lines were used. The HEK293T cell line, isolated from human embryonic kidney cell line is highly transfectable and capable of producing high titre retroviruses from retrovirus vector DNA (American Type Culture Collection ATCC® CRL3216™).The following reagent was obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: TZM-bl from Dr. John C. Kappes, Dr. Xiaoyun

Wu and Tranzyme Inc(297–301). The TZM-bl cell line which is a HeLa cell line derivative, is an immortalized cell line derived from cervical cancer cells that expresses the CD4 and CCR5 cell surface receptors necessary for HIV-1 viral entry(302).

TZM-bl cells have also been stably transfected with a reporter plasmid harbouring the Firefly luciferase gene under the control of a Tat inducible HIV-1 LTR promoter. After entry, reverse transcription, integration and protein expression from HIV-1 viruses/pseudoviruses in the TZM-bl cell line, the Tat protein binds to the LTR in the reporter system resulting in the expression of luciferase. Measurement of this luciferase by the addition of a luciferin containing reagent (Bright-Glo™, Promega, WI, USA) to the infected cells, allows for the quantification of viral infectivity via measurement of chemiluminescence. Bright-Glo™ reagent facilitates the lysis of the TZM-bl cells and contains beetle luciferin, the substrate for Firefly luciferase which is converted into oxyluciferin and releases photons of light during the reaction (Figure 2.3).

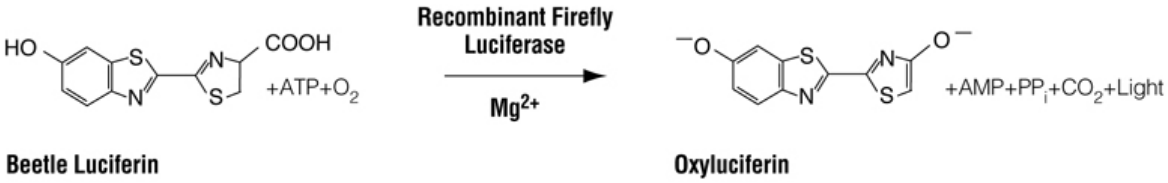


Figure 2.3 Schematic of the bioluminescent reaction catalysed by Firefly luciferase. Figure taken from the Bright-Glo™ Luciferase Assay System Technical Manual found at https://www.promega.com/-/media/files/resources/protocols/technical-manuals/0/bright-glo-luciferase-assay-system-protocol.pdf.

Both cell lines were maintained at 37˚C in 5% CO_2 (v/v) (hereinafter referred to as standard incubation conditions), in complete Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% heat-inactivated foetal bovine serum (FBS) (Thermo Scientific), 50µg/ml gentamicin antibiotic (Lonza). This media was referred to as D10+. Cells were grown in a humidified incubator in T75 filter-cap tissue culture flasks and passaged approximately every 3 days or when cell confluency exceeded 80%. A number of studies have shown that passaging cells many times can lead to changes in cell morphology, growth characteristics, protein expression and may negatively affect transfection efficiency. Therefore, cells were not used after passage 30 and were instead discarded.

HEK293T cells were passaged as follows: Media was removed from the culture flask leaving the cell monolayer intact. Cells were washed with 8ml 1X Dulbecco's Phosphate-Buffered Saline (DPBS) lacking calcium (Ca^{2+}) and magnesium (Mg^{2+}) ions as these inhibit Trypsin enzymatic function and are contained within D10+ media. Trypsin is a serine protease, which cleaves peptide chains at the C-termini of lysine or arginine amino acid residues, which are responsible for cell adhesion to the culture flask. Ethylenediaminetetraacetic acid (EDTA) is a chelating agent of divalent cations and therefore enhances Trypsin efficiency. A volume of 1ml of 1X Trypsin-EDTA (0.25% Trypsin, EDTA 1mM) was added to the flask and incubated for 1 minute at room temperature. 10ml of D10+ was then added to the flask to inactivate the trypsin and resuspend cells. To count and determine percentage viability of cells, a 0.4% Trypan Blue Stain (Invitrogen, CA, USA) was performed. This is an azo dye and works on the principle that dead cells have compromised cell membranes and therefore take up the dye whereas live cells with intact membranes do not. The staining procedure was performed as follows: A 1:2 dilution of cells was made with Trypan Blue, 10µl of this mixture was added to a Countess cell counting chamber (Invitrogen) slide and cell count was determined using a Countess automated cell counter. A total of 5 x 10^5 to 1 x 10^6 cells were seeded into each new flask in a final volume of 12ml of D10+ media.

TZM-bl cells were passaged following the same procedure with the following exception; 3ml of Trypsin-EDTA was added to the flask after the monolayer was washed with 1X DPBS. Cells were then incubated for 1 minute at room temperature after which 1ml of Trypsin-EDTA was removed from the flask. The flask was then incubated at 37˚C for an additional 5 minutes before the addition of D10+ media. Addition of media inactivates the activity of the trypsin.

2.8.2 Suspension cell lines

The following reagent was obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: Jurkat Clone E6-1 from Dr. Arthur Weiss. The Jurkat cell line is an immortalized cell line established from the peripheral blood of a 14-year-old patient with T cell leukaemia. The specific Jurkat E6-1 clone utilised in this study is a CD4+

suspension T cell line which expresses high levels of IL-2 upon stimulation with both phorbol esters and either monoclonal antibodies against to the T3 antigen or lectins(303).

The following reagent was obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: CEM.NK^{R} CCR5+ Cells from Dr. Alexandra Trkola(304–306). The CEM.NK^{R} CCR5+ cell line is a Human T-lymphoblastoid cell line that is a CEM CD4+ suspension cell line variant resistant to natural killer (NK) cell killing.

Jurkat E6-1 and CEM.NK^{R} CCR5+ cells were maintained under standard incubation conditions, in complete Roswell Park Memorial Institute (RPMI-1640) (Lonza, Basel, Switzerland) media supplemented with 10% heat-inactivated FBS (Thermo Scientific) and 50µg/ml gentamicin antibiotic. This media was referred to as R10+.

Jurkat E6-1 and CEM.NK^{R} CCR5+ cells were passaged as follows: Media containing cells was removed from the culture flask, transferred to a 15ml conical flask and centrifuged at 887 x g for 8 minutes at room temperature. The supernatant was subsequently discarded and the cell pellet was resuspended in 4ml of R10+ media. Cells were counted as described for the HEK293T cell line. An aliquot of the resuspended cells was then seeded into a new T25 cell culture flask at a concentration of 2.5 x 10^{5} live cells/ml of R10+ media in a total volume of 8ml. This procedure was carried out every 72-96 hours or when cell confluency exceeded 3 x 10^{6} cells/ml.

2.9 Generation of pseudovirus stocks

To produce the RGH pseudoviruses, HEK293T cells were co-transfected with the pHEF-VSVg envelope protein expressing plasmid and each pRGH LTR clone or one of the pRGH-WT, pRGH-ΔeGFP and pRGH-ΔCMV/ΔmCherry plasmids. The Lipofectamine 3000 (Thermo Scientific) transfection reagent was used for transfection of the cells. Lipofectamine 3000 is a positively charged cationic lipid that interacts with the phosphate backbone of DNA molecules mediating interaction of the DNA and the cell membrane of target cells. The Lipofectamine/DNA complex therefore fuses with the negatively charged cell membrane facilitating entry via endocytosis(307–309).

This transfection reagent allows for high transfection efficiency, with multiple plasmids, and induces minimal cell death.

For each clone, 8µg of the pRGH plasmid and 2µg of the pHEF-VSVg plasmid (4:1) was mixed with 5µl of the P3000 Reagent and 12µl of Opti-MEM™-Reduced Serum Media (Gibco®, Thermo Scientific) in a 15ml conical tube. Opti-MEM™ is used as it maintains a more consistent pH when compared to D10+ and serum in media reduces transfection efficiency(310). A volume of 5µl of the Lipofectamine 3000 Reagent and 125µl of Opti-MEM™ was mixed in a separate 15ml conical tube.

The Lipofectamine mixture was then added to the plasmid mixture, mixed by vortexing and incubated at room temperature for 15 minutes to allow for Lipofectamine/DNA complexes to form. During incubation, HEK293T cells were seeded at 1.2 x 10^6/well in a 6-well plate in a total volume of 1ml of D10+ media. The Lipofectamine/plasmid mixtures were then directly added to separate wells and each well made up to a final volume of 2ml by the addition of D10+ media. Plates were incubated for 48 hours under standard incubation conditions. Pseudovirus stocks were harvested by centrifugation of culture medium at 887 x g for 8 minutes to remove cells/cell debris. 200µl aliquots of the pseudovirus containing supernatant were transferred to cryovials and stored at -80˚C.

Each pseudovirus underwent titration in three different cell lines to confirm functionality and the necessary viral volume for infection.

2.9.1 TZM-bl titration

Pseudovirus stocks were titrated on TZM-bl cells to evaluate/screen for functionality. A 50% Tissue Culture Infectious Dose ($TCID_{50}$) assay was performed to measure the amount of virus required to produce a cytopathic effect (CPE) in 50% of TZM-bl cells inoculated with serially diluted viral stocks. Each viral stock underwent a 1:5 serial dilution in D10+ media containing Diethylaminoethyl (DEAE)-dextran at a concentration of 40µg/ml. DEAE-dextran is a polycation that stabilises viral adsorption to the cell membrane of TZM-bl cells thereby promoting cell-free viral uptake. Dilutions

were performed in duplicate over the first seven wells in a 96-well tissue culture microtiter plate in a final volume of 100µl. Eight background control wells were included per plate (eighth row) containing only 100µl of D10+ media with no virus. TZM-bl cells were then seeded into to each well at $1x10^4$ cells in 100µl of D10+ such that the final volume in each well containing cells and virus was 200µl with a final DEAE-dextran concentration of 20µg/ml.

Following a 48-hour incubation under standard incubation conditions approximately 80µl (to account for evaporation) of supernatant was removed from each well such that 100µl of cells/media/virus remained. 100µl of BrightGlo reagent (Promega) was added and the plate was incubated at room temperature in the dark for 2 minutes. The contents of each well were thoroughly mixed by pipette action and 150µl were transferred to an opaque black 96-well microtiter plate. Luminescence readings were then acquired as relative light units (RLU) in a GLOMAX® 96-Microplate Luminometer using the Promega BrightGlo™ Protocol; Promega, WI, USA. Viral titres (infectious units of virus per ml) were calculated following the Reed-Muench method(311).

2.9.2 Jurkat E6-1 & CEM.NK^R CCR5+ Titration

It was recommended that 10-20% eGFP expression one day post infection should yield infection with only single viral integrations within cells (Marcel Ooms, personal communication). To determine the necessary volume required to infect cells such that 10-20% of cells expressed eGFP one day post-infection, Jurkat E6-1 or CEM.NK^R CCR5+ cells were infected with two titres of each viral stock. A total of 1.0×10^6 Jurkat E6-1 cells per well in a 6-well tissue culture plate were infected with either 50µl or 150µl of each viral supernatant, in a final volume of 2ml of D10+ media supplemented with 4µg/ml Polybrene. Polybrene is a cationic polymer which increases viral infection by enhancing adsorption of virus to target cells membranes in a receptor independent manner(312). Cells were then subjected to spinoculation at 500 x g for 90 minutes at room temperature. Plates were finally incubated for 24 hours under standard incubation conditions.

A total of 5.0 x 10^5 CEM.NK^R CCR5+ cells per well in a 12-well tissue culture plate were infected with either 100µl or 250µl of each viral supernatant, in a final volume of 2ml of D10+ media supplemented with 4µg/ml Polybrene. Cells were then subjected to spinoculation at 1000 x g for 90 minutes at room temperature. Plates were finally incubated for 24 hours under standard incubation conditions.

For both cell lines eGFP expression on day one post-infection at different viral volumes was used to extrapolate the approximate viral titre required to obtain 10-20% eGFP expression assuming a linear relationship between viral titre and eGFP expression.

2.9.3 Confirmation of 5'LTR replacement

To confirm that the LTR cloned into the 3'LTR position of pRGH was successfully regenerated at the 5'LTR position upstream of the *gag*-eGFP cassette, the 5'LTR was amplified and sequenced from the integrated proviral within genomic DNA of cells infected with each pseudovirus. Jurkat E6-1 cells were infected as described in section 2.9.2, however to increase the frequency of integration events without inducing high levels of cell death, a volume of pseudovirus that yielded approximately 50% eGFP expression one day post-infection (based on data obtained from experiments in section 2.9.2) was used. Cells were maintained under standard incubation conditions over a period of 4 days to allow for integration to take place. A volume of 500µl of culture media was removed and replaced with fresh media after 48 hours. On day 4 post-infection, genomic DNA was extracted from infected cells using the QIAamp DNA Mini and Blood Mini Kit, following the DNA Purification from Blood or Body Fluids Spin Protocol. DNA was extracted from a total of approximately 4 x 10^6 Jurkat E6-1 cells and purified DNA was eluted in 50µl of AE buffer.

The integrated proviral 5'LTR sequence was amplified from purified genomic DNA via PCR following the Phusion High-Fidelity PCR Kit protocol. Primers were designed to a region complimentary to the start of the LTR and to the matrix region of *gag*. The cLTR_F forward and SQ16RC reverse primers were used (Table A2.1, Appendix 2). Final concentrations of reagents were as described in section 2.6.2. The following thermal cycling conditions were used; 98˚C for 2 min followed by 35 cycles of 98˚C for

30 sec, 62˚C for 30 sec and 72˚C for 30 min, followed by a final elongation step at 72˚C for 10 min. The PCR products were run on an agarose gel as described in section 2.4. Amplicons identified as approximately 1010bp in size, corresponding to the amplified proviral LTR, were excised from the gel and purified following the QIAquick Gel Extraction Kit Protocol. The purified DNA amplicon was Sanger sequenced (section 2.3) using the cLTR_F forward and SQ16RC reverse primers (Table A2.1, Appendix 2).

2.10 eGFP profile over time

To determine which day post-infection fluorescent protein expression stabilised and latency potential should be measured most accurately, cells were infected with pRGH pseudovirus and eGFP expression was measured over time. Jurkat E6-1 cells were infected (section 2.9.2) with the pRGH-WT pseudovirus at viral supernatant volumes ranging from 0-1000µl. eGFP expression was measured one, two, three, four, seven and eight days post-infection for each sample where 500µl of culture was prepared for flow cytometric analysis (section 2.11.3) on each day post-infection with replacement of 500µl of R10+ media. CEM.NK^{R} CCR5+ cells were also infected (section 2.9.2) with three participant LTR-pseudotyped RGH viruses and eGFP expression was measured for each infection over a six-day period where 500µl of culture was prepared for flow cytometric analysis (section 2.11.3) on each day post-infection with replacement of 500µl of R10+ media.

2.11 Measurement of Latency Potential

2.11.1 LTR panel infection

Jurkat E6-1 cells were infected with each pseudovirus to result in 10-20% of cells expressing eGFP on day one post-infection (section 2.9.2). Cells were maintained in culture under standard incubation conditions over an eight-day period where 1ml of culture media was removed and replaced with fresh media every other day until day

six post infection. On day seven post infection, each infected cell culture sample was divided equally into two subsets and transferred to new 6-well culture plates. One subset was treated with phorbol 12-myristate 13-acetate (PMA) and Ionomycin (Iono) at final concentrations of 4ng/ml and 1µM dissolved in Dimethyl sulfoxide (DMSO), respectively (stimulated). The other subset was treated with the equivalent volume of DMSO (unstimulated). PMA and Iono are compounds capable of inducing T cell activation via the activation of the PKC cascade. All wells for both subsets were made up to a final volume of 2ml. On day eight, all samples underwent flow cytometric analysis (section 2.11.1). This eight-day assay was performed in triplicate on three separate occasions.

CEM.NK^{R} CCR5+ cells were infected as previously described (section 2.9.2). Cells were maintained as above, however, 500µl of culture media was removed and replaced every second day, and on day seven post infection cell culture samples were divided and transferred to 12-well culture plates and all wells were made up to 1ml after addition of treatment compounds. This assay was performed in triplicate on three separate occasions.

2.11.2 Compensation control pseudovirus infections

Jurkat E6-1 and CEM.NK^{R} CCR5+ cells were infected with each compensation control pseudovirus (section 2.9.2), pRGH-ΔCMV/ΔmCherry and pRGH-ΔeGFP, one and four days prior to flow cytometric acquisition of each latency potential assay (section 2.11), respectively. Cells were maintained in culture under standard incubation conditions. For the pRGH-ΔeGFP pseudovirus infected cells, 1ml or 500µl of culture media was removed and replaced with fresh media on day two post-infection for Jurkat E6-1 and CEM.NK^{R} CCR5+ cells, respectively. On day three post-infection, pRGH-ΔeGFP pseudovirus infected cells were treated with PMA/Iono (section 2.11) to stimulate high levels of expression of mCherry. On the corresponding day eight of each latency potential assay, these samples also underwent flow cytometric analysis.

2.11.3 Flow Cytometry Acquisition and Analysis

On the selected day post-infection, infected Jurkat-E6 and CEM.NK^{R} CCR5+ cells were pelleted by centrifugation at 887 x g for 8 minutes at room temperature and resuspended in 1% (v/v) paraformaldehyde and incubated at room temperature for 10 minutes to allow cell fixation prior to flow cytometric acquisition.

Flow cytometric data for this project was acquired on either a Becton Dickinson (BD) FACSCalibur™ using BD CellQuest™ Pro software for the GFP titrations or a BD LSRFortessa™ using FACSDIVA™ software for all other data acquisition. Data was analysed on FlowJo version 10.2 (Treestar Inc.)

2.11.4 Gating strategy

For all samples run on the FACSCalibur™, a total of 40 000 live events were recorded. Dead cells were excluded as determined by the forward scatter area vs. side scatter area profiles [FSC-A/SSC-A] of the cells. Cells were then gated on eGFP expression based on uninfected cell populations (Figure 2.4).

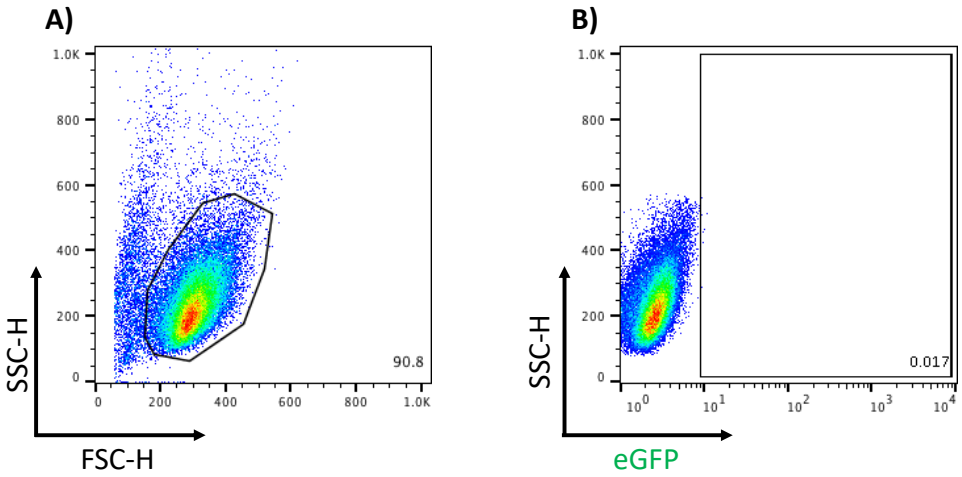


Figure 2.4 Gating strategy to identify the population of eGFP expressing cells. Gates are based on uninfected Jurkat E6-1 cells. Cells were gated on **A**) Live Cells and **B**) eGFP expression.

For all samples run on the LSRFortessa™, a total of 100 000 total events were recorded. Cells were first gated on time to ensure there were no shifts in fluorescence

intensity during the acquisition run. Dead cells were then excluded as determined by the forward scatter area vs. side scatter area profiles [FSC-A/SSC-A] of the cells. Doublet events were excluded as determined by the forward scatter area vs. forward scatter height profiles [FSC-A/FSC-H] of the cells. Finally, cells were gated for single or co-expression of eGFP and mCherry expression based on uninfected cell populations (Figure 2.5).

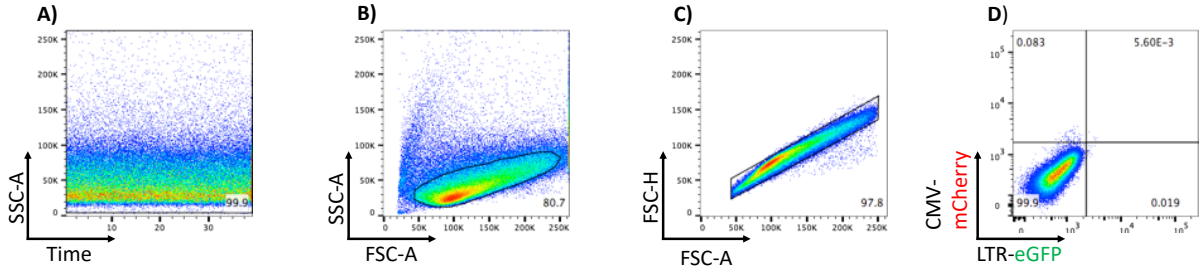


Figure 2.5 Gating strategy to identify populations of mCherry and eGFP expressing cells. Uninfected Jurkat E6-1 cells treated with DMSO for 24 hours. Cells were gated on Time **A**), Live Cells **B**), Singlet events **C**) and finally mCherry and eGFP expression **D**)

Latency potential was defined as the ratio of latently infected cells to actively infected cells and was thus calculated as follows:

$$Latency\ Potential = \frac{Proportion\ of\ mCherry\ only\ expressing\ cells}{Proportion\ of\ mCherry\ \&\ eGFP\ co\text{-}expressing\ cells}$$

2.12 Measurement of LTR activity

The panel of LTR cloned pGL4.10 plasmids described in section 2.7 was used to measure the basal and Tat-induced expression of each participant LTR for which latency potential had been determined in section 2.11.

HEK293T cells were co-transfected with 100ng of the each of the pGL4.10 LTR plasmids, 25ng of the pRL-SV40 vector and a range of concentrations of the pcDNA™3.1/V5-His-TOPO vector, with the CAP063 *tat* gene cloned into the multiple cloning site (Tat-TOPO vector, D.R. Chopera, UCT), (0, 10, 25, 50 and 100ng). The

tat sequence was cloned from a subtype C infected participant who is a rapid progressor in the CAPRISA 002 cohort(313).

PolyFect transfection reagent (QIAGEN) was used for transfection of these cells. For each participant, the pGL4.10, pRL-SV40 and Tat-TOPO vectors were diluted in D10+ media, mixed and added to each of the corresponding wells of a 96-well plate. PolyFect was added at a concentration of 40µg/ml per well in a total volume of 100µl (PolyFect and vector mixes). Plates were incubated at room temperature for 10 minutes to allow for PolyFect-DNA complexes to form. HEK293T cells were then seeded at 2.0 x 10^4 cells per well, in 100µl of D10+ media, into each well of the 96-well plate such that each well was made up to 200µl containing PolyFect at a final concentration of 20µg/ml. Transfections with a HIV-1 Subtype B BaL LTR sequence cloned into the pGL4.10 vector and a pGL4.10 with no LTR cloned in at each Tat-TOPO concentration were included as normalisation controls. Each transfection reaction was set up in triplicate. Plates were incubated for 48 hours under standard incubation conditions.

Following incubation, 120µl of supernatant was removed from each well such that 80µl of culture media remained. A volume of 75µl of Dual-Glo® Luciferase Assay Reagent (Promega), containing beetle luciferin, was added to each well of the plate and the plate was incubated at room temperature in the dark for 10 minutes. The contents of each well were thoroughly mixed by pipette action and 100µl were transferred to an opaque black 96-well microtiter plate. Firefly luminescence readings were then obtained as relative light units (RLU) in a GLOMAX® 96-Microplate Luminometer using the Promega Dual-Glo Protocol; Promega, WI, USA. A volume of 75µl of Dual-Glo® Stop & Glo® Reagent (Promega), containing coelenterazine, was added to each well of the 96-well plate and the plate was incubated at room temperature in the dark for a further 10 minutes.

Renilla luminescence readings were then obtained as relative light units (RLU) as described above. Transfection efficiency was controlled for by normalising luminescence readings against *Renilla* luciferase expression. Luciferase reactions are shown in Figure 2.6. LTR activity was calculated as a response ratio relative to the activity of the BaL isolate LTR sequence (provided by P Selhorst, UCT), to provide a method to standardise readings and control for inter-experiment variation, and expressed as a factor of BaL activity(314). Therefore, LTR activity for each participant LTR was calculated using the following equation (developed by P. Selhorst, UCT):

$$LTR\ Activity = \frac{\left(\frac{FLL\ (Participant\ LTR)}{RLL\ (Participant\ LTR)} - \frac{FLL\ (NC)}{RLL\ (NC)}\right)}{\left(\frac{FLL(BaL\ LTR)}{RLL\ (BaL\ LTR)} - \frac{FLL\ (NC)}{RLL\ (NC)}\right)}$$

FLL: Firefly luciferase luminescence

RLL: *Renilla* luciferase luminescence

NC: Negative control (pGL4.10 with no LTR)

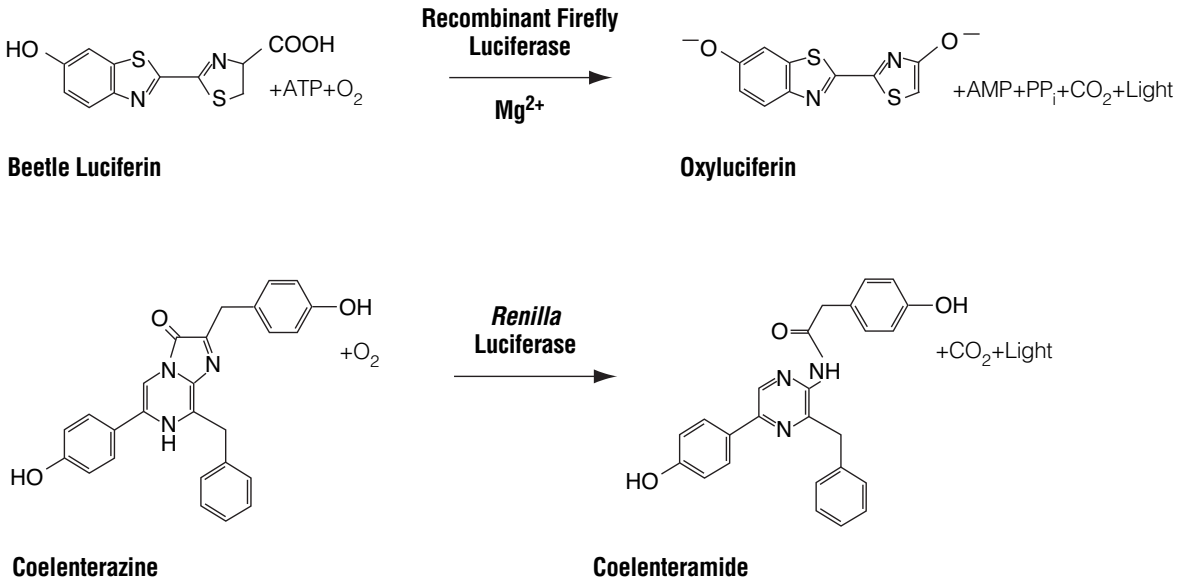


Figure 2.6 Bioluminescent reactions catalyzed by Firefly and *Renilla* luciferases. Beetle luciferin undergoes mono-oxygenation, catalyzed by Firefly luciferase, to oxyluciferin and light. Coelenterazine undergoes mono-oxygenation, catalyzed by *Renilla* luciferase, to coelenteramide and light. Bright-Glo™ Luciferase Assay System Technical Manual found at: https://www.promega.com/-/media/files/resources/protocols/technical-manuals/0/dual-glo-luciferase-assay-system-protocol.pdf.

2.13 Statistical analysis

Graphs were generated and statistical analysis performed in GraphPad Prism 7.0a (GraphPad Software, La Jolla California USA, www.graphpad.com). Spearman correlation (CI=95%) tests were performed for all correlations due to the small sample sizes. Statistical significance was calculated using the Wilcoxon matched-pairs signed rank test and a p-value of <0.05 was considered significant. Where multiple comparisons were made, a one-way analysis of variance (ANOVA), followed by a Tukey multiple comparison test was performed and a p-value of <0.05 was considered significant.

Chapter 3: Results

Introduction

This study investigated LTR latency potential and promoter activity of the HIV-1 LTR DNA sequence isolated from recently infected participants from the CAPRISA 004 cohort (estimated duration of infection: median 5 weeks, range 3-12) (acute/early infection) (section 2.2)(287). LTR clones and sequences (n=26) (HXB2 positions 9086-9621) were available from acute/early time points (provided by P. Selhorst, UCT), and in addition, a 330bp U3 region sequence (HXB2 position 9086-9415) generated directly from plasma virus from acute/early infection and 12 months post-infection respectively was available for all participants (Chopera *et al.*)(286).

3.1 Selection of a subtype C LTR panel

To determine whether the LTR sequences were representative of subtype C South African viruses, phylogenetic analysis was performed. The 26 acute/early LTR clone sequences were compared to 50 subtype C South African sequences and to each subtype, A through G, consensus LTR sequence (obtained from the Los Alamos National Laboratories HIV Sequence Database http://www.hiv.lanl.gov/) (Figure 3.1). The participant LTR sequences clustered exclusively with the subtype C sequences and were dispersed amongst South African subtype C sequences. This confirmed that they were all subtype C, that there were no recent phylogenetic linkages between selected sequences, and they were representative of South African subtype C viruses.

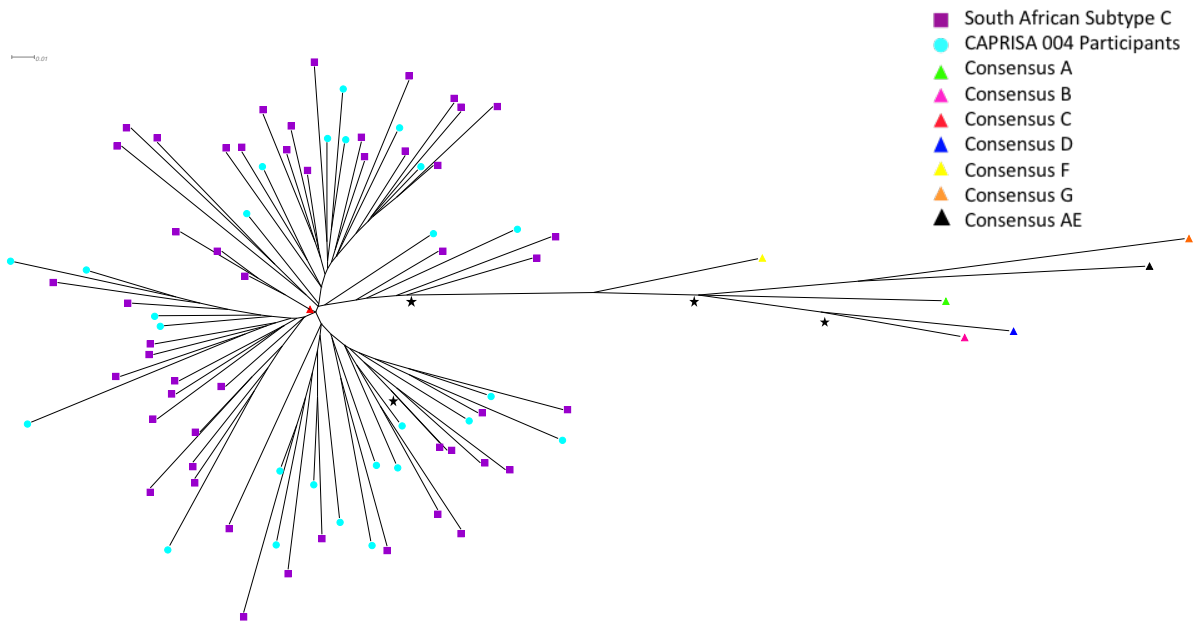


Figure 3.1 Subtype diversity of the LTR. The evolutionary history was inferred using the Maximum Likelihood method based on the Tamura-Nei model(315). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The panel of 26 selected CAPRISA 004 LTR sequences, a selection of 50 South African subtype C sequences and a consensus sequence for each subtype was included. Different sequences are identified by coloured shapes detailed in the key. A total of 451 nucleotide positions were included in the final dataset, covering the U3 and R regions of the LTR. Scale bar (top left of figure) = 0.01. Black stars represent nodes at which bootstrap support was ≥80.

The LTR clones originated from plasma virus, cultured in PBMCs over a median of 25 days (range 25-58) (Selhorst *et al.*)(285). The HIV-1 viral replication cycle takes approximately 2 days, and the viral reverse transcriptase enzyme has an inherently low fidelity and is prone to introducing mutations during replication(316, 317). There is therefore potential for viral evolution during culture. We thus investigated whether the cloned LTR sequences represented the *in vivo* plasma viruses from these women at a matched time point using DNA distance analyses. The mean intra-participant pairwise DNA distance between the cloned sequence and matched acute plasma virus sequence was 0,0036 (range 0-0.032) (Figure 3.2). One participant, CAP352, had a higher than expected DNA distance for intra-participant quasispecies diversity (0.032).

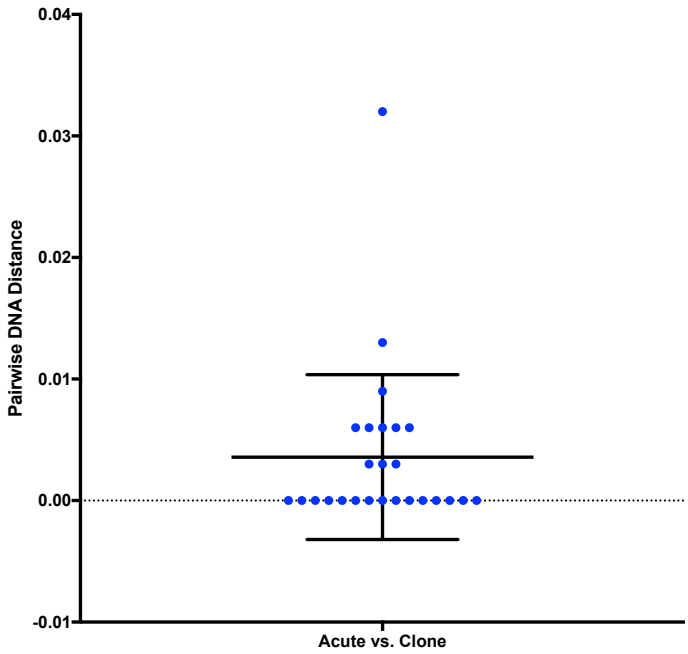


Figure 3.2 Scatter plot of pairwise DNA distances between the acute and matched pGL4.10 cloned matched LTR sequences. The number of base substitutions per site between sequences are shown. Analyses were conducted using the Maximum Composite Likelihood model(318). A total of 329 nucleotide positions were analysed. The central bar shows the mean and the error bars represent the standard deviation of the mean.

Furthermore, to investigate whether the difference between the proviral and plasma sequence was greater than we would expect for intra-person variation, we generated a maximum likelihood phylogenetic tree to compare cloned LTR and acute/early infection sequences to those generated from 12 months post-infection (Figure 3.3). Participant sequences grouped together, confirming their identity. Of the 26 participants: all three sequences from different time points/sources were identical for six participants; the cloned sequence was identical to or grouped more closely with the acute sequence for 13 participants; and in seven participants there was higher diversity between the cloned and acute sequences compared to the acute and 12-month post-infection sequences. Of the latter seven, cloned sequences were equidistant from the acute and 12-month post-infection virus for three participants.

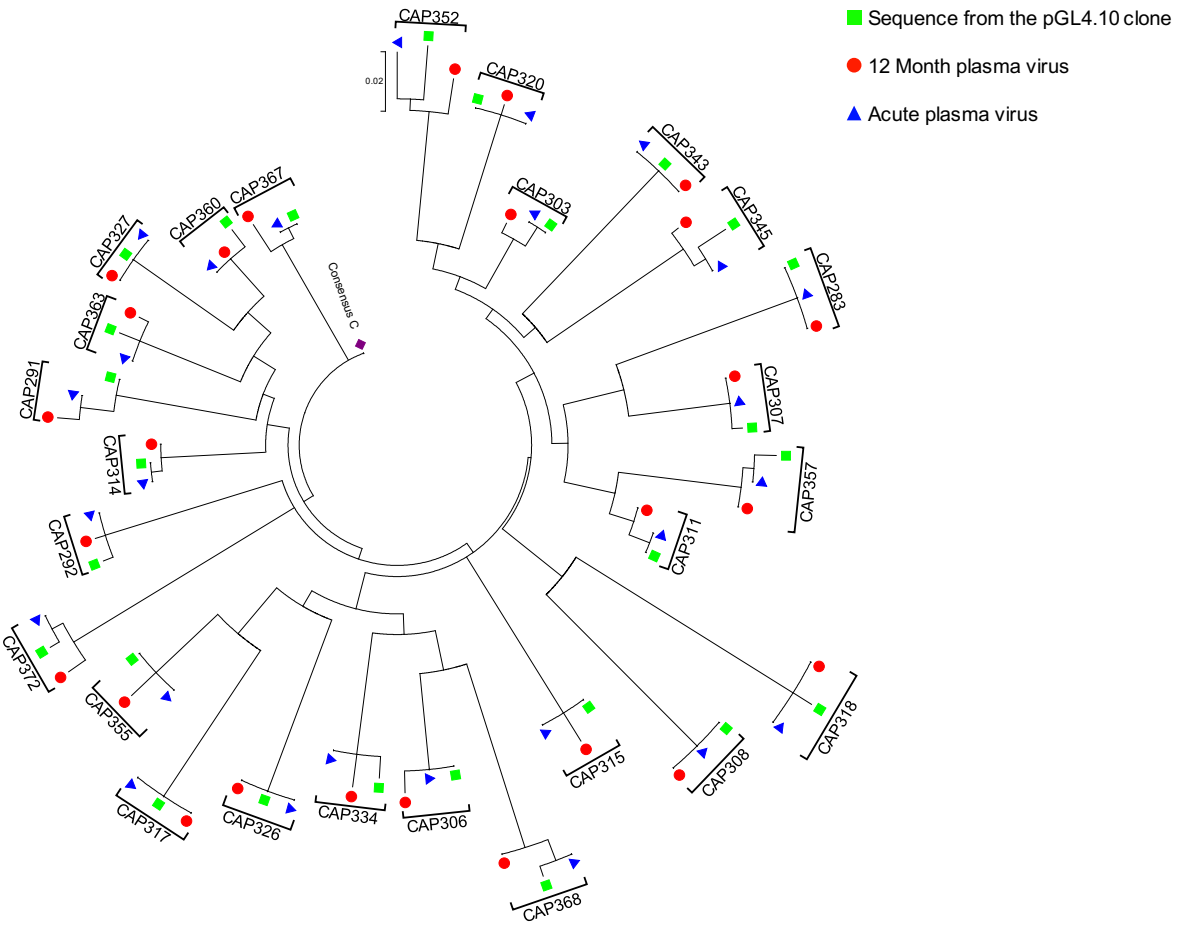


Figure 3.3 Phylogenetic tree of CAPRISA 004 LTR sequences. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model(315). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 76 nucleotide sequences. A panel of 26 CAPRISA 004 participants were included with sequences obtained during acute infection (at participant enrolment) (blue triangles), at chronic infection (at 12 months post-infection) (red circles), and sequences from the panel of pGL4.10 cloned LTRs (green squares). There were a total of 322 positions in the final dataset including the entire overlapping *nef*/U3 region of the LTR. The tree was rooted on the global Consensus C LTR sequence (purple diamond). Scale bar (top centre of figure) = 0.02.

Participant LTR sequences that were a 100% match or were more similar to the corresponding acute plasma sequence than to the 12 months post-infection plasma virus sequence were selected for further investigation (n=18). One additional participant (CAP360) was included who was identified as being infected by multiple transmitted variants of HIV-1(288). Therefore, based on the results above, a total of 19 participant-derived LTR clones were utilized in further experiments (Table A.1, Appendix 3).

3.2 Generation of a panel of LTR-pseudotyped RGH viruses

The effect of the LTR genotype on viral latency was evaluated using an *in vitro* model where the participant-derived LTRs were cloned into the pRGH-ΔU3 vector (Figure A2.3, Appendix 2). This plasmid harbours a *gag*-eGFP gene under the control of the LTR and an mCherry gene under the control of a constitutive CMV promoter (see section 1.8.2.3). This construct was used to generate chimeric HIV-1 pseudoviruses. Following infection, simultaneous detection of eGFP and mCherry was indicative of active viral expression, whereas expression of mCherry alone indicated viral integration but no HIV-1 expression (latent infection). Therefore, pRGH allowed for the simultaneous detection of actively and latently HIV-1 infected CD4+ T cells.

3.2.1 Restriction-Free Cloning Optimisation

To assess the effect of genotypic variation within the panel of subtype C LTRs on the ability to establish latency, participant LTRs were cloned into the RGH plasmid. As the RGH plasmid lacked unique restriction sites flanking the desired insertion site, an RF-cloning strategy was implemented. RF-cloning is a PCR-based cloning procedure where sequences to be cloned are incorporated during the PCR reaction via flanking sequences complimentary to the target plasmid.

Participant LTRs were amplified from the pGL4.10 clone panel and an approximately 650bp DNA fragment (HXB2 positions 9014-9621) corresponding to the U3, R and some flanking sequence was generated from all participants (Figure 3.4).

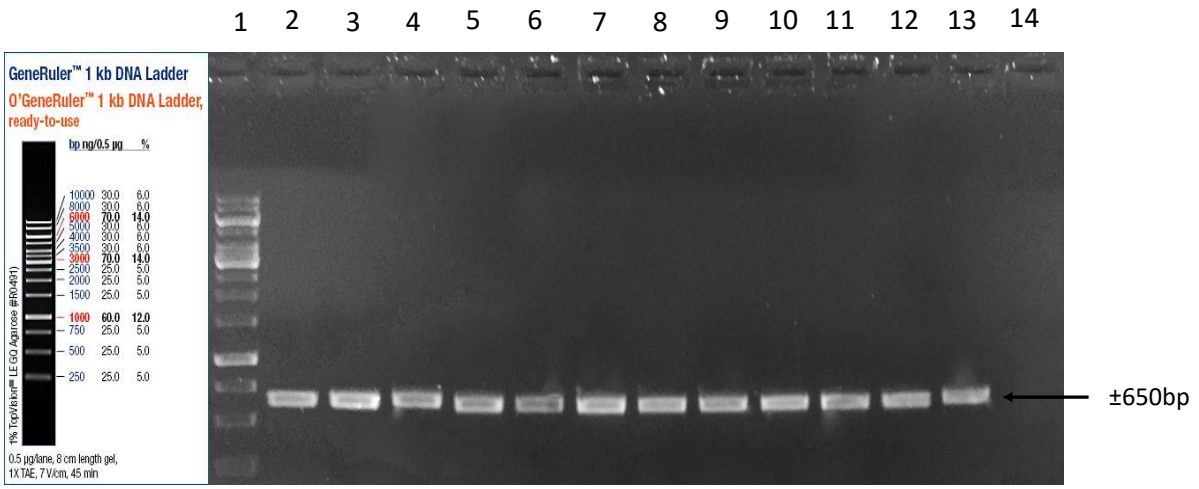


Figure 3.4 CAPRISA 004 participant LTRs amplified from the pGL4.10 panel. A 1% agarose gel of HIV-1 3'LTR including 72bp of pGL4.10 vector upstream of the LTR start site and the entire U3 and R region of the LTR producing a band of approximately 650bp. Lane 1: The O'GeneRuler^{TM} 1 kb DNA Ladder (Thermo Scientific), Lanes 2-13 PCR amplified LTR sequences from participants: CAP283, CAP303, CAP306, CAP311, CAP314, CAP315, CAP320, CAP326, CAP343, CAP360, CAP367, CAP372 PCR respectively, Lane 14 Negative PCR control (no template DNA added to reaction). Electrophoresis was performed for 40 minutes at 100V. DNA was stained with GelRed^{TM} and the gel was visualised under UV light.

The LTRs were cloned into the 3'LTR site of pRGH-ΔU3 (Figure A2.3, Appendix 2). During the viral replication cycle, the 3'LTR is transferred to the 5' end of the genome (see section 1.6.1) to generate the U3 and R region of the 5'LTR upstream of the *gag*-eGFP cassette. Due to the size of the target RGH-ΔU3 plasmid (approx. 13.5kb), the RF-cloning reaction was inefficient. Therefore, a number of conditions were tested in order to develop the most efficient cloning protocol possible, which yielded the most positive clones following bacterial transformation and colony screening (Table 3.1). Of three condition sets tested, condition set 3 was found to yield the highest colony positivity rate of between 10-25%. A total of seven different participant LTRs were used for this optimization. Increased DpnI and digestion time appeared to reduce the number of clones which produced the parent plasmid banding pattern.

Table 3.1 Optimsiation of conditions used for RF-cloning.

Condition	Condition Set 1	Condition Set 2	Condition Set 3
DNA polymerase	Phusion high fidelity	Phusion high fidelity	Phusion high fidelity
Denaturation time	10 seconds	30 seconds	30 seconds
Annealing time	30 seconds	50 seconds	50 seconds
Amount of insert	125ng	50ng	50ng
Amount of DpnI	10 units/reaction	10 units/reaction	20 units/reaction
DpnI digestion time	1 hour	1 hour	2 hours
Colony Positivity Rate	0%	<10%	10-25%

3.2.2 Confirmation of cloning and pseudovirus generation

Three methods were used to confirm insertion of the LTR into the correct position of pRGH: (i) DNA amplification, (ii) restriction enzyme mapping and (iii) Sanger sequencing. DNA amplification of bacterial colonies containing LTR inserts, using primers designed to flank the cloned LTR, produced a band of approximately 740bp, whereas unsuccessful clones produced a band of only 310bp or alternate patterns (Figure 3.5A). Following restriction enzyme mapping with HindIII, positive clones produced bands of approximately 6470bp, 2620bp, 1930bp and 1450bp when run on an agarose gel. An extra band of approximately 1020bp or missing bands indicated unsuccessful insertion of the participant LTR (Figure 3.5B). In order to obtain positive clones, up to 20 colonies in three RF-cloning reactions per participant were screened.

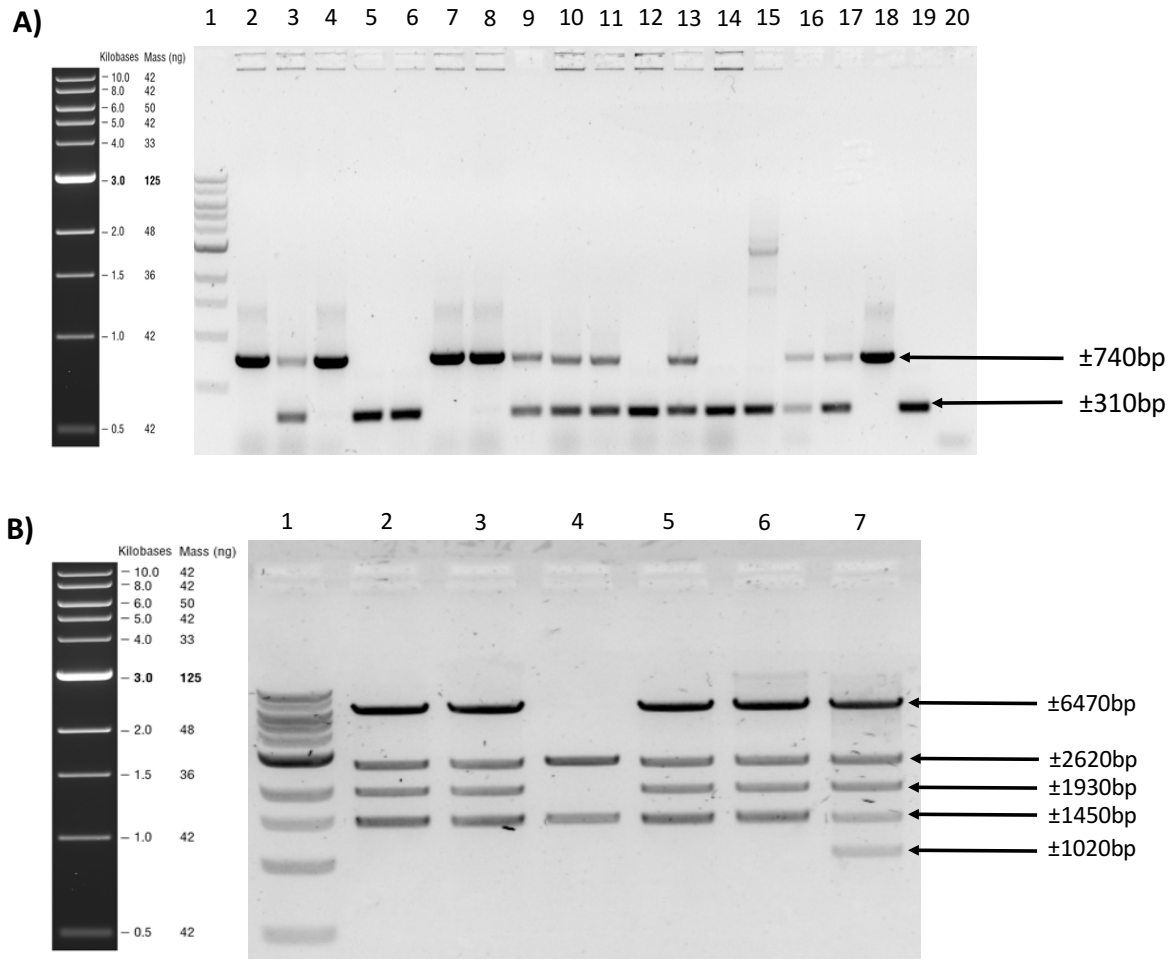


Figure 3.5 Screening for clones with participant LTRs incorporated into the 3'LTR position of the pRGH-ΔU3 plasmid after RF-cloning. A) 1% agarose gel of DNA amplified from bacterial colonies, transformed with pRGH-ΔU3 containing RF-cloned participant LTRs, which were directly inoculated into PCR master mixes. Screening primers were designed to flank the entire 3'LTR region, specifically, to the *nef* gene and a region downstream of the 3'U5 region within the plasmid. A band of approximately 740bp was expected in clones where the participant LTR had been inserted (same as the pRGH-WT) and a band of approximately 310bp was expected in clones with unsuccessful insertion (same as the pRGH-ΔU3). Clones displaying both bands were also considered as unsuccessful insertion. Lane 1: The 1kb DNA ladder (NEB), Lane 2-7: CAP314 LTR RF clone transformed colonies, Lane 8-17: CAP320 LTR RF clone transformed colonies, Lane 18: Amplified region from the pRGH-WT plasmid, Lane 19: Amplified region from pRGH-ΔU3, Lane 20: Negative PCR control (no template DNA/bacterial colony added to reaction). **B)** 1% agarose gel of HindIII digested plasmid DNA isolated from transformed colonies displaying the correct banding pattern in A). Bands of approximately 6470bp, 2620bp, 1930bp and 1450bp were expected from clones with successful insertion of the 3'LTR (same as the pRGH-WT) and an extra band of approximately 1020bp (same as pRGH-ΔU3) or missing bands indicated unsuccessful insertion of the participant LTR. Lane 1: The 1kb DNA ladder (NEB), Lane 2-5: Plasmid DNA isolated from colonies corresponding to lane 2, 4, 7 and 8 in A), Lane 6: Digested pRGH-WT plasmid DNA, Lane 7: Digested pRGH-ΔU3 plasmid DNA. For both gels electrophoresis was performed for 40 minutes at 100V. DNA was stained with GelRed™ and the gel was visualised under UV light.

We successfully cloned 11 of the 19 participant LTRs into the RGH-ΔU3 vector (Table A3.1, Appendix 3). Sequencing of the 11 clones confirmed that the inserted LTR

sequence matched the sequences amplified from the pGL4.10 panel. Pseudoviruses were subsequently generated from these clones.

3.3 Cloned LTRs remain intact after infection and integration

The generation of the LTR during replication is a multi-step, complex procedure which takes place prior to proviral integration (sections 1.6.1 and 3.2.1). Participant LTRs were cloned into the 3' LTR position of the RGH plasmid with the intention of the U3 and R regions of the participant LTR sequence being transferred to the 5' LTR position upstream of the *gag*-eGFP cassette during replication and integration. To confirm that the LTR was correctly regenerated, we sequenced LTRs of two participants, CAP303 and CAP311. To do this, genomic DNA was extracted from Jurkat E6-1 cells infected with LTR-pseudotyped viruses, CAP303 or CAP311. The integrated proviral 5'LTR sequence was amplified from this genomic DNA, and Sanger sequencing of these amplicons revealed that the LTR had been generated and that the sequence matched that of cloned participant LTRs.

3.4 Construction of flow cytometry compensation controls

As the read out of the latency assay was performed by flow cytometric detection of fluorescent proteins (section 3.6), pseudoviruses expressing each of the fluorescent proteins individually were generated to be used for compensation of the fluorescence spectral overlap (Figure A2.1 Appendix 2). To achieve this, pRGH-WT (Figure 1.5), was modified to express only mCherry or eGFP.

To produce an mCherry-only expressing pRGH-WT clone, the eGFP gene was deleted by inverse PCR. To confirm deletion of the eGFP gene, DNA was amplified from transformed bacterial colonies, with primers designed to flank the region of the eGFP gene in pRGH. An amplicon of approximately 720bp was expected in clones in which the gene had been successfully deleted (Lanes 2, 3, 6, 7 and 10, Figure 3.6A). DNA isolated from bacterial colonies which produced the correct sized band was restriction enzyme digested with HindIII. Clones were confirmed as having the eGFP deletion if bands of 6470bp, 2620bp, 1450bp and 1210bp were visualised following

agarose gel electrophoresis (Figure 3.6B). Clones with successful deletion were designated as pRGH-ΔeGFP.

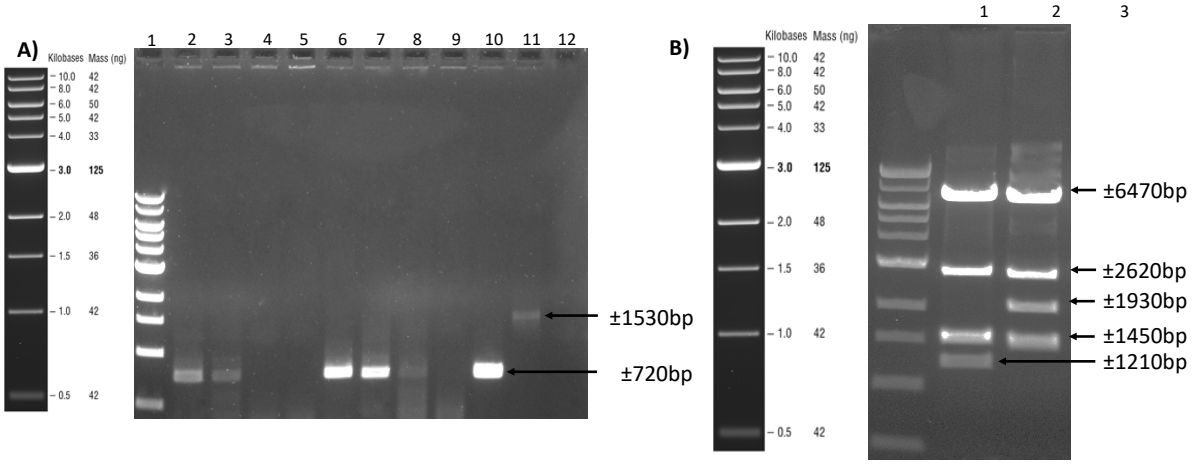


Figure 3.6 Screening for an mCherry only expressing pRGH plasmid. A) 1% agarose gel of DNA amplified from bacterial colonies, transformed with pRGH-WT plasmids after deletion of the eGFP gene (pRGH-ΔeGFP) by inverse PCR, that were directly inoculated into PCR master mixes. Screening primers were designed to flank the site of the eGFP gene within the pRGH-WT plasmid, specifically, to the p17 and p24 regions of *gag*. A band of approximately 720bp was expected in clones where the eGFP gene had been deleted and a band of approximately 1530bp was expected in clones where the deletion was unsuccessful. Lane 1: The 1kb DNA ladder (NEB), Lane 2-10: pRGH-ΔeGFP clone transformed colonies, Lane 11: Amplified region from the pRGH-WT plasmid, Lane 12: Negative PCR control (no template DNA/bacterial colony added to reaction). **B)** 1% agarose gel of HindIII digested plasmid DNA isolated from transformed colonies displaying the correct banding pattern in A). Bands of approximately 6470bp, 2620bp, 1450bp and 1210bp were expected in clones with successful deletion of the eGFP gene and bands of approximately 6470bp, 2620bp, 1930bp and 1450bp indicated unsuccessful deletion. Lane 1: The 1kb DNA ladder (NEB), Lane 2: Plasmid DNA isolated from the bacterial colony corresponding to lane 10 in A), Lane 3: pRGH-WT DNA. For both gels electrophoresis was performed for 40 minutes at 100V. DNA was stained with GelRed™ and the gel was visualised under UV light.

An eGFP-only expressing clone was created by deletion of the CMV/mCherry region by restriction enzyme digestion with BlpI and XhoI. A band of approximately 12470bp, corresponding to the remainder of the parent plasmid, was gel extracted following gel electrophoresis and subsequently re-ligated (Figure 3.7A). To confirm deletion, DNA was amplified, from transformed bacterial colonies, with primers designed to flank the CMV/mCherry region in pRGH. An amplicon of approximately 960bp was observed in clones in which the gene had been successfully deleted (Lane 3, Figure 3.7B). DNA isolated from bacterial colonies which produced the correct sized band was restriction enzyme digested with KpnI. Clones were confirmed as having the CMV/mCherry deleted if bands of 7300bp, 4830bp and 330bp were visualised following agarose gel

electrophoresis (Figure 3.7C). Clones with successful deletion were designated as pRGH-ΔCMV/ΔmCherry.

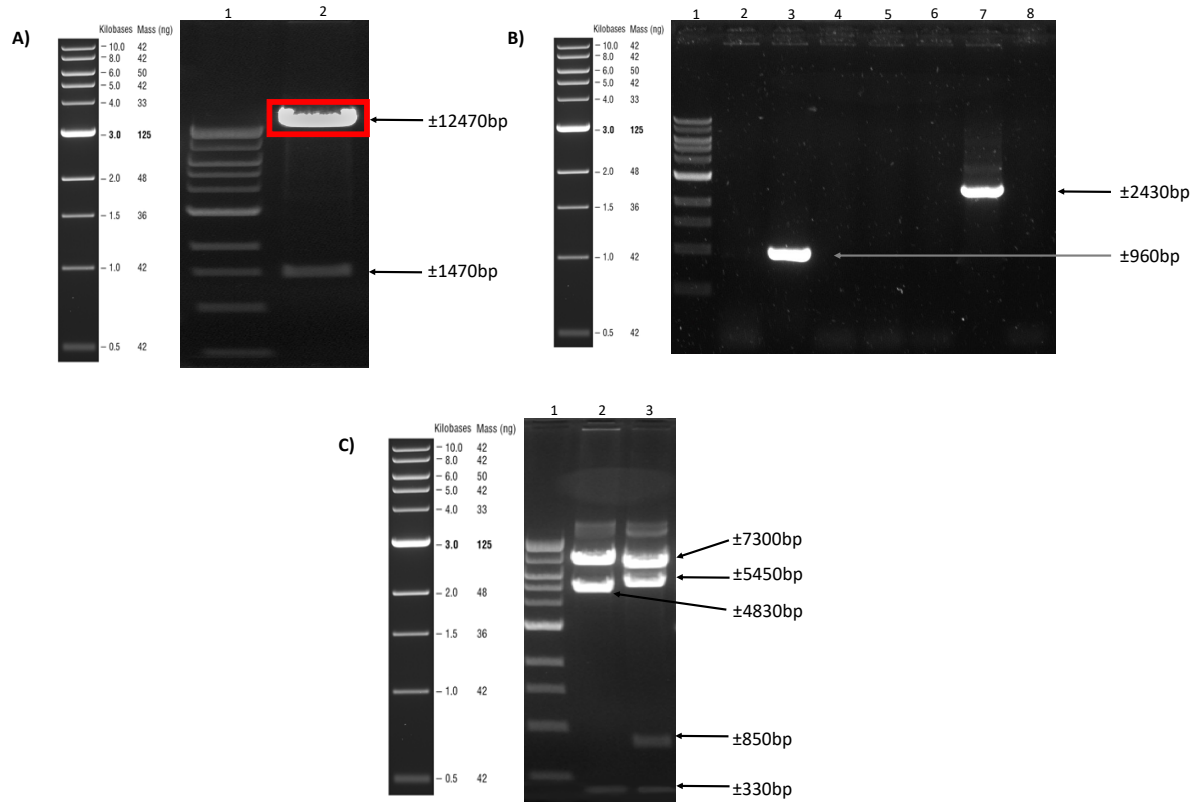


Figure 3.7 Generation and screening for an eGFP only expressing pRGH plasmid. A) 1% agarose gel of the pRGH-WT plasmid restriction digested with BlpI and XhoI to remove the CMV promoter and mCherry gene. A band of approximately 1470bp was expected containing the CMV/mCherry sequence, and a larger band of approximately 12470bp containing the remainder of the parent plasmid. The larger band (highlighted) underwent gel extraction/purification for subsequent ligation and transformation. **B)** 1% agarose gel of DNA amplified from bacterial colonies, transformed with re-ligated pRGH-WT plasmid after deletion of the CMV/mCherry sequence (pRGH-ΔCMV/ΔeGFP), that were directly inoculated into PCR master mixes. Screening primers were designed to flank the site of the CMV/mCherry region, specifically, to the *nef* region and a region downstream of the 3'U5 region within the plasmid. A band of approximately 960bp was expected in clones where the CMV/mCherry sequence had been deleted and a band of approximately 2430bp was expected in clones where the deletion was unsuccessful. Lane 1: The 1kb DNA ladder (NEB), Lane 2-5: pRGH-ΔCMV/ΔmCherry clone transformed colonies, Lane 6: Amplified region from the pRGH-WT plasmid, Lane 12: Negative PCR control (no template DNA/bacterial colony added to reaction). **C)** 1% agarose gel of KpnI digested plasmid DNA isolated from transformed colonies displaying the correct banding pattern in B). Bands of approximately 7300bp, 4830bp and 330bp were expected in clones with successful deletion of the CMV/mCherry sequence and bands of approximately 7300bp, 5450bp, 850bp and 330bp indicated unsuccessful deletion. Lane 1: The 1kb DNA ladder (NEB), Lane 2: Plasmid DNA isolated from the colony corresponding to lane 2 in A), Lane 3: pRGH-WT DNA. For all gels electrophoresis was performed for 40 minutes at 100V. DNA was stained with GelRed™ and the gel was visualised under UV light.

Sanger sequencing confirmed deletion of either of the reporter genes, and that clones maintained intact flanking regions. Pseudoviruses were generated and expression of either eGFP or mCherry was confirmed by flow cytometry (Figure 3.8).

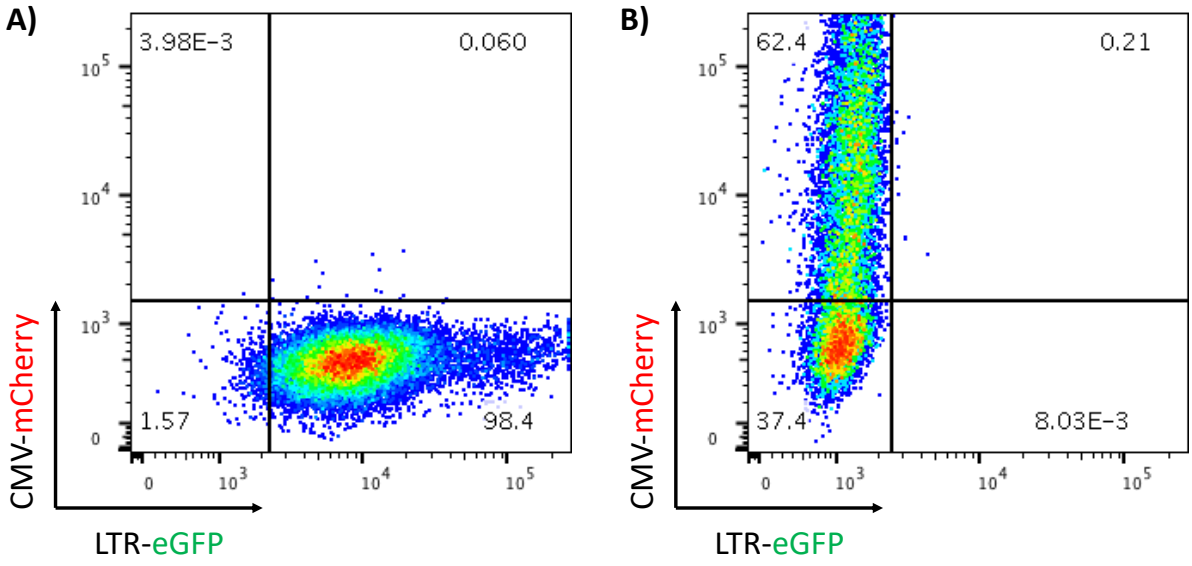


Figure 3.8 Single fluorescent protein expression of flow cytometry compensation control pseudoviruses. Pseudocolour dot plots of Jurkat E6-1 cells infected with **A)** pRGH-ΔCMV/ΔmCherry and **B)** pRGH-ΔeGFP pseudoviruses. Cells were infected and analysed by flow cytometry 1 and 4 days post-infection in plot A and B respectively. Cells infected with pRGH-ΔeGFP were stimulated with PMA/Iono on day 3 post-infection.

3.5 Optimization of RGH pseudovirus infection

Pseudoviruses were initially confirmed to be infectious by viral titration in TZM-bl cells with titres ranging from 1398-6988 $TCID_{50}$/ml. These pseudoviruses contain a Gag-eGFP fusion protein which can be detected following infection of CD4+ T cells by flow cytometry. Thus during early infection, a strong eGFP signal reflects cell-free virions and not LTR-driven expression and is therefore not an indication of eGFP integrated viruses. Furthermore, it was determined by Dahabieh *et al.*(156), that to obtain single-copy integrations per cell, a viral concentration/titre that yields an infection rate of approximately 10-20% eGFP expressing cells one day post-infection should be used. Obtaining single integration events was necessary as, in a scenario where more than

one provirus was present per cell, it would not be possible to determine if proviruses within the same cell were displaying the same levels of transcriptional activity.

Therefore, it was necessary to determine the length of time that this signal took to wane such that only fluorescent protein expressed from integrated provirus was measured. The viral titre required to yield the necessary infection rate was also determined for each virus.

3.5.1 Determination of time to stable state of latency

To determine the time taken for the initial eGFP signal to disappear and for stable expression of novel eGFP from integrated proviruses, Jurkat E6-1 cells were infected with a wild type RGH pseudovirus at volumes ranging from 50-1000µl and the eGFP signal was monitored over time by flow cytometry. As expected, there was an initial spike in eGFP at all volumes on day one post-infection. This signal gradually decreased over time and the signal appeared to stabilise between day four and eight post-infection depending on the volume (Figure 3.9).

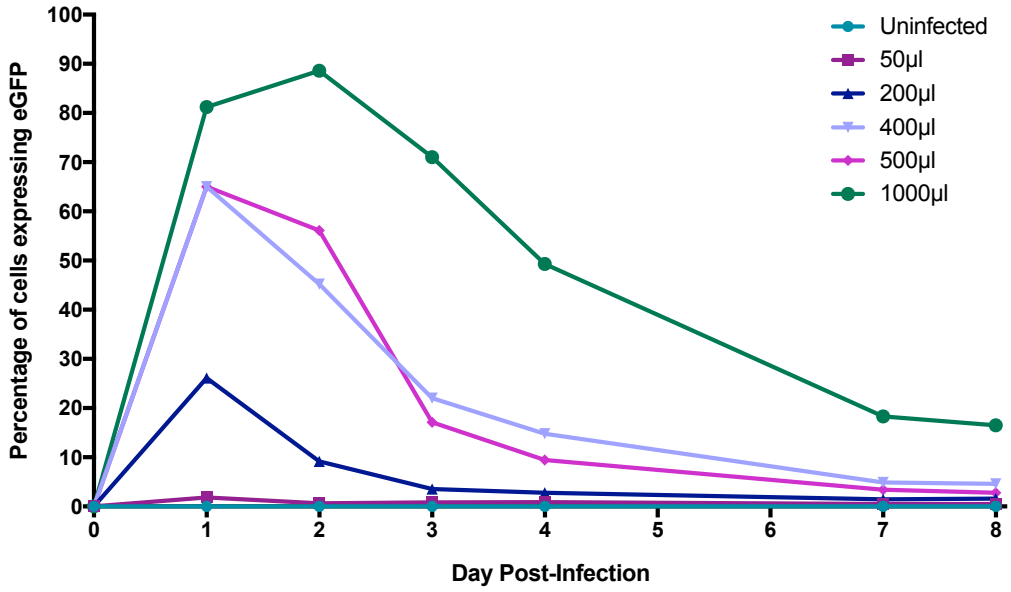


Figure 3.9 eGFP profile over time in Jurkat cells infected at a range of viral input volumes (pRGH-WT). Expression was measured to determine when incoming capsid eGFP had decayed. Key values represent volumes (µl) of viral supernatant added at infection. Infected cells were maintained under standard incubation conditions throughout the course of the sampling period and flow cytometric measurement of the percentage of cells expressing eGFP was performed on 1, 2, 3, 4, 7 and 8 days post infection.

This experiment was repeated with the CEM.NK^{R} CCR5+ cell line. Using a different pseudovirus, CAP303, at a volume determined to be optimal in the Jurkat E6-1 cell line (section 3.5.2), similar kinetics of decay were observed in the CEM.NK^{R} CCR5+ cell line when compared to Jurkat E6-1 cells. We observed a spike of eGFP expression on day one post-infection which gradually decreased till day six where it began to stabilize (Figure 3.10).

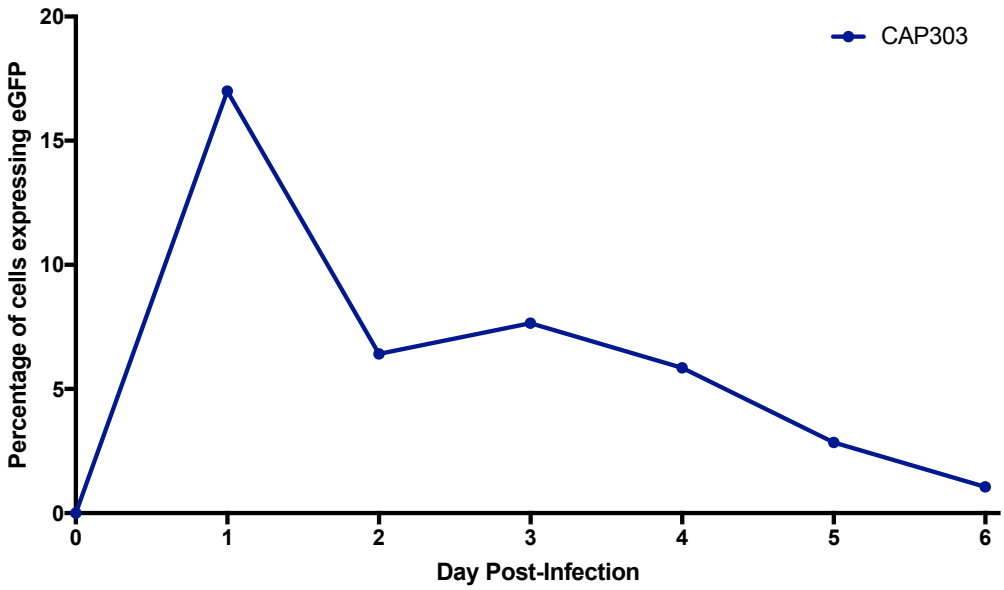


Figure 3.10 eGFP profile over time in CEM.NK^{R} CCR5+ cells infected with the CAP303 pseudovirus. Cells were infected at pseudovirus titres previously determined to be optimal for each pseudovirus in the Jurkat E6-1 cell line. Expression was measured to determine when incoming capsid eGFP had decayed. Infected cells were maintained under standard incubation conditions throughout the course of the sampling period and flow cytometric measurement of the percentage of cells expressing eGFP was performed each day post-infection.

Using profiles from Jurkat E6-1 and CEM.NK^{R} CCR5+ cells, the measurement of latency potential was taken at day eight post-infection for both cells lines.

3.5.2 Determination of viral titre

To determine the optimal viral volume to obtain 10-20% of cells expressing eGFP one day post-infection thus ensuring single integration events per cell, Jurkat E6-1 cells and CEM.NK^{R} CCR5+ cells were infected with two volumes of viral supernatant, 50µl or 150µl, and 100µl or 250µl, respectively. eGFP expression was subsequently

measured by flow cytometry one day post-infection. Preliminary experiments suggested that CEM.NK^{R} CCR5+ cells were more difficult to infect, and higher volumes of pseudovirus, increased spinoculation speed and smaller cell culture plates were used to infect these cells (Table A3.2, Appendix 3). A range of infection rates were obtained for each pseudovirus in both cell lines (Figure 3.11).

Based on these results, a viral volume was estimated for infections such that eGFP expression fell between 10 and 20%, on day one post-infection. Final volumes used for infection are shown in (Table A3.2, Appendix 3).

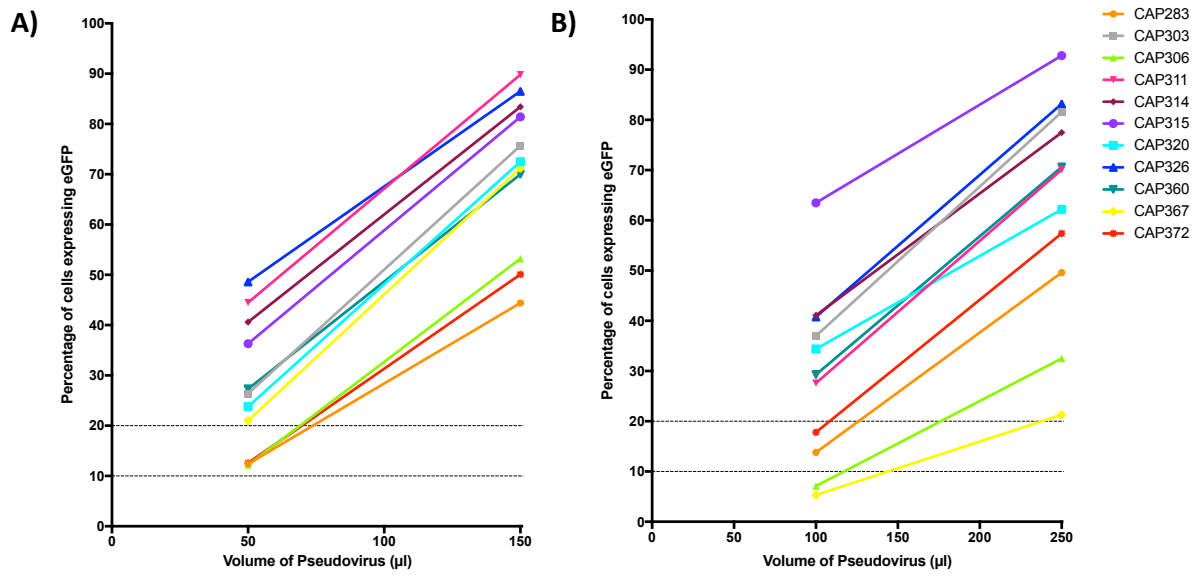


Figure 3.11 Determination of pseudovirus volume input. Percentage of GFP expressing cells at different viral volume for each participant LTR-pseudotyped virus. **A)** Jurkat E6-1 or **B)** CEM.NK^{R} CCR5+ cells were each infected with two volumes of viral supernatant, 50 or 150µl, and 100 or 250µl respectively and GFP expression measured 1 day post-infection by flow cytometry.

3.6 Subtype C participant LTR genotype is associated with a range of latency potentials

The effect of intra-subtype genotypic LTR variation on the frequency of latently infected cells early post-infection was assessed, using LTR-pseudotyped viruses (section 3.2). Jurkat E6-1 cells were infected with each participant pseudotyped virus at a volume which yielded approximately 10-20% of cells expressing eGFP on day

one post-infection (section 3.5.2). Fluorescent protein expression (mCherry and eGFP) was measured by flow cytometry at eight days post-infection (section 3.5.1). The latency potential of each LTR was measured by determining the ratio of latently infected cells (mCherry only) (red) to actively infected cells (mCherry and eGFP) (yellow) (Figure 3.12).

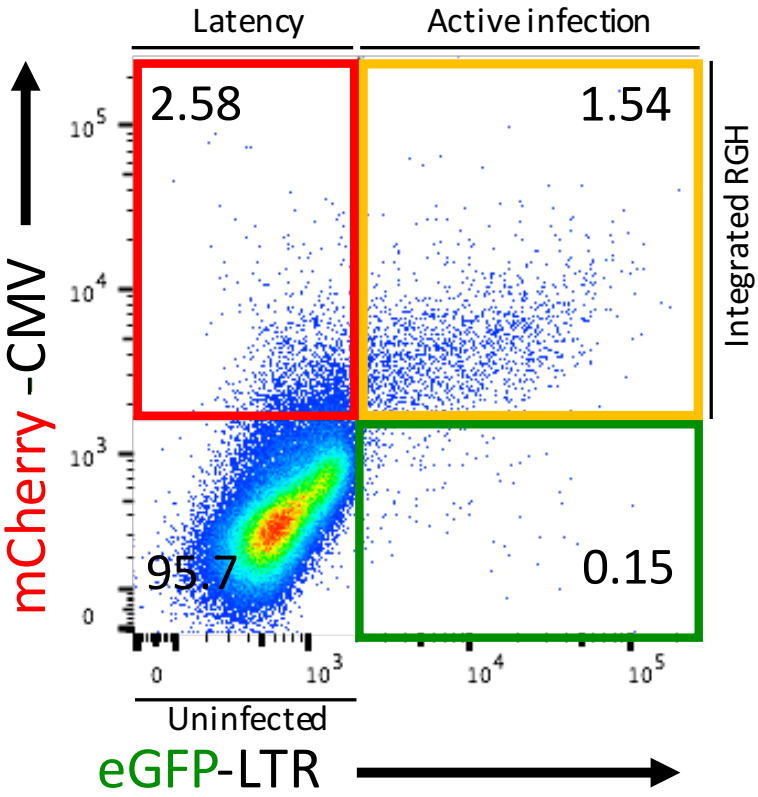
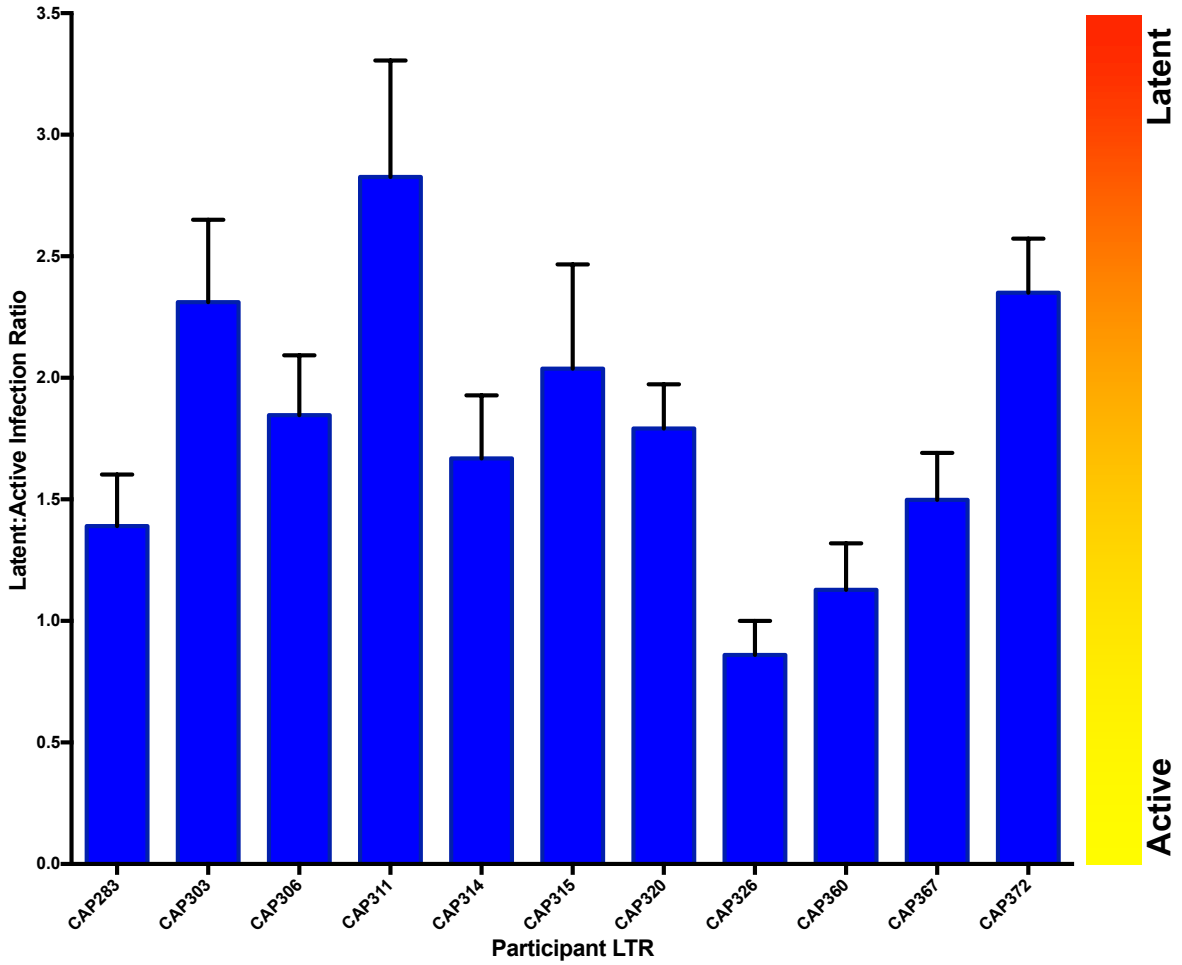


Figure 3.12 Depiction of pRGH infected cell populations detected by flow cytometric analysis. Pseudocolour dot plot of a pRGH infected population of Jurkat E6-1 cells. Cells within the upper left quadrant represent infected cells containing integrated RGH DNA with transcriptionally active CMV promoters but silent LTR's (latent infection). Cells within the upper right quadrant represent infected cells containing integrated RGH DNA with transcriptionally active CMV promoters and LTR's (active infection). Cells within the bottom left quadrant expressed neither fluorescent protein and were considered uninfected.

In three independent experiments, the median latent:active infection ratio of pseudoviruses in the Jurkat E6-1 cell line was 1.97 with a range of 0.86-2.83 (Figure 3.13). Representative flow cytometric pseudocolour dot plots are shown in Figure A3.3, Appendix 3. A latent:active ratio median above 1 is demonstrative of a majority of infected cells harbouring latent proviruses upon infection across participants. The latency potential of CAP311 was significantly different from that of three other

participants, while the latency potential of CAP326 differed significantly from that of two other participants (ANOVA, followed by a Tukey multiple comparison test).



Set Comparison	p-value	p-Value Summary
CAP283 vs. CAP311	0.048	*
CAP303 vs. CAP326	0.044	*
CAP311 vs. CAP326	0.003	**
CAP311 vs. CAP360	0,011	*
CAP326 vs. CAP372	0.036	*

Figure 3.13 RGH demonstrates the differences in latency observed across HIV-1 participant LTRs in Jurkat E6-1 cells. Jurkat E6-1 cells were infected with the participant pRGH variants. At 7 days post-infection, cells were treated with DMSO for 24 hours and subsequently analysed by flow cytometry. Latency potential is expressed as the ratio of latent (mCherry only) to actively (mCherry and eGFP) infected cells. Error bars represent standard errors of the results of triplicate experiments. Participant pairs significantly different from each other are shown below (one way ANOVA, followed by a Tukey multiple comparison test).

To confirm that silent LTRs within infected cells were transcriptionally competent, half of each of the infected cultures was treated with PMA/Iono on day seven post-infection

to induce T cell activation, thereby stimulating transcription. PMA and Iono cooperatively enhance the activation of cellular Protein Kinase C pathway resulting in T cell activation (described in section 1.9.4). Indicative of an increase of cells with active LTRs, a decrease in the latent:active ratio proportion decreased to below 1 in all cases in Jurkat E6-1 infected cells, with a median of 0.55 (range: 0.46-0.78) (Figure 3.14).

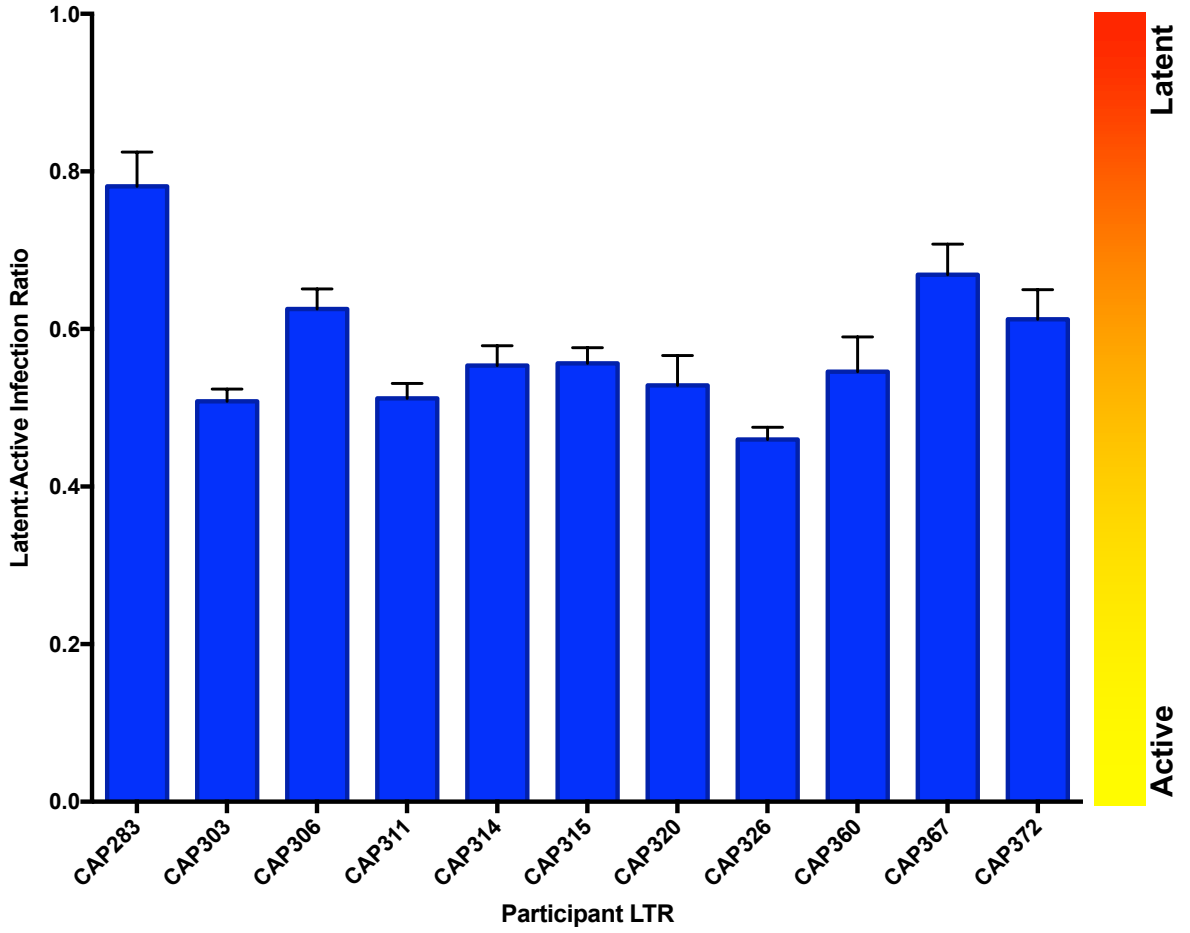


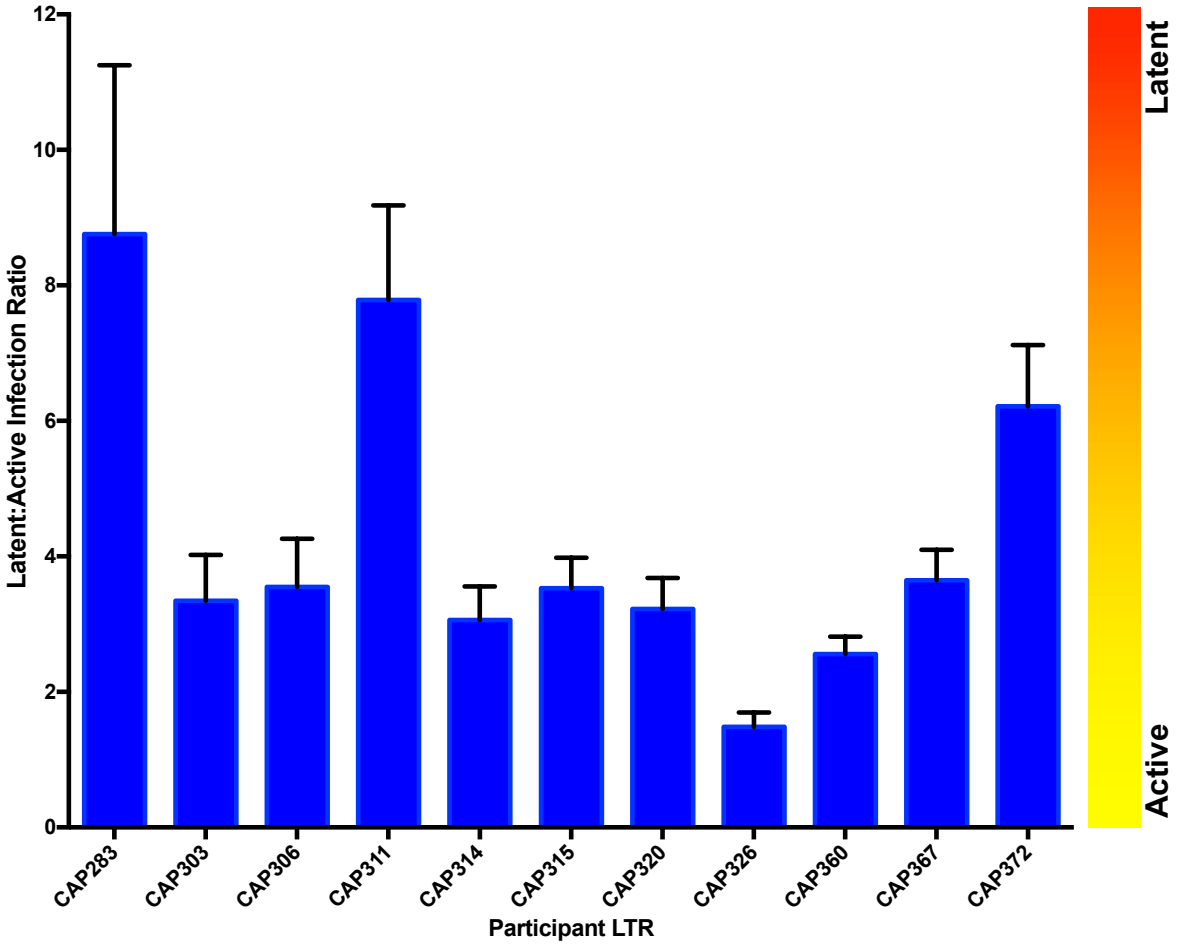
Figure 3.14 Latency potential in stimulated cells. Jurkat E6-1 cells were infected with the participant pRGH variants. At 7 days post-infection, cells were treated with PMA/Iono for 24 hours and subsequently analysed by flow cytometry. Latency potential is expressed as the ratio of latent (mCherry only) to actively (mCherry and eGFP) infected cells. Error bars represent standard errors of the results of triplicate experiments.

The fold decrease in latency potential differed across participant LTRs, with a median fold decrease of 3 (range 1.78-5.52).

Interestingly, not all proviruses in latently infected cells were reactivated as not all infected cells expressed eGFP post-stimulation. Furthermore, following stimulation, the proportion of cells expressing mCherry increased from a median of 2.5 to 9.8.

3.7 LTR-associated latency potential evaluated in an alternate cell line

To determine if the latency potential profile (latent:active ratio) (section 3.6) was a property of specific LTR genotype and not a property of Jurkat E6-1 infections, the latency assay was repeated in the CEM.NK^{R} CCR5+ cell line. CEM.NK^{R} CCR5+ cells were infected with the same panel of LTR-pseudotyped viruses (n=11) and protein fluorescence measured by flow cytometry. In three independent experiments, the median latent:active ratio was 3.15 with a range of 1.46-8.8 (Figure 3.15). Representative flow cytometric pseudocolour dot plots are shown in Figure A3.4, Appendix 3. In this cell line, a greater proportion of latently infected cells than actively infected cells were observed with all participant LTRs. The latency potential of CAP283 was significantly different from that of eight other participants, while the latency potential of CAP311 differed significantly from that of two other participants(ANOVA, followed by a Tukey multiple comparison test).



Set Comparison	p-value	p-Value Summary
CAP283 vs. CAP303	0.028	*
CAP283 vs. CAP306	0.038	*
CAP283 vs. CAP314	0.018	*
CAP283 vs. CAP315	0.037	*
CAP283 vs. CAP320	0.023	*
CAP283 vs. CAP326	0.001	**
CAP283 vs. CAP360	0.008	**
CAP283 vs. CAP367	0.044	*
CAP311 vs. CAP326	0.007	**
CAP311 vs. CAP360	0.037	*

Figure 3.15 RGH demonstrates the differences in latency observed across HIV-1 participant LTRs in CEM.NK^R CCR5+ cells. CEM.NK^R CCR5+ cells were infected with the participant RGH variants. At 7 days post-infection, cells were treated with DMSO for 24 hours and subsequently analysed by flow cytometry. Latency potential is expressed as the ratio of latent (mCherry only) to actively (mCherry and eGFP) infected cells. Error bars represent standard errors of the results of triplicate experiments. Participant pairs significantly different from each other are shown below (one way ANOVA, followed by a Tukey multiple comparison test).

We determined if the latency potential hierarchy was maintained across the CEM.NK^R CCR5+ and Jurkat E6-1 cell lines. While there was no significant relationship when comparing data from all 11 participants (p = 0.18, r = 0.44) (Figure 3.16A), the hierarchy across participants was largely maintained when an outlier, participant CAP283 was excluded from the analysis (Spearman correlation test p = 0.03, r = 0.68) (Figure 3.16B). CAP283 was identified as an outlier as this LTR displayed the third lowest latency potential in Jurkat E6-1 cells but the highest in CEM.NK^R CCR5+ cells. This observation provides some evidence of a specific LTR sequence having an inherent latency potential independent of cell line.

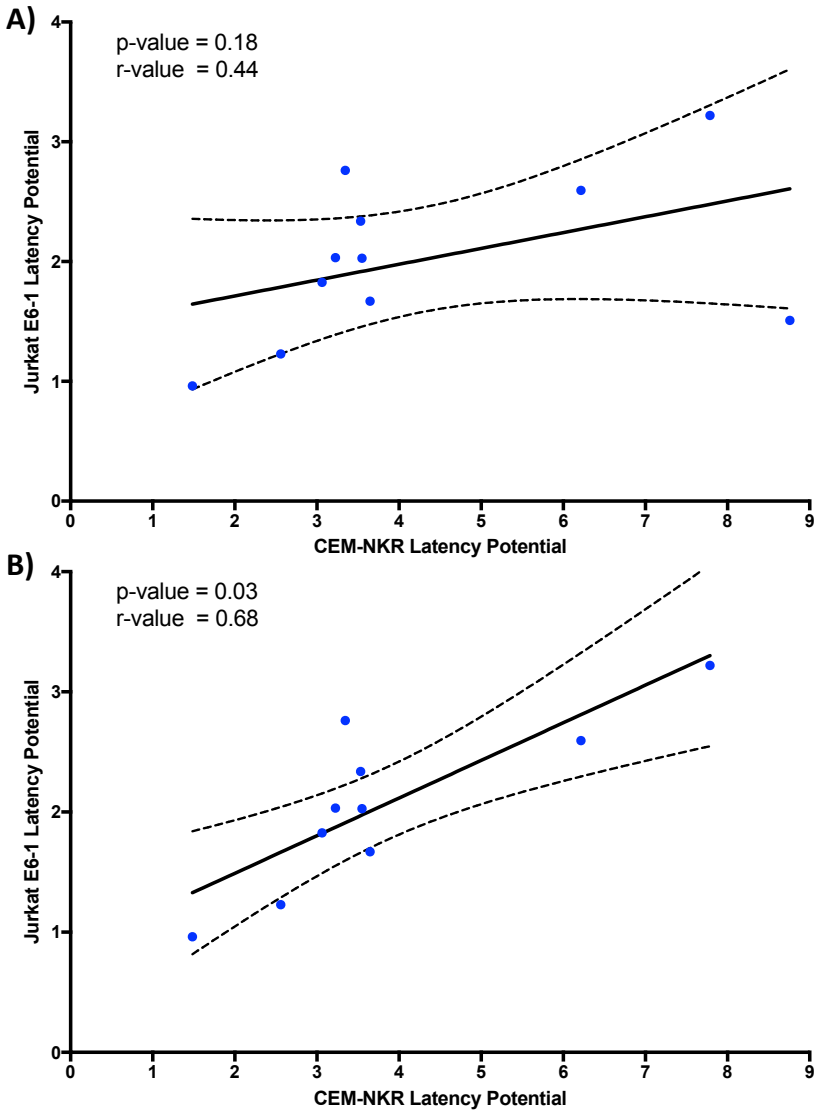


Figure 3.16 The association between latency potential measured in two different cell lines. Jurkat E6-1 and CEM.NK^R CCR5+ cells were infected with participant LTR-pseudotyped RGH viruses and latency potential measured by flow cytometry as described in in sections 3.6 and 3.7. **A)** All participants data included **B)** CAP283, identified as an outlier, was excluded. A Spearman correlation test was performed. The dashed curves indicate the 95% confidence intervals.

Similar to Jurkat E6-1 cells, after stimulation of CEM.NK^R CCR5+ cells, the total proportion of mCherry expressing cells increased (from median of 2.8 to 7.4). However, unlike Jurkat E6-1 cells, the latent:active ratio increased to 4.5, indicating that the stimulation was not activating the latent proviruses. To determine whether this was due to an insufficient concentration of PMA, RGH pseudovirus-infected cells were treated with a range of increasing concentrations of PMA. Iono concentration remained constant as the initial concentration used was higher than that used for standard stimulation of T cells to induce viral expression(319). Increasing the PMA concentration 25-fold up to 100ng/ml did not reactivate these latent proviruses as an approximately 1.5-fold increase in the latent:active ratio compared to DMSO-treated cells was observed with all three PMA concentrations (Figure 3.17). Therefore, it was not possible to assess whether all silent infections were truly representative of latency in CEM.NK^R CCR5+ cells. Thus, these data were not used in subsequent analyses.

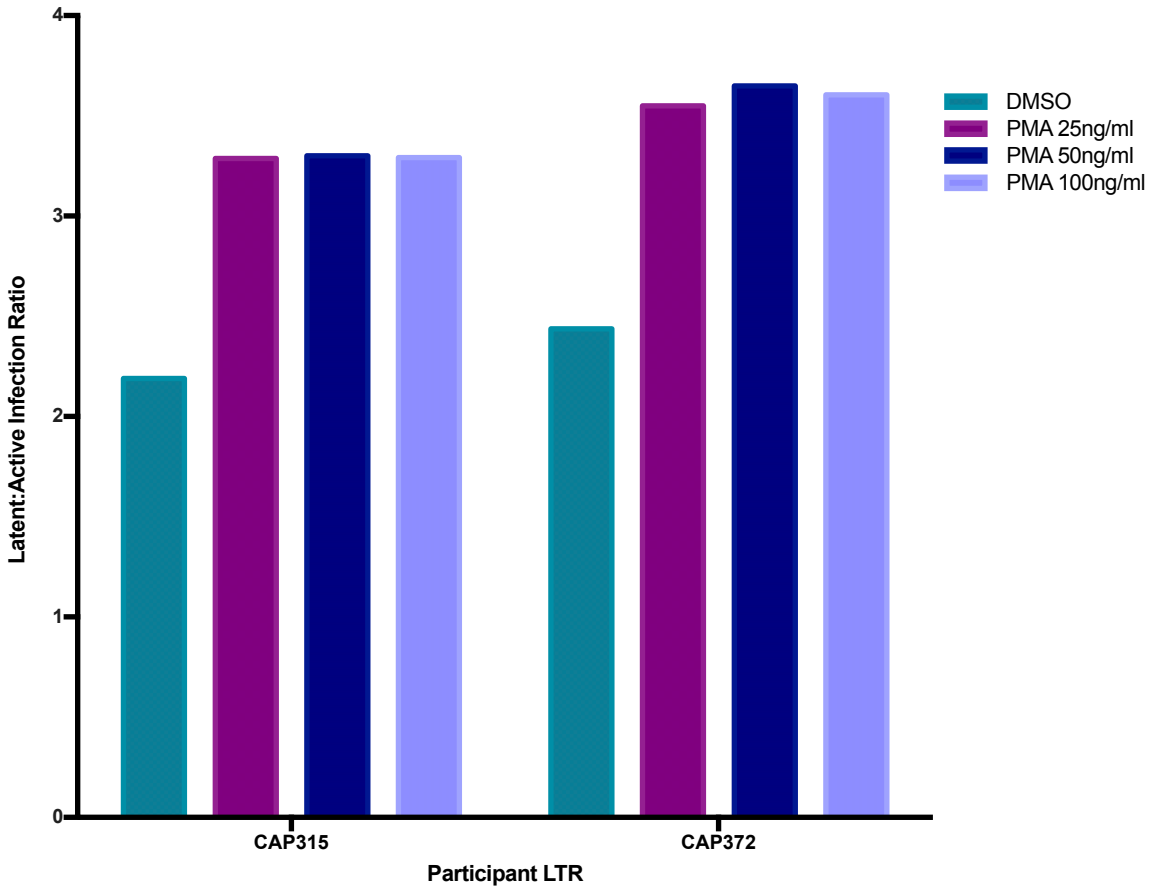


Figure 3.17 The effect of PMA concentration on RGH-infected CEM.NK^{R} CCR5+ cells. Four CEM.NK^{R} CCR5+ cell cultures were each infected with either CAP315 or CAP372 LTR-pseudotyped RGH viruses. At 7 days post-infection, each infected culture was treated with 25, 50 or 100ng/ml of PMA and 1µM Iono or DMSO as a control for 24 hours and subsequently analysed by flow cytometry. Latency potential is expressed as the ratio of latent (mCherry only) to actively (mCherry and eGFP) infected cells.

3.8 Evaluation of basal and Tat-induced LTR promoter activities

We were interested to know whether LTR promoter activity influenced viral latency. Following HIV-1 infection, the LTR-dependent HIV-1 transcription can be divided into two phases. In the initial phase, which occurs immediately after integration, the LTR has a basal level of activity that is mediated by cellular transcription factors that interact with the LTR. The second phase, following the translation of HIV-1 proteins, is largely dependent on Tat, whereby the presence of Tat enhances gene expression.

We measured basal activity by transfecting HEK293T cells with a plasmid where the LTR was cloned upstream of a luciferase gene (LTR-pGL4.10) (section 2.7) (Figure

A2.2 Appendix 2) (plasmids provided by P Selhorst, UCT, unpublished). The Tat-dependent activity was measured by co-transfecting the LTR-pGL4.10 clones with a plasmid that expressed a subtype C Tat protein (Tat-TOPO). The optimum amount of Tat-TOPO needed to induce promoter activity within the linear range was evaluated using LTRs from 11 participants. An optimum amount of 50ng of the Tat-TOPO plasmid per reaction was determined (Figure A3.1, Appendix 3) and used for all experiments.

The panel of 11 participant LTRs displayed variable basal activity ranging from 0.65 to 2.4 times BaL-LTR activity median 1.871 (Figure 3.18A and Table A3.3, Appendix 3). Following Tat induction, significantly higher levels of LTR activity were observed (p-value = 0.001, Wilcoxon matched-pairs signed rank test) with a 4.8-fold mean increase (range 2.5-7.4). Tat-induced LTR activity had a median of 8.9 times BaL-LTR activity with a range of 1.61-15.83 (Figure 3.18A).

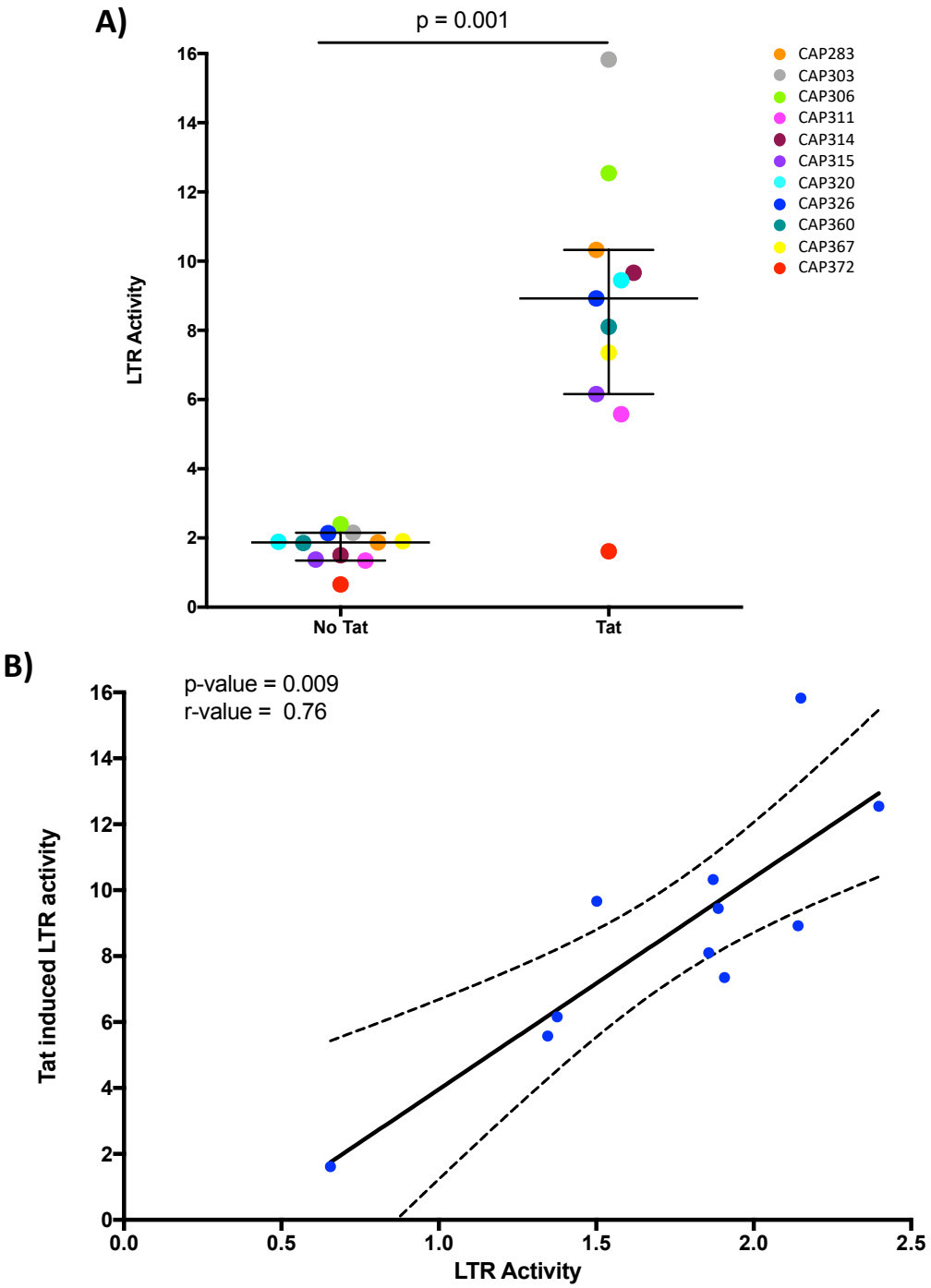


Figure 3.18 Basal and Tat-induced LTR promoter activity. A) Scatter plot of promoter activity of participant derived LTR sequences in the absence and presence of the HIV-1 Tat protein. HEK293 cells were transfected with each of the panel of pGL4.10 LTR clones or co-transfected with these clones and 50ng of the Tat-TOPO plasmid. A *Renilla* luciferase expressing vector was used to normalise transfection efficiency in all cases. LTR activity was measured by determining the luciferase production after the addition of luciferase at 48 h post-transfection and expressed as a relative response ratio to the HIV-1 subtype B Bal strain LTR luciferase expression. The statistical comparison was performed using the Wilcoxon matched-pairs signed rank test. The bars indicate the median and error bars are representative of the interquartile range. **B)** The association of basal and Tat-induced LTR promoter activity. A Spearman correlation test was performed. The dashed curves indicate the 95% confidence intervals.

There was one participant LTR, CAP372, which showed limited activation by Tat and which also had the lowest basal activity. Overall, the activity hierarchy was largely maintained across participant sequences in the presence and absence of Tat (Spearman correlation test; p-value = 0.009, r-value = 0.76) (Figure 3.18B). Of interest, basal transcriptional activity had a limited range, whereas this range was greatly expanded after Tat induction.

As genotypically different LTRs had different promoter activities as well as latency potentials, we next wanted to determine if any relationship existed between these two properties. No significant correlation was observed between latency potential in Jurkat E6-1 cells and basal or Tat-induced LTR promoter activity (Spearman correlation tests, basal p = 0.25, r = -0.38, Tat-induced p = 0.42, r = -0.27) (Figure 3.19)

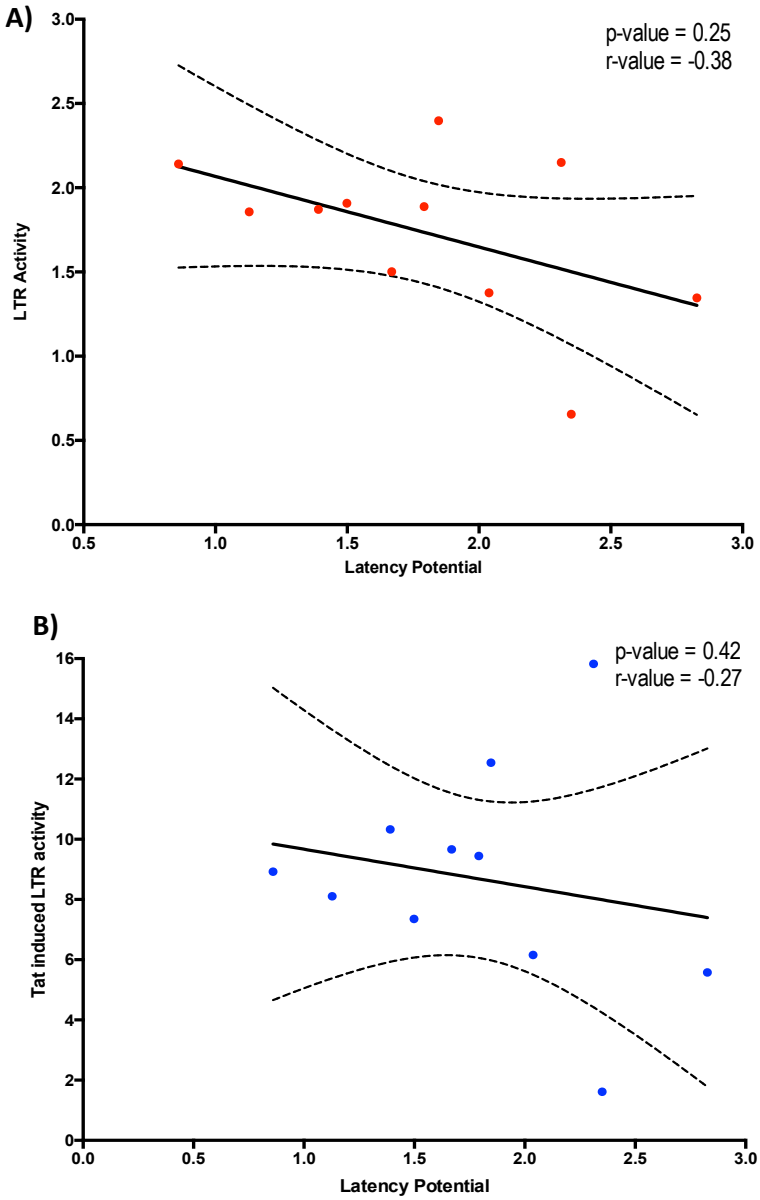


Figure 3.19 The relationship between latency potential and LTR activity. HEK293T cells were transfected with **A)** the panel of pGL4.10 luciferase expression clones or **B)** co-transfected with the panel of pGL4.10 luciferase expression clones and 50ng of the Tat-TOPO plasmid. A *Renilla* luciferase expressing vector was used to normalize transfection efficiency. LTR activity was measured by luciferase expression after the addition of luciferin at 48 h post-transfection and was expressed as a response ratio relative to the HIV-1 BaL strain LTR luciferase expression. Latency potential is expressed as the latent:active ratio of DMSO treated samples as in Figure 3.13. The dashed curves indicate the 95% confidence intervals.

3.9 Influence of LTR genotype on latency potential

It has been demonstrated that a minor genotypic deviation from the consensus in domains of the LTR associated with binding of regulatory cellular factors was sufficient

to alter latency establishment(101). To investigate the underlying factors contributing to the differences in LTR-dependent latency, we investigated if there were genotypic features associated with LTR-dependent latency.

In order to characterise and identify the differences from the consensus C sequence within each participant-specific LTR, (HXB2 position 1-532) a *Highlighter* plot (www.hiv.lanl.gov) was generated. Participant LTRs largely resembled the subtype C consensus sequence with differences away from consensus distributed across the length of each LTR (Figure 3.20). CAP360, CAP303 and CAP320 each harboured a 15-21bp insertion between HXB2 site 326-327, with CAP320, harbouring an additional NF-κB-like motif inserted in this site, upstream of the canonical three sites in the enhancer region. One participant LTR, CAP372, had a single base pair difference from consensus in the AP-1 binding site within the enhancer region.

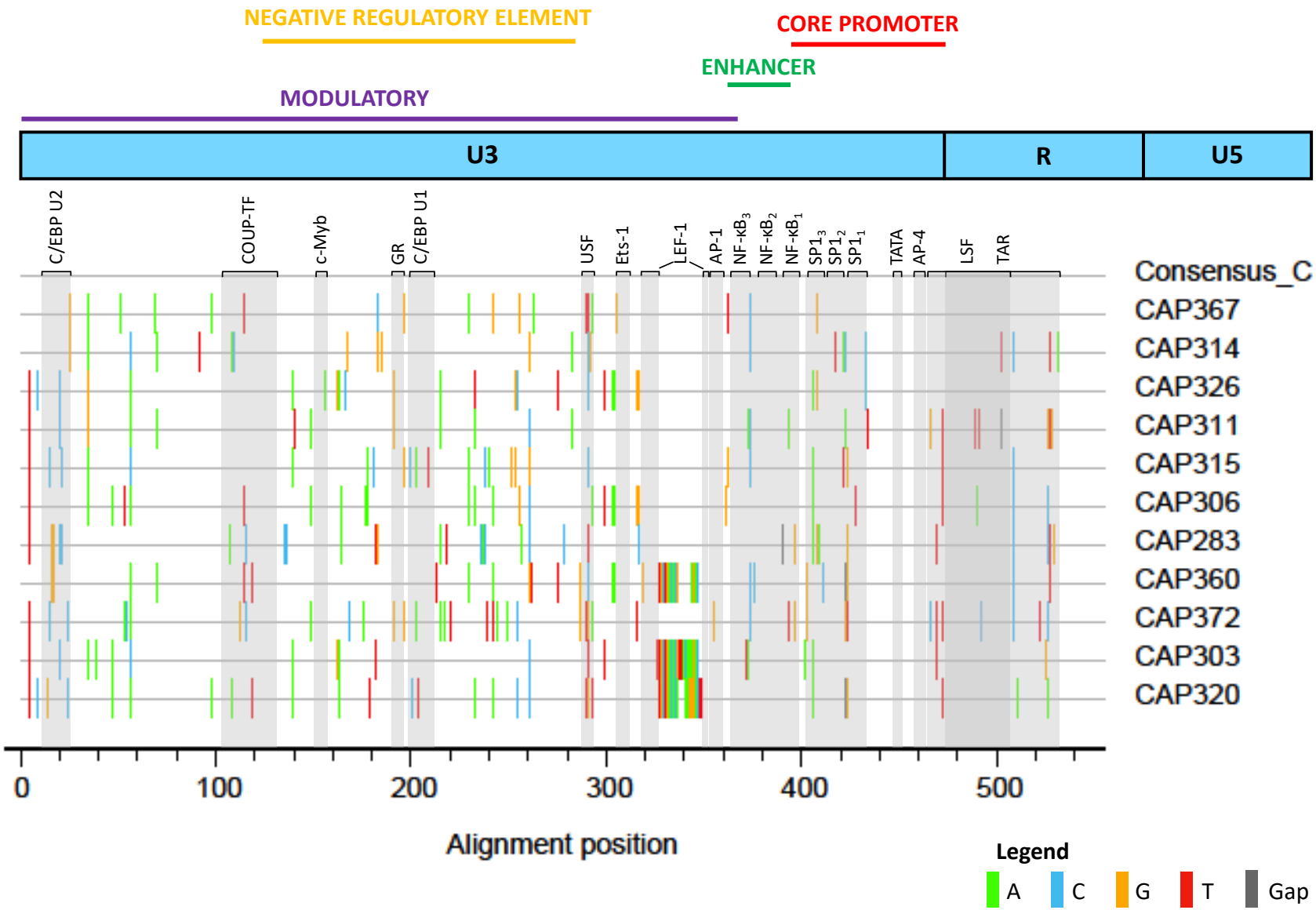


Figure 3.20 CAPRISA 004 LTR genotypic differences away from a consensus C LTR. Highlighter (www.HIV.LANL.gov) plot of the U3 and R regions of the panel of participant LTR sequences. Tic marks represent nucleotide substitutions as compared to the top-most Subtype C Consensus LTR sequence. Sequences are listed from top to bottom in order of least to most distant from consensus. Sites of well-characterised cis-regulatory elements, where published consensus binding sites could be identified within the consensus C LTR, are indicated in the highlighted grey boxes.

Additionally, we measured the pairwise DNA distance of each participant LTR to a consensus C LTR (HXB2 position 1-532). There was higher overall DNA distance from LTR consensus C (median 0.056; range 0.039-0.094), compared to the 19 domains associated with regulation (median 0.047; range 0.026-0.087) (Figure 3.21). One participant LTR, CAP372, was highly divergent from the consensus sequence.

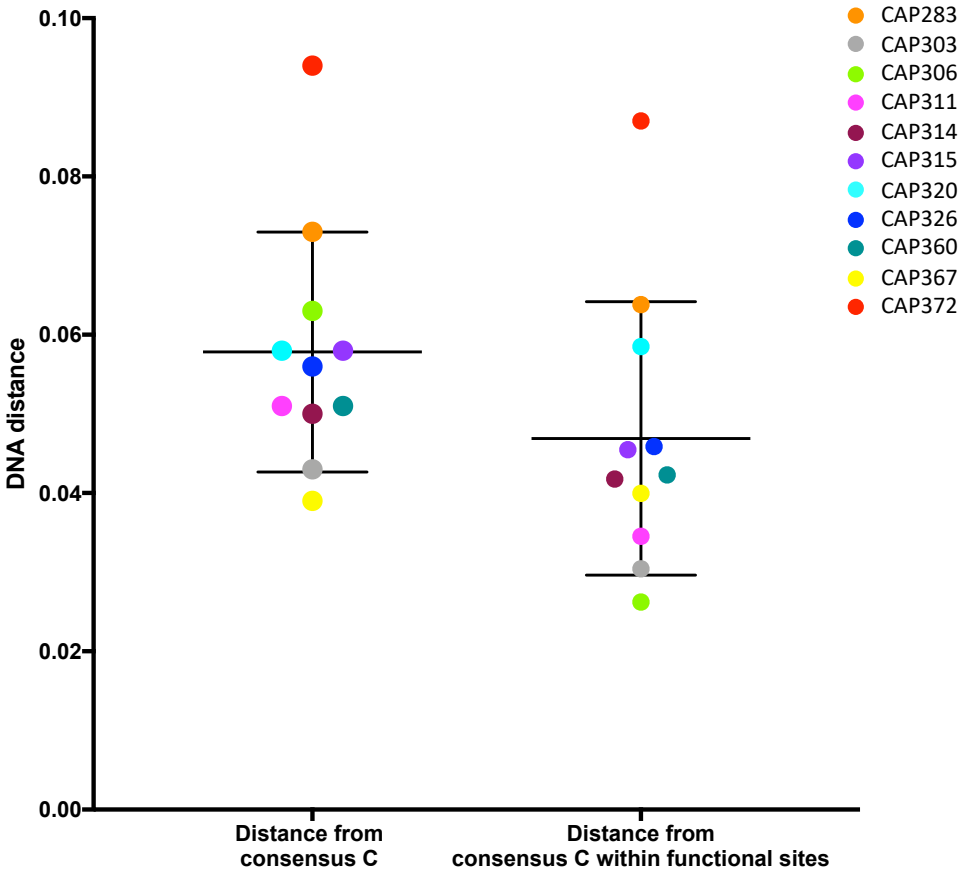


Figure 3.21 Scatter plot of pairwise DNA distances between a consensus C LTR sequence, and each of the selected panel of pGL4.10 cloned participant LTR sequences. Distances were calculated across the entire sequence and within functional sites. The number of base substitutions per site between sequences are shown. Analyses were conducted using the Maximum Composite Likelihood model(318). A total of 532 positions were analysed. The bars indicate the mean and error bars represent the standard deviation.

Interestingly, the average number of mutations per functional site was higher in the modulatory region than both the enhancer and core promoter regions. The enhancer region was the most highly conserved with an average DNA distance of 0.021 per functional site (Figure 3.22).

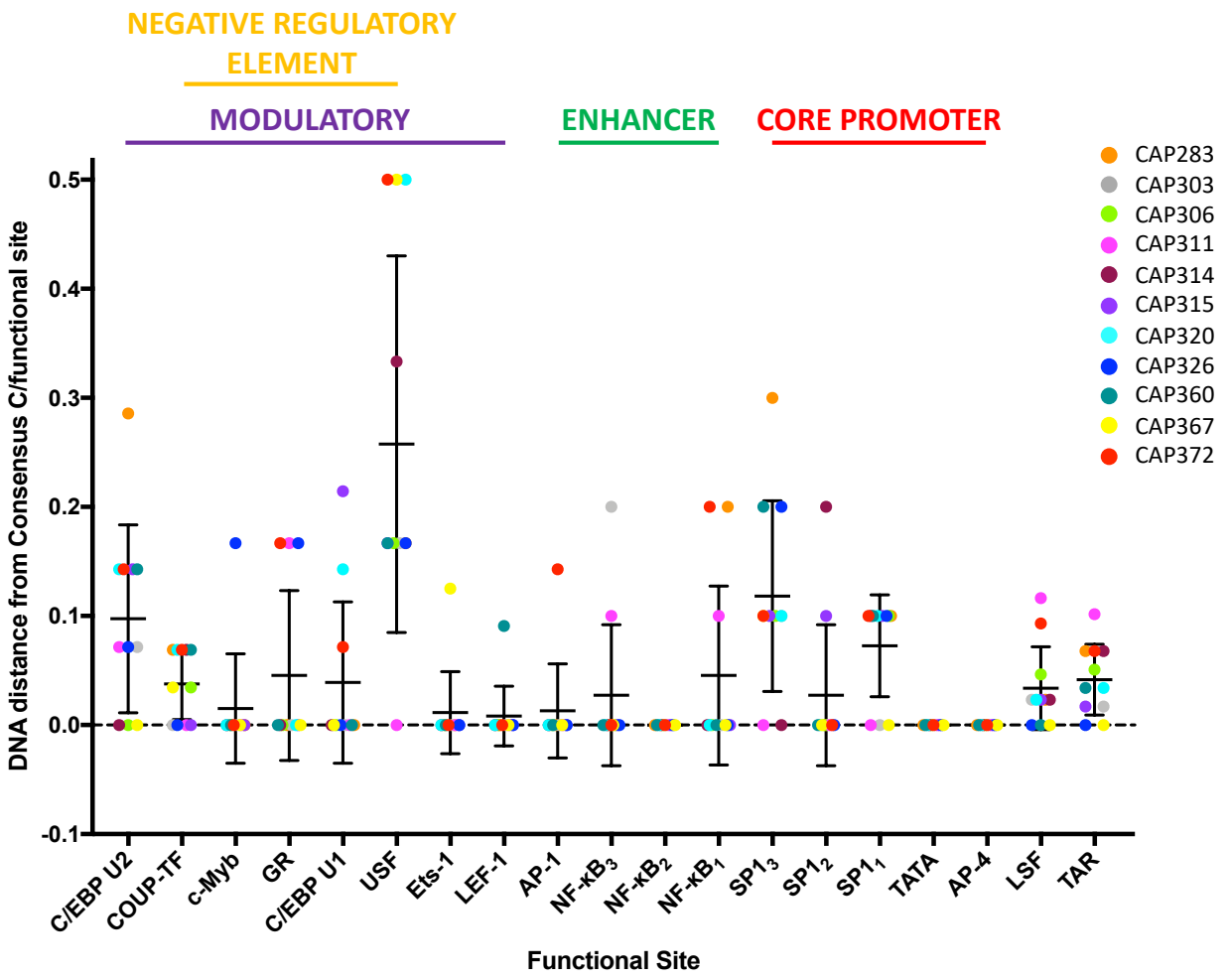


Figure 3.22 CAPRISA 004 LTR DNA distance from a consensus C sequence within each functional site. The bars indicate the mean and error bars represent the standard deviation.

We next wanted to determine if any association existed between the average DNA distance per functional site from a consensus C LTR and participant-specific latency potential, given the variability within the LTR. No significant correlation was observed, Spearman correlation test Jurkat E6-1, $p = 0.47$ $r = -0.25$ (Figure 3.23).

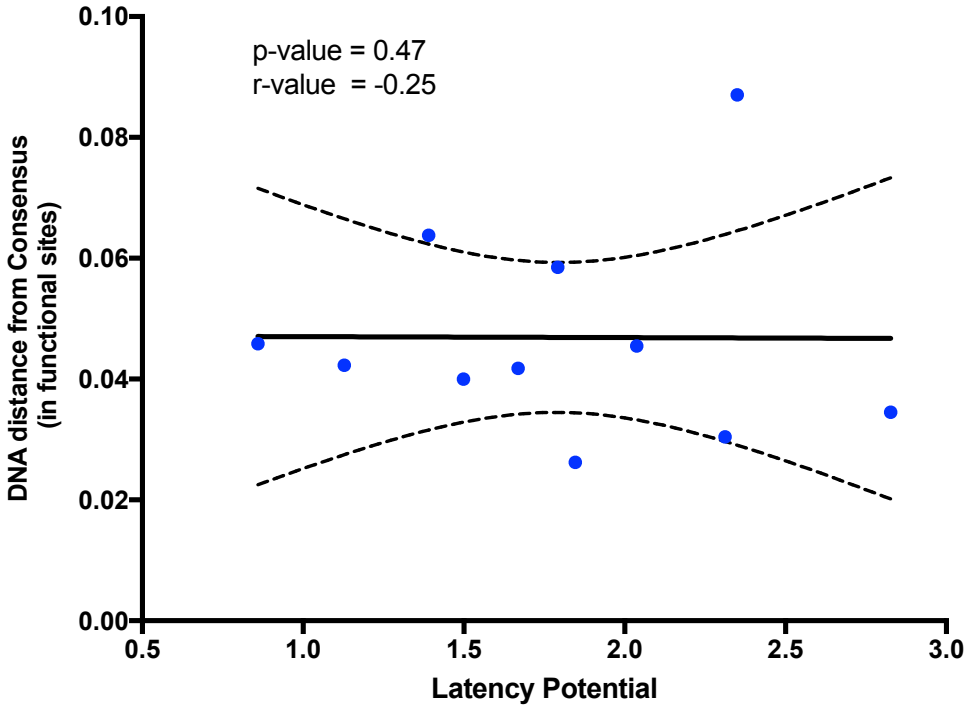


Figure 3.23 The association between CAPRISA 004 LTR latency potential and DNA distance from a consensus C LTR sequence within functional sites. Latency potential is expressed as the latent:active ratio of DMSO treated samples as in Figure 3.12. The dashed curves indicate the 95% confidence intervals.

Chapter 4: Discussion

HIV-1 latency remains the greatest barrier to eradication of the virus from an infected individual. Latency allows viral DNA to persist in the presence of all current therapeutic strategies, and for re-emergence of viremia after treatment interruption. Transcriptional silencing of the HIV-1 promoter element, the LTR, plays an important role in the occurrence of viral latency(79, 147, 320). However, the role of viral factors in latency remains poorly understood. In this study, we investigated the role of LTR genotype on HIV-1 latency establishment. Using LTRs from acutely infected South African women, representative of the subtype C epidemic, we implemented an *in vitro* model to measure the proportion of latently infected cells shortly after infection (latency potential) and evaluate how this potential relates to promoter strength. We found that latency potential varied across subtype C LTRs and was not influenced by promoter activity as measured by transcription in the presence and absence of Tat. Furthermore, we found that viruses that were divergent from the most common sequence (subtype C consensus) did not have significantly different latency potentials. To our knowledge, this is the first study to characterize the influence of inter-participant, intra-subtype C genotypic diversity on HIV-1 viral latency.

We used a dual reporter latency model developed by Dahabieh *et al.*, (2013) to measure latency(156). This model enables identification of three populations of cells, including, uninfected cells, cells harbouring latent virus and those with actively replicating viruses. This is achieved using two independent reporter genes, mCherry (integration marker) and eGFP (viral expression marker) that are under the control of a constitutive CMV promoter and the HIV-1 LTR promoter respectively. This dual reporter system has an advantage over other single reporter based systems, such as those developed by, Jordan *et al*.(240), and Burke *et al*.(255), in that it provides a positive marker for latent infection and negates the need for selection of only latently infected cells.

We show that latency potential varied across LTR sequences derived from 11 subtype C acutely HIV-1 infected women, with significant differences across multiple individuals.

However, the range of intra-subtype C latency potential was smaller than that observed in LTRs from different subtypes(156), supporting the hypothesis that the degree of variation of the LTR plays an important role in influencing latency potential.

Latency establishment is a phenomenon reliant on a variety of molecular mechanisms, acting in synergy to induce the latent state(79, 147, 320). Some of these mechanisms do not directly target the LTR. In the current model, the majority of cellular conditions were assumed to be homogenous, and in Jurkat E6-1 cells, the latency potential hierarchy within the 11 participant LTRs was maintained across replicate experiments. This demonstrates that there is an inherent genotype-dependent property of HIV-1 LTRs that influences the latency potential of a specific virus. However, although conditions in tissue culture are assumed to be uniform, it was not possible to completely control for all intracellular factors such that the LTR genotype of the infecting virus was the only variable factor in the system. This may therefore account for variations across samples.

Consistent with findings by Dahabieh *et al.*(156), we observed that, days after infection of the Jurkat E6-1 cell line, the majority of infected cells harboured transcriptionally silent proviral DNA. This suggests that the virus has the capacity to extensively establish latency early post-infection. However, it is important to note that a range of other *in vivo* factors may influence which cells become latently infected. Therefore, this model provides the advantage of investigation of viral factors in isolation. Given that the RGH model only measures the proportion of infected cells in which latency is established within a few days of infection, and that mechanisms of latency establishment are unknown, the effect of longer culture on the proportion of latently infected cells cannot be speculated.

We were able to detect latently infected cells in our cultures eight days post-infection. However, it is possible that latency was established earlier but could not be measured due to masking by residual eGFP. Additionally, the cell lines utilised for infections were actively proliferating and not resting populations. Although latency is thought to be established when actively infected CD4+ T cells transition to a resting memory state *in vivo*, detection of latent provirus in our study cells disputes this theory, and supports the

multiple studies that report that actively infected CD4+ T cells can harbour transcriptionally silent HIV-1 provirus(146, 147, 151, 157).

Post-stimulation, the proportion of latently infected cells decreased in the Jurkat E6-1 cell line. This confirms that latency was successfully established in a subset of cells and these viruses were inducible via T cell activation. If all infected cells were capable of actively transcribing mCherry and eGFP, cells expressing both proteins would be expected to be present post-stimulation. However, since mCherry-only expressing cells remained, we hypothesised that (i) PMA/Iono stimulation was insufficient to activate all silent LTRs, (ii) the mechanism of silencing was not directly related to transcription/the LTR, or (iii) some cells were infected with a provirus with a non-functional eGFP cassette. Unlike Jurkat E6-1 cells, stimulation of the CEM.NKR CCR5+ cells did not decrease the latent:active ratio even at higher concentrations of PMA. These cells were either resistant to PMA activation, or the cells were already partially activated, and PMA did not increase the activation threshold sufficiently to reactivate the silent LTRs. Alternatively, latency could have been established by an alternate mechanism to that in Jurkat E6-1 cells and was not reversible by effects induced by PMA/Iono in this cell line. To our knowledge, only one study has investigated the effect of PMA/Iono on CEM.NKR CCR5+ cells and it was shown that it induced internalization of the CXCR4 co-receptor, however this is not likely to impact the activation state of the cell or ability to reverse latency(321).

Although the CMV promoter is considered a strong promoter, it has previously been demonstrated, albeit in myeloid cell lineages, that the promoter can be silenced via histone modifications(322, 323). It has also been shown that the CMV promoter is sensitive to T cell activation, with increased transcription of the immediate early genes under the control of this promoter observed upon PMA treatment(324, 325). Interestingly, we observed an overall increase in the proportion of infected cells (mCherry expressing), in both cell lines, following PMA/Iono stimulation. This observation suggests that not all infected cells expressed mCherry constitutively. Furthermore, this indicates that PMA/Iono is capable of activating the CMV promoter. It is therefore possible that the model may underestimate the number of HIV-1 infected cells and, it would be pertinent

to replace the CMV promoter with an alternative promoter such as the EF1α, ACTB or PGK promoters that have been observed to be more stable in a range of cell lines(326).

It has been demonstrated that the replicative capacity and fitness of a virus is influenced by subtype-level genotypic variation of its LTR(77, 78). This was proven to be due to the presence of specific motifs e.g. additional NF-κB site in subtype C or an additional TATAA box sequence in subtype E. Conservation of a binding sequence for AP-1, present in some HIV-1 subtype LTRs but not in others, has been shown to promote the ability of the different viruses to establish latency(101). This AP-1 site was highly conserved across our participant LTRs, with the exception of a single difference away from consensus in one participant. Interestingly, this sequence displayed one of the highest latency potentials in our panel. It is possible that this single site does not act independently on latency establishment and that there is an interplay of *cis*-regulatory factors during this process. Furthermore, no correlation was observed overall between latency potential and the frequency of nucleotide differences in LTR functional sites or overall DNA distance from the consensus C sequence. It is likely that latency is influenced by a range of complex mechanisms and differences from the consensus may have a positive, negative or negligible effect on viral latency. Further studies are needed to evaluate this.

We hypothesized that highly transcriptionally active LTRs would be more resistant to transcriptional silencing and would therefore be less prone to viral latency. However, no relationship between latency potential and either basal or Tat-induced promoter activity was observed. This finding suggests that Tat interaction may not directly influence the ability of the virus to establish latency however, this conclusion cannot be drawn due to the small sample size of our study. Assuming that *tat* and the TAR loop co-evolve to allow optimal interaction, we attempted to reduce diversity at this level by using a subtype C Tat-expressing plasmid with the *tat* sequence isolated from a CAPRISA 002 patient. However, this *tat* sequence was not matched to any of the participant LTRs, and it is possible that utilising a participant-matched Tat may more accurately represent *in vivo* conditions. Alternatively, using the subtype B Tat expressed in the RGH model may allow for better comparison between activity and latency potential.

A general limitation of our study was the low number of participant LTRs in our panel, resulting in weak statistical power particularly with correlation analyses. The limited number was primarily due to difficulties cloning the LTRs into the RGH vector. Therefore, future studies should aim to further optimize cloning of LTRs into the RGH vector, in order to develop a high-throughput method of generating a large panel of participant-specific LTR clones. This could potentially be achieved by replacing the RF-cloning procedure with a standard restriction cloning via the use of a shuttle vector system.

In addition, many studies on latency have been performed with on-therapy patients. While the composition of the reservoir is still being characterised, studies indicate that viruses in the reservoir resemble those from chronic infection(327). Whereas, LTRs in this study were isolated from acutely infected patients. Therefore, this model can be made more *in vivo* relevant by investigating chronic LTRs, longer term culture of RGH pseudovirus infected samples and the infection of primary cells isolated from donors.

In conclusion, our study provides evidence to suggest that subtype C viruses, can directly and rapidly establish latency within a large proportion of non-resting CD4+ T cells. Inter-participant genotypic variation within the LTR of an infecting virus was shown to have a direct role on the proportion of viruses which established latency, i.e. each LTR had an inherent latency potential. The precise mechanisms of latency remain largely unknown. However, since latency potential was independent of basal and Tat-induced LTR activity, other factors such as regulatory element interaction and the efficiency of recruitment of molecules responsible for establishing latency, such as histone modifiers, may play a role. Identification of these cellular factors, and the early latently infected cells themselves in infected patients, could provide useful insights into targets for treatment and prevention of latency and could contribute to curative and remission strategies.

Appendices

Appendix 1. Reagents and buffers

The materials described hereafter and the methods in which they can be used are described in *Molecular Cloning: A Laboratory Manual*(328).

Tris/Acetate/EDTA (TAE) buffer: (10x 1 Liter) (stored at room temperature)

1. Dissolve 48.4 g Tris and 11.42mL glacial acetic acid in 900 ml distilled water
2. Add 3.722g of Na_2EDTA (pH 8.0)
3. Adjust volume to 1 Liter

Carbenicillin: (100 mg/mL) (stored at -20˚C)

1. Weigh 1000 mg carbenicilin disodium salt (98-100% anhydrous)
2. Dissolve in 10 mL distilled, de-ionized H_2O in a 15 mL tube
3. Vortex to ensure that all the salt is dissolved
4. Using a syringe, pass the solution through a 0.22 µm filter into a fresh, sterile tube

Luria-Bertani Broth: (stored at room temperature)

1. Add the following to 800 mL H_2O:

   10g Bacto-tryptone

   5g yeast extract

   10g NaCl
2. Adjust pH to 7.5 with NaOH or HCl
3. Adjust the volume to 1 L with dH_2O
4. Sterilize by autoclaving at 121˚C for 15 minutes

Luria-Bertani Agar: (stored at 4˚C)

1. Add the following to 800ml H_2O:

    10g Bacto-tryptone

    5g yeast extract

    10g NaCl
2. Adjust pH to 7.5 with NaOH or HCl
3. Add 15g agar and dissolve by heating in the microwave
4. Adjust the volume to 1l with dH_2O
5. Sterilize by autoclaving at 121˚C for 15 minutes

Preparation of agar plates:

1. Prepare LB Agar as above and allow to cool
2. Add 1µl carbenicilin (100 mg/ml as above) per ml of agar using sterile technique
3. Add to sterile petri dishes until the height of the liquid ~1cm thick
4. Allow to set before storing the plate with the lid facing down

Appendix 2. Supplementary methods

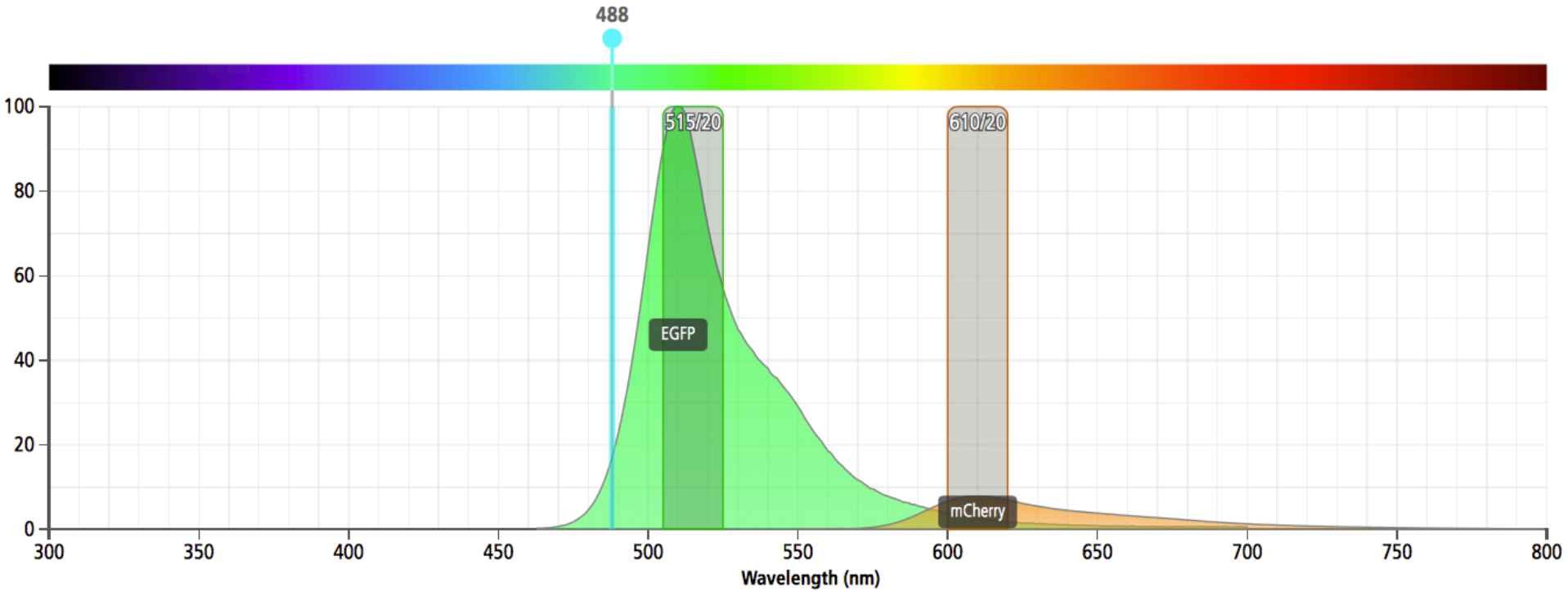


Figure A2.1 Emission spectra of eGFP and mCherry fluorescent proteins excited at 488nm. The 515/20 and 610/20 bandpass filters for detection of eGFP and mCherry respectively, are indicated. Upon excitation at 488nm, the eGFP emission spectra spilt over into the 610/20 filter.

Table A2.1 List of all PCR primers used for this study. Reverse primer HXB2 positions are indicated in red.

Primer Name	Sequencing Direction	Primer Sequence	Length (mer)	HXB2 position	Purpose
^eGFP-pRGH-FP	Forward	5'-AGC CAA AAT TAC CCT ATA GTG CAG AAC CTC-3'	30	1174→1203	C
^eGFP-pRGH-InvPCR(R)	Reverse	5'-GAC CTG GCT GTT GTT TCC TGT GTC AC-3'	26	1148←1173	C
cLTR_F	Forward	5'-AAC TAC ACA CCG GGA CCA GGG-3'	21	87→107	PA/S
gagB-R	Reverse	5'-CTC CCT GAC ATG CTG TCA TCA T-3'	22	1825←1846	CPS/S
gagD-F	Forward	5'-TCT CTA GCA GTG GCG CCC G-3'	19	626→644	S
gagD-R	Reverse	5'-AAT TCC TCC TAT CAT TTT TGG-3'	21	2382←2402	S
GF40	Forward	5'-GAC ACC AAG GAA GCC TTA GA-3'	20	1075→1094	CPS/S
GF80	Forward	5'-AGA GAA CCA AGG GGA AGT GA-3'	20	1477→1493	S
Nef F	Forward	5′-CCT AGA AGA ATA AGA CAG GGC TT-3′	22	8754→8775	CPS
pRGH Colony RP	Reverse	5'-TCA TGT TTG ACA GCT TAT CAT CGC C-3'	25	N/A	CPS/S
pRGH_LTR_F	Forward	5'-CCA TCT CGA GGT GCC TTT AAG ACC AAT GAC-3'	30	9006→9035	PA/CPS/S
pRGH_LTR_R	Reverse	5'-TAC CAA CAG TAC CGG ATT GC-3'	20	N/A	PA
SQ16RC	Reverse	5'-CTT GTC TAG GGC TTC CTT GGT-3'	21	1078←1098	PA/S

C = Cloning

PA = PCR Amplification

S = Sequencing

CPS = Colony PCR Screening

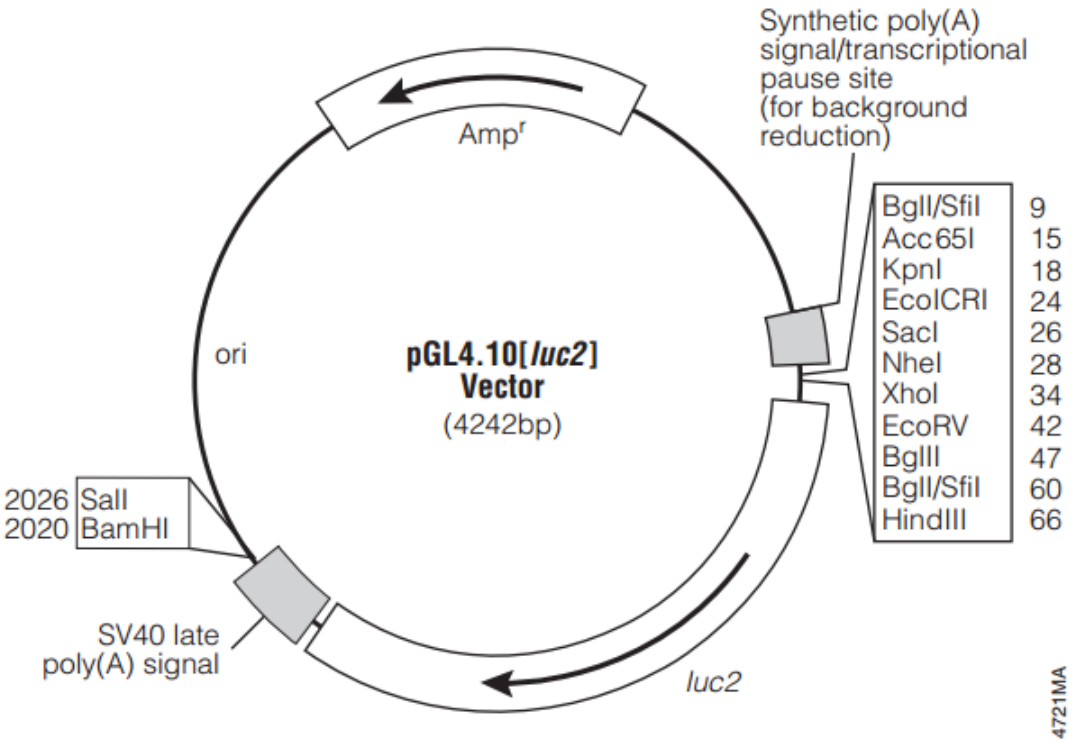


Figure A2.2 pGL4.10[luc2] Vector. Figure obtained at : https://worldwide.promega.com/-/media/files/resources/protocols/product-information-sheets/a/pgl4-10-vector.pdf

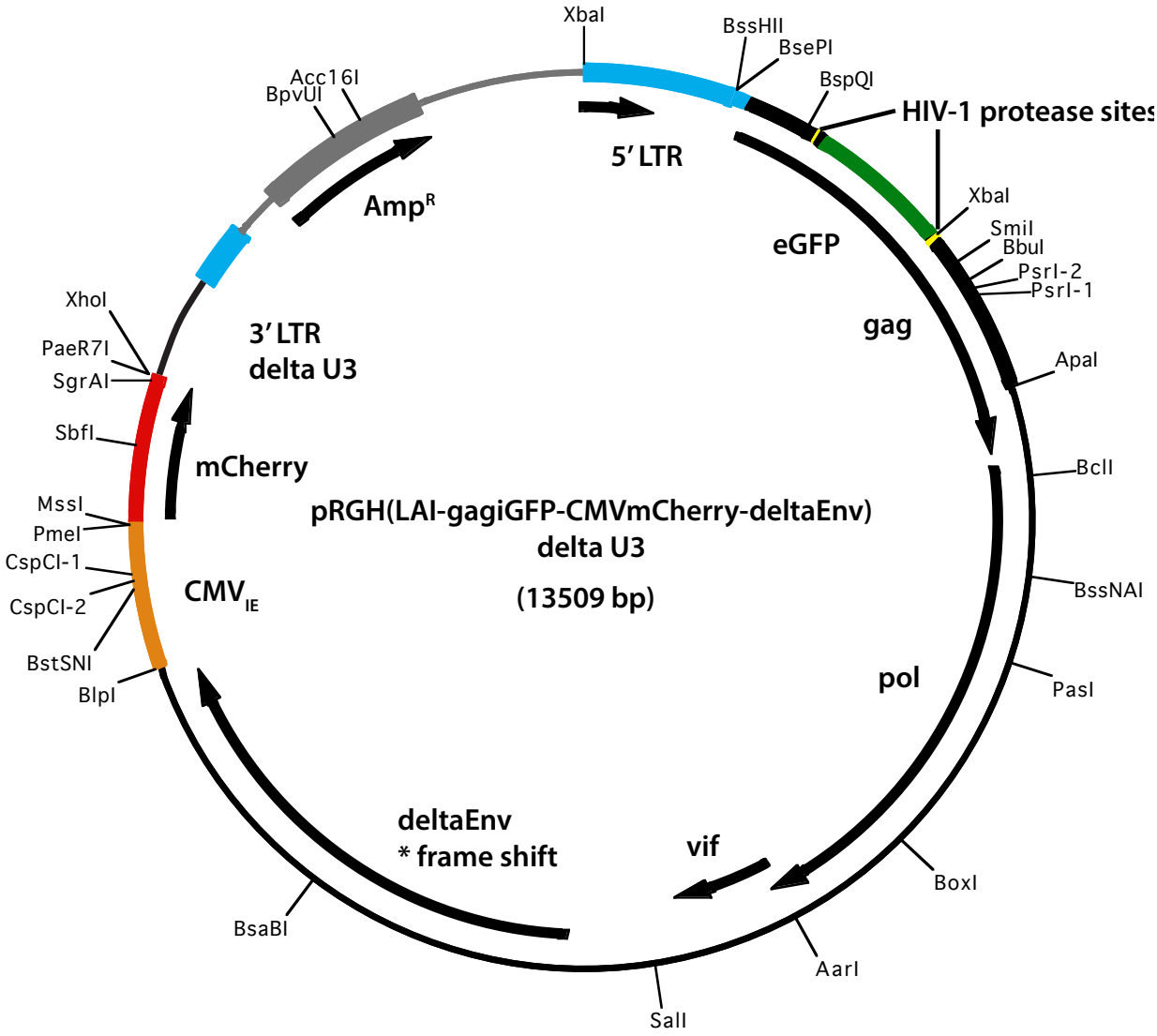


Figure A2.3 Schematic depiction of the 3'LTR-ΔU3 RGH plasmid. Plasmid used for generation of LTR-pseudotyped viruses. pRGH is derived from pLAI full length replication competent, infectious HIV-1 subtype B LAI molecular clone. An eGFP gene has been inserted in-frame between the matrix and capsid regions of the *gag* gene. The *env* open reading frame was disrupted by a single base pair deletion at position 7136 resulting in a frameshift. The *nef* gene has been partially deleted and been replaced with a CMV_{IE} driven mCherry gene construct. The U3 region of the 3' LTR has been deleted. Figure obtained at: http://www.aidsreagent.org/support_docs/12430_map.pdf.

Appendix 3. Supplementary data

Table A3.1 CAPRISA 004 participants selected for inclusion in the LTR panel

PID*	Weeks Post Infection**	Year of Infection	Distance from Consensus C***	Viral Load (copies/ml)	CD4 count (cells/µl)	LTR Cloned Successfully
CAP283	7	2007	0.073	12900	403	Yes
CAP292	5	2008	0.061	458000	510	No
CAP303	4	2008	0.043	344000	469	Yes
CAP306	8	2008	0.063	141000	853	Yes
CAP308	9	2008	0.072	37100	473	No
CAP311	6	2008	0.051	18600	346	Yes
CAP314	12	2008	0.050	166000	358	Yes
CAP315	3	2008	0.058	108000	519	Yes
CAP317	7	2008	0.078	351000	377	No
CAP320	5	2008	0.058	37500	851	Yes
CAP326	3	2008	0.056	31500	470	Yes
CAP327	3	2008	0.057	91600	689	No
CAP343	7	2009	0.056	130000	332	No
CAP345	6	2009	0.081	130000	379	No
CAP355	3	2009	0.048	19700	466	No
CAP360	3	2009	0.051	80600	450	Yes
CAP367	3	2009	0.039	3360000	228	Yes
CAP368	4	2009	0.084	1180	841	No
CAP372	3	2009	0.094	107000	424	Yes

* PID = Participant Identification Number

** Weeks post-infection is estimated based on last negative and first positive HIV-1 test result.

*** Pairwise DNA distance of each participant LTR sequence to a global subtype C LTR (described in section 3.9.1).

Table A3.2 eGFP expression in CEM.NK^{R} CCR5+ cells infected with the CAP283 and CAP306 pseudoviruses one day post-infection. Cells were infected at pseudovirus volumes previously determined to be optimal for each pseudovirus in the Jurkat E6-1 cell line.

PID*	Percentage eGFP one day post-infection (%)
CAP283	4.0
CAP303	2.3

* PID = Participant Identification Number

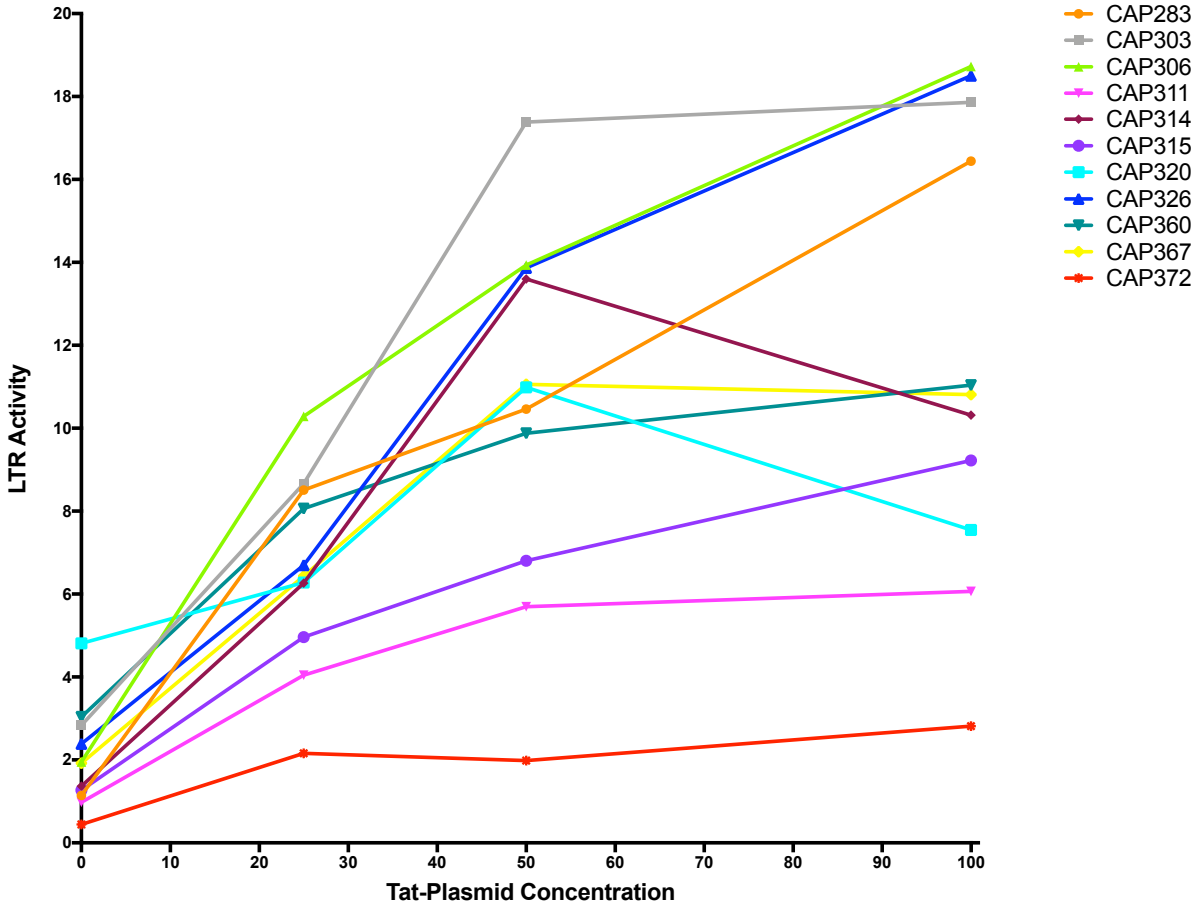


Figure A3.1 Determination of Tat-TOPO plasmid concentration input. HEK293T cells were co-transfected with each pGL4.10 participant LTR plasmid and the Tat-TOPO at a range of concentrations. A *Renilla* luciferase expressing vector was used to normalize transfection efficiency in all cases. LTR activity was measured by luciferase expression after the addition of luciferin at 48 h post-transfection, and is expressed as a response ratio relative to the HIV-1 Bal strain LTR luciferase expression.

Table A3.3 Volumes of each participant pseudovirus used for infections in the latency potential assays.

	Jurkat E6-1		CEM.NKR CCR5+	
PID*	Volume of PSV** (µl)	Percentage eGFP one day post-infection (%)	Volume of PSV** (µl)	Percentage eGFP one day post-infection (%).
CAP283	50	15.2	100	19.5
CAP303	21	8.5	41	16.8
CAP306	50	12.2	150	17.4
CAP311	15	7.2	54	15.3
CAP314	15	6.9	45	13.6
CAP315	18	11.3	25	8.76
CAP320	25	9.7	45	12,0
CAP326	12	6,0	40	10.7
CAP360	24	9.5	52	11.5
CAP367	32	11.7	230	26.5
CAP372	50	12.0	90	16.8

* PID = Participant Identification Number

** PSV = Pseudovirus

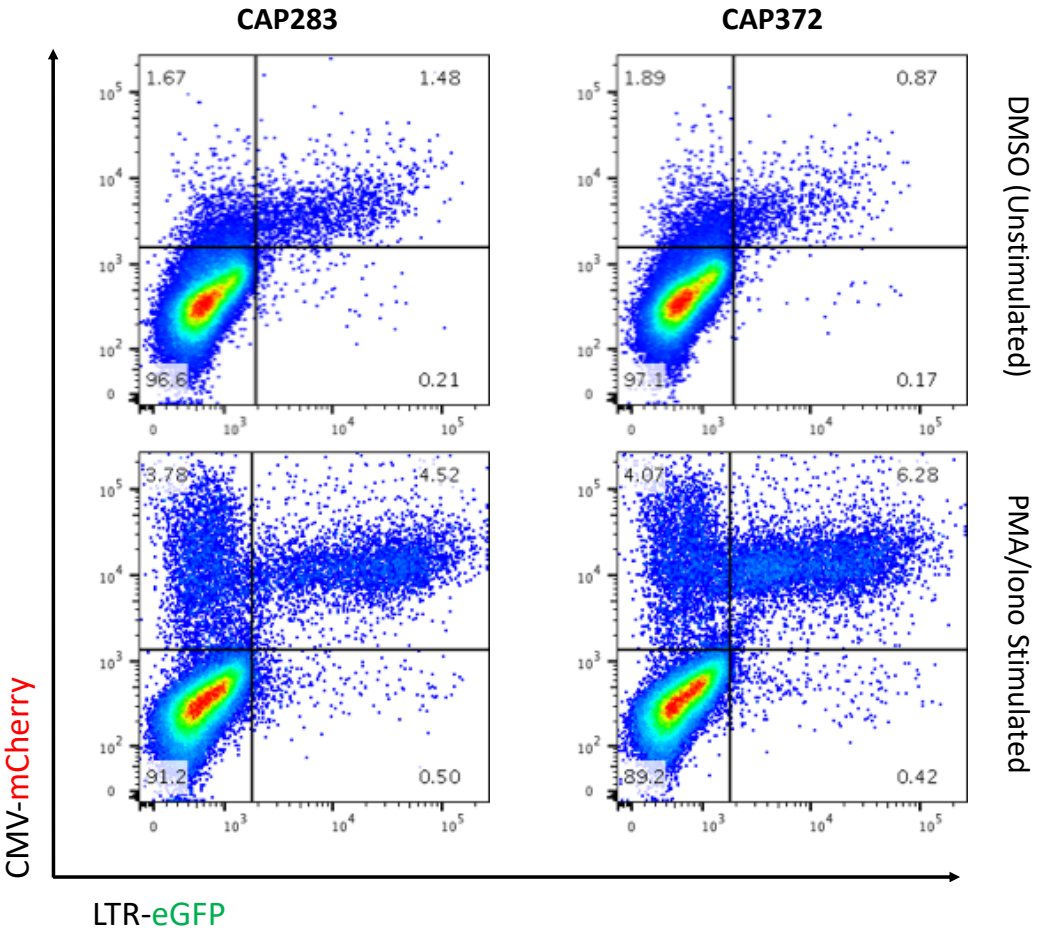


Figure A3.2 Pseudocolour dot plots of populations of Jurkat E6-1 cells infected with the CAP283 and CAP372 RGH pseudoviruses. Jurkat E6-1 cells were infected with each participant RGH variants. At 7 days post-infection, cells were treated with either DMSO or PMA/Iono for 24 hours and subsequently analysed by flow cytometry.

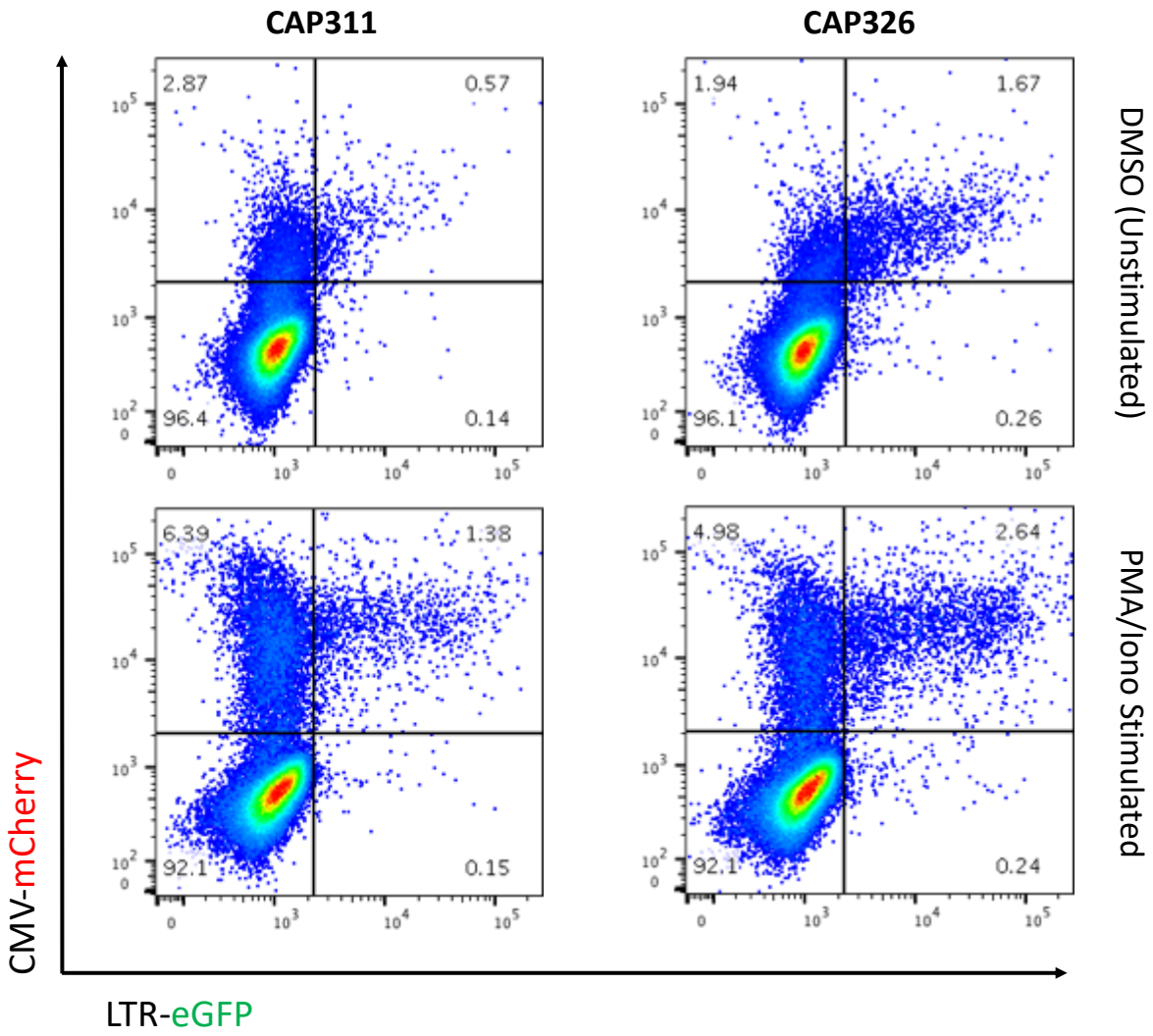


Figure A3.3 Pseudocolour dot plots of populations of CEM.NK^{R} CCR5+ cells infected with the CAP311 and CAP326 RGH pseudoviruses. Jurkat E6-1 cells were infected with each participant RGH variants. At 7 days post-infection, cells were treated with either DMSO or PMA/Iono for 24 hours and subsequently analysed by flow cytometry.

Table A3.4 Basal and Tat-induced LTR activity relative to BaL of each participant LTR

PID*	Basal Activity	Tat-induced activity
CAP283	1.87	10.33
CAP303	2.15	15.83
CAP306	2.40	12.55
CAP311	1.35	5.58
CAP314	1.50	9.66
CAP315	1.38	6.16
CAP320	1.89	9.45
CAP326	2.14	8.92
CAP360	1.86	8.10
CAP367	1.91	7.35
CAP372	0.66	1.61

* PID = Participant Identification Number

References

1. Barré-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C. 1983. Isolation of a T-Lymphotropic Retrovirus from a Patient at Risk for Acquired Immune Deficiency Syndrome (AIDS). Science 220:868–871.
2. Levy JA., Hoffman AD, Kramer SM, Landis JA, Shimbakuro JM, Oshiro LS. 1984. Isolation of Lymphocytopathic Retroviruses from San Francisco Patients with AIDS. Science 225:840–842.
3. 2018. UNAIDS Fact Sheet: South Africa, accessed on 12 March 2018 at: http://www.unaids.org/en/regionscountries/countries/southafrica.
4. Attia S, Egger M, Müller M, Zwahlen M, Low N. 2009. Sexual transmission of HIV according to viral load and antiretroviral therapy: Systematic review and meta-analysis. AIDS 23:1397–1404.
5. Hirsch MS, Conway B, Aquila RTD, Johnson VA, Clotet B, Demeter LM, Hammer SM. 1998. Antiretroviral Drug Resistance Testing in Adults With HIV Infection Implications for Clinical Management. Jama 279:1984–1991.
6. Furtado MR, Callaway DS, Phair JP, Kunstman KJ, Stanton JL, Macken CA, Perelson AS, Wolinsky SM. 1999. Persistence of HIV-1 Transcription In Peripheral-Blood Mononuclear Cells in Patients Receiving Potent Antiretrovial Therapy. N Engl J Med 340:1614–1622.
7. Wong JK, Yukl SA. 2016. Tissue Reservoirs of HIV. Curr Opin HIV AIDS 11:362–370.
8. Kagee A, Remien RH, Berkman A. 2011. Structural barriers to ART adherence in Southern Africa: challenges and potential ways forward. Glob Public Health 6:1–12.
9. De Jong MD, De Boer RJ, De Wolf F, Foudraine NA, Boucher CAB, Goudsmit J, Lange JMA. 1997. Overshoot of HIV-1 viraemia after early discontinuation of antiretroviral treatment. AIDS 11:79–84.
10. García F, Plana M, Vidal C, Cruceta A, O'Brien WA, Pantaleo G, Pumarola T, Gallart T, Miró JM, Gatell JM. 1999. Dynamics of viral load rebound and immunological changes after stopping effective antiretroviral therapy. AIDS 13:79–86.

11. Xing S, Siliciano RF. 2013. Targeting HIV latency: pharmacologic strategies toward eradication. Drug Discov Today 18:541–551.
12. Hütter G, Nowak D, Mossner M, Ganepola S, Müßig A, Allers K, Schneider T, Hofmann J, Kücherer C, Blau O, Blau IW, Hofmann WK, Thiel E. 2009. Long-Term Control of HIV by CCR5 Delta32/ Delta32 Stem-Cell Transplantation. N Engl J Med 360:692–698.
13. Henrich TJ, Hanhauser E, Marty FM, Sirignano MN, Keating S, Lee TH, Robles YP, Davis BT, Li JZ, Heisey A, Hill AL, Busch MP, Armand P, Soiffer RJ, Altfeld M, Kuritzkes DR. 2014. Antiretroviral-Free HIV-1 Remission and Viral Rebound After Allogeneic Stem Cell Transplantation: Report of 2 Cases. Ann Intern Med 161:319–327.
14. Persaud D, Gay H, Ziemniak C, Chen YH, Piatak M, Chun T-W, Strain M, Richman D, Luzuriaga K. 2013. Absence of Detectable HIV-1 Viremia after Treatment Cessation in an Infant. N Engl J Med 369:1828–1835.
15. Sáez-Cirión A, Bacchus C, Hocqueloux L, Avettand-Fenoel V, Girault I, Lecuroux C, Potard V, Versmisse P, Melard A, Prazuck T, Descours B, Guergnon J, Viard JP, Boufassa F, Lambotte O, Goujard C, Meyer L, Costagliola D, Venet A, Pancino G, Autran B, Rouzioux C. 2013. Post-Treatment HIV-1 Controllers with a Long-Term Virological Remission after the Interruption of Early Initiated Antiretroviral Therapy ANRS VISCONTI Study. PLoS Pathog 9:e1003211.
16. Hayden EC. 2015. Teen is healthy 12 years after ending HIV drugs. Nature 523:393.
17. Frange P, Faye A, Avettand-Fenoël V, Bellaton E, Descamps D, Angin M, David A, Caillat-Zucman S, Peytavin G, Dollfus C, Le Chenadec J, Warszawski J, Rouzioux C, Sáez-Cirión A. 2016. HIV-1 virological remission lasting more than 12 years after interruption of early antiretroviral therapy in a perinatally infected teenager enrolled in the French ANRS EPF-CO10 paediatric cohort: A case report. Lancet HIV 3:e49–e54.
18. Cohen J. 2017. What can science learn from a child who has controlled HIV without drugs for more than 8 years? Sci Mag.
19. Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, Cummins LB, Arthur LO, Peeters M, Shaw GM, Sharp PM, Hahn BH. 1999. Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. Nature 397:436–441.

20. Popper SJ, Sarr AD, Travers KU, Guèye-Ndiaye A, Mboup S, Essex ME, Kanki PJ. 1999. Lower Human Immunodeficiency Virus (HIV) Type 2 Viral Load Reflects the Difference in Pathogenicity of HIV-1 and HIV-2. J Infect Dis 180:1116–1121.
21. Keele BF, Van Heuverswyn F, Li Y, Bailes E, Takehisa J, Santiago ML, Bibollet-Ruche F, Chen Y, Wain L V., Liegeois F, Loul S, Ngole EM, Bienvenue Y, Delaporte E, Brookfield JFY, Sharp PM, Shaw GM, Peeters M, Hahn BH. 2006. Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. Science 313:523–526.
22. Van Heuverswyn F, Li Y, Neel C, Bailes E, Keele BF, Liu W, Loul S, Butel C, Liegeois F, Bienvenue Y, Ngolle EM, Sharp PM, Shaw GM, Delaporte E, Hahn BH, Peeters M, Pan C. 2006. Human immunodeficiency viruses: SIV infection in wild gorillas. Nature 444:164.
23. Vallari A, Holzmayer V, Harris B, Yamaguchi J, Ngansop C, Makamche F, Mbanya D, Kaptue L, Ndembi N, Devare S, Brennan CA, Diagnostics A, Park A, Gurtler L, Devare S, Brennan CA. 2011. Confirmation of Putative HIV-1 Group P in Cameroon. J Virol 85:1403–1407.
24. Hemelaar J, Gouws E, Ghys PD, Osmanov S. 2011. Global trends in molecular epidemiology of HIV-1 during 2000-2007. AIDS 25:679–689.
25. HIV Circulating Recombinant Forms (CRFs) accessed on 27 September 2018 at: https://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html.
26. Taylor B, Sobieszczyk M, McCutchan F, Hammer SM. 2008. The Challenge of HIV-1 Subtype Diversity. N Engl J Med 358:1590–1602.
27. Amornkul PN, Karita E, Kamali A, Rida WN, Sanders EJ, Lakhi S, Price MA, Kilembe W, Cormier E, Anzala O, Latka MH, Bekker L-G, Allen SA, Gilmour J, Fast PE. 2013. Disease progression by infecting HIV-1 subtype in a seroconverter cohort in sub-Saharan Africa. AIDS 27:2775–2786.
28. Domingo E, Holland JJ. 1997. RNA Virus Mutations and Fitness For Survival. Annu Rev Microbiol 51:151–178.
29. Ariën KK, Abraha A, Quiñones-mateu ME, Vanham G, Arts EJ, Arie KK, Quin ME, Kestens L. 2005. The Replicative Fitness of Primary Human Immunodeficiency Virus Type 1 (HIV-1) Group The Replicative Fitness of Primary Human Immunodeficiency Virus Type 1 (HIV-1) Group M , HIV-1 Group O , and HIV-2 Isolates † 1:8979–8990.

30. Vasan A, Renjifo B, Hertzmark E, Chaplin B, Msamanga G, Essex M, Fawzi W, Hunter D. 2006. Different rates of disease progression of HIV type 1 infection in Tanzania based on infecting subtype. Clin Infect Dis 42:843–852.
31. Easterbrook PJ, Smith M, Mullen J, O'Shea S, Chrystie I, De Ruiter A, Tatt ID, Geretti AM, Zuckerman M. 2010. Impact of HIV-1 viral subtype on disease progression and response to antiretroviral therapy. J Int AIDS Soc 13:1–9.
32. Venner CM, Nankya I, Kyeyune F, Demers K, Kwok C, Chen P-L, Rwambuya S, Munjoma M, Chipato T, Byamugisha J, Van Der Pol B, Mugyenyi P, Salata RA, Morrison CS, Arts EJ. 2016. Infecting HIV-1 Subtype Predicts Disease Progression in Women of Sub-Saharan Africa. EBioMedicine 13:305–314.
33. Göttlinger HG. 2001. HIV-1 Gag: a Molecular Machine Driving Viral Particle Assembly and Release. HIV Seq Compend 2–28.
34. Jacks T, Power MD, Masiarz FR, Luciw PA, Barr PJ, Varmus HE. 1988. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. Nature 331:280–283.
35. Henderson LE, Bowers MA, Sowder RC, Serabyn SA, Johnson DG, Bess JW, Arthur LO, Bryant DK, Fenselau C. 1992. Gag Proteins of the Highly Replicative MN Strain of Human Immunodeficiency Virus Type 1: Posttranslational Modifications, Proteolytic Processings, and Complete Amino Acid Sequences. J Virol 66:1856–65.
36. Blood GAC. 2016. Human Immunodeficiency Virus (HIV). Transfus Med Hemotherapy 43:203–222.
37. Wyatt R, Kwong PD, Hendrickson WA, Sodroski JG. 1999. Structure of the Core of the HIV-1 gp120 Exterior Envelope Glycoprotein. Los Alamos Rev 1–7.
38. Kao S-Y, Calman AF, Luciw PA, Peterlin BM. 1987. Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. Nature 330:489–493.
39. Addo MM, Altfeld M, Rosenberg ES, Eldridge RL, Philips MN, Habeeb K, Khatri A, Brander C, Robbins GK, Mazzara GP, Goulder PJ, Walker BD. 2001. The HIV-1 regulatory proteins Tat and Rev are frequently targeted by cytotoxic T lymphocytes derived from HIV-1-infected individuals. Proc Natl Acad Sci 98:1781–1786.
40. Karn J. 2000. Tat, a novel regulator of HIV transcription and latency. HIV Seq Compend 2–18.

41. Pollard VW, Malim MH. 1998. The HIV-1 Rev Protein. Annu Rev Microbiol 52:491–532.
42. Fernandes J, Jayaraman B, Frankel A. 2012. The HIV-1 Rev response element. RNA Biol 9:6–11.
43. Yu Q, Landau NR, König R. 2003. Vif and the Role of Antiviral Cytidine Deaminases in HIV-1 Replication. HIV Seq Compend 1–13.
44. Goh WC, Rogel ME, Matthew Kinsey C, Michael SF, Fultz PN, Nowak MA, Hahn BH, Emerman M. 1998. HIV-1 Vpr increases viral expression by manipulation of the cell cycle: A mechanism for selection of Vpr in vivo. Nat Med 4:65–71.
45. Connor RI, Chen BK, Choe S, Landau NR. 1995. Vpr is required for efficient replication of human immunodeficiency virus type-1 in mononuclear phagocytes. Virology.
46. Neil SJD, Zang T, Bieniasz PD. 2008. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. Nature 451:425–430.
47. Willey RL, Maldarelli F, Martin MA, Strebel K. 1992. Human immunodeficiency virus type 1 Vpu protein induces rapid degradation of CD4. J Virol 66:7193–7200.
48. Casartelli N, di Matteo G, Potesta M, Rossi P, Doria M. 2003. CD4 and Major Histocompatibility Complex Class I Downregulation by the Human Immunodeficiency Virus Type 1 Nef Protein in Pediatric AIDS Progression. J Virol 77:11536–11545.
49. Ariën KK, Verhasselt B. 2008. HIV Nef: role in pathogenesis and viral fitness. Curr HIV Res 6:200–208.
50. Piguet V, Trono D. 1999. A Structure-function Analysis of the Nef Protein of Primate Lentiviruses. Rev Med Virol 9:111–120.
51. Sundquist W, Kräusslich H-G. 2012. HIV-1 Assembly, Budding, and Maturation. Cold Spring Harb Perspect Med 2:1–25.
52. Rowland-Jones S. 1999. HIV infection: Where have all the T cells gone? Lancet 354:5–7.
53. McCune JM. 2001. The dynamics of CD4+ T-cell depletion in HIV disease. Nature 410:974–979.
54. Landau NR, Warton M, Littman DR. 1988. The envelope glycoprotein of the human immunodeficiency virus binds to the immunoglobulin-like domain of CD4. Nature 336:403–405.

55. Sattentau QJ, Moore JP. 1991. Conformational changes induced in the human immunodeficiency virus envelope glycoprotein by soluble CD4 binding. J Exp Med 174:407–415.
56. Moore JP, Trkola A, Dragic T. 1997. Co-receptors for HIV-1 entry. Curr Opin Immunol 9:551–562.
57. Doms RW, Moore JP. 2000. HIV-1 membrane fusion: targets of opportunity. J Cell Biol 151:F9-14.
58. Iordanskiy S, Bukrinsky M. 2007. Reverse transcription complex: the key player of the early phase of HIV replication. Future Virol 2:49–64.
59. Roberts JD, Bebenek K, Kunkel TA. 1988. The Accuracy of Reverse Transcriptase. Am Assoc Adv Sci 242:1171–1173.
60. Bukrinsky MI, Sharova N, Dempsey MP, Stanwick TL, Bukrinskaya AG, Haggerty S, Stevenson M. 1992. Active nuclear import of human immunodeficiency virus type 1 preintegration complexes. Proc Natl Acad Sci U S A 89:6580–6584.
61. Bukrinsky M. 2004. A Hard Way to the Nucleus. Mol Med 10:1–5.
62. Laspia MF, Rice AP, Mathews MB. 1989. HIV-1 Tat protein increases transcriptional initiation and stabilizes elongation. Cell 59:283–292.
63. Chou S, Upton H, Bao K, Schulze-Gahmen U, Samelson AJ, He N, Nowak A, Lu H, Krogan NJ, Zhou Q, Alber T. 2013. HIV-1 Tat recruits transcription elongation factors dispersed along a flexible AFF4 scaffold. Proc Natl Acad Sci 110:E123-131.
64. Haase AT. 1986. Pathogenesis of lentivirus infections. Nature 322:130–136.
65. Van Lint C, Bouchat S, Marcello A. 2013. HIV-1 transcription and latency: an update. Retrovirology 10:1–38.
66. Michael NL, D'Arcy L, Ehrenberg PK, Redfield RR. 1994. Naturally Occurring Genotypes of the Human Immunodeficiency Virus Type 1 Long Terminal Repeat Display a Wide Range of Basal and Tat-Induced Transcriptional Activities. J Virol 68:3163–3174.
67. Krebs FC, Hogan TH, Quiterio S, Gartner S, Wigdahl B. 2001. Lentiviral LTR-directed expression, sequence variation, and disease pathogenesis. HIV Seq Compend 29–70.
68. Guntaka R V. 1993. Transcription termination and polyadenylation in retroviruses. Microbiol Rev 57:511–521.

69. Huber HE, Richardson CC. 1990. Processing of the Primer for Plus Strand DNA Synthesis by Human Immunodeficiency Virus 1 Reverse Transcriptase ". J Biol Chem 265:10565–10573.
70. Esposito F, Corona A, Tramontano E. 2012. HIV-1 Reverse Transcriptase Still Remains a New Drug Target: Structure, Function, Classical Inhibitors, and New Inhibitors with Innovative Mechanisms of Actions. Mol Biol Int 2012:1–23.
71. Klaver B, Berkhout B. 1994. Comparison of 5' and 3' long terminal repeat promoter function in human immunodeficiency virus. J Virol 68:3830–3840.
72. Fujinaga K, Cujec TP, Peng J, Garriga J, Price DH, Graña X, Peterlin BM. 1998. The ability of positive transcription elongation factor B to transactivate human immunodeficiency virus transcription depends on a functional kinase domain, cyclin T1, and Tat. J Virol 72:7154–7159.
73. Harrich D, Mavankal G, Mette-snider A, Gaynor RB. 1995. Human immunodeficiency virus type 1 TAR element revertant viruses define RNA structures required for efficient viral gene expression and replication. J V irology 69:4906–4913.
74. Gaynor R. 1992. Cellular transcription factors involved in the regulation of HIV-1 gene expression. AIDS 6:347–363.
75. Pereira LA, Bentley K, Peeters A, Churchill MJ, Deacon NJ. 2000. A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter. Nucleic Acids Res 28:663–668.
76. Lu YC, Stenzel M, Sodroski JG, Haseltine WA. 1989. Effects of Long Terminal Repeat Mutations on Human Immunodeficiency Virus Type-1 Replication. J Virol 63:4115–4119.
77. Jeeninga RE, Hoogenkamp M, Armand-Ugon M, de Baar M, Verhoef K, Berkhout B. 2000. Functional differences between the long terminal repeat transcriptional promoters of human immunodeficiency virus type 1 subtypes A through G. J Virol 74:3740–51.
78. Opijnen T Van, Jeeninga RE, Boerlijst MC, Pollakis GP, Zetterberg V, Salminen M, Berkhout B. 2004. Human Immunodeficiency Virus Type 1 Subtypes Have a Distinct Long Terminal Repeat That Determines the Replication Rate in a Host-Cell-Specific Manner. J Virol 78:3675–3683.
79. Colin L, Van Lint C. 2009. Molecular control of HIV-1 postintegration latency: Implications for the development of new therapeutic strategies. Retrovirology

6:1–29.

80. Rittner K, Churcher MJ, Gait MJ, Karn J. 1995. The human immunodeficiency virus long terminal repeat includes a specialised initiator element which is required for tat-responsive transcription. J Mol Biol 248:562–580.
81. Jones KA, Kadonaga JT, Luciw PA, Tjian R. 1986. Activation of the AIDS retrovirus promoter by the cellular transcription factor, Sp1. Science 232:755–759.
82. Marban C, Redel L, Suzanne S, Van Lint C, Lecestre D, Chasserot-Golaz S, Leid M, Aunis D, Schaeffer E, Rohr O. 2005. COUP-TF interacting protein 2 represses the initial phase of HIV-1 gene transcription in human microglial cells. Nucleic Acids Res 33:2318–2331.
83. Marban C, Suzanne S, Dequiedt F, De Walque S, Redel L, Van Lint C, Aunis D, Rohr O, Walque D, Redel L, Lint C Van, Aunis D, Rohr O. 2007. Recruitment of chromatin-modifying enzymes by CTIP2 promotes HIV-1 transcriptional silencing. EMBO J 26:412–423.
84. Suñé C, García-Blanco MA. 1995. Sp1 transcription factor is required for in vitro basal and Tat-activated transcription from the human immunodeficiency virus type 1 long terminal repeat. J Virol 69:6572–6576.
85. Ou S-HI, Garcia-Martinez LF, Paulssen EJ, Gaynor RB. 1994. Role of flanking E box motifs in human immunodeficiency virus type 1 TATA element function. J Virol 68:7188–7199.
86. Imai K, Okamoto T. 2006. Transcriptional repression of human immunodeficiency virus type 1 by AP-4. J Biol Chem 281:12495–12505.
87. Naghavi MH, Schwartz S, Sonnerborg A, Vahlne A. 1999. Long terminal repeat promoter/enhancer activity of different subtypes of HIV type 1. AIDS Res Hum Retroviruses 15:1293–1303.
88. Hiscott J, Kwon H, Génin P. 2001. Hostile takeovers: Viral appropriation of the NF-κB pathway. J Clin Invest 107:143–151.
89. Bachu M, Yalla S, Asokan M, Verma A, Neogi U, Sharma S, Murali R V., Mukthey AB, Bhatt R, Chatterjee S, Rajan RE, Cheedarla N, Yadavalli VS, Mahadevan A, Shankar SK, Rajagopalan N, Shet A, Saravanan S, Balakrishnan P, Solomon S, Vajpayee M, Shivappa K, Kundu TK, Jeang KT, Ranga U. 2012. Multiple NF-κB sites in HIV-1 subtype C long terminal repeat confer superior magnitude of transcription and thereby the enhanced viral predominance. J Biol

Chem 287:44714–44735.

90. Nabel G, Baltimore D. 1987. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. Nature 326:711–713.
91. Sheppard K-A, Rose DW, Haque ZK, Kurokawa R, McInerney E, Westin S, Thanos D, Rosenfeld MG, Glass CK, Collins T. 1999. Transcriptional Activation by NF-κB Requires Multiple Coactivators. Mol Cell Biol 19:6367–6378.
92. Perkins ND, Edwards NL, Duckett CS, Agranoff AB, Schmid RM, Nabel GJ. 1993. A cooperative interaction between NF-κB and Sp1 is required for HIV-1 enhancer activation. EMBO J 12:3551–3558.
93. Pazin M.J., Sheridan P.L., Cannon K, Cao Z., Keck J.G., Kadonaga J.T., Jones KA. 1996. Nf-κB-mediated chromatin reconfiguration and transcriptional activation of the HIV-1 enhancer in vitro. Genes Dev 10:37–49.
94. Kim YK, Bourgeois CF, Pearson R, Tyagi M, West MJ, Wong J, Wu S, Chiang C, Karn J. 2006. Recruitment of TFIIH to the HIV LTR is a rate- limiting step in the emergence of HIV from latency. EMBO J 25:3596–3604.
95. Gerritsen ME, Williams AJ, Neish AS, Moore S, Shi Y, Collins T. 1997. CREB-binding protein/p300 are transcriptional coactivators of p65. Proc Natl Acad Sci U S A 94:2927–2932.
96. Lusic M, Marcello A, Cereseto A, Giacca M. 2003. Regulation of HIV-1 gene expression by histone acetylation and factor recruitment at the LTR promoter. EMBO J 22:6550–6561.
97. Ott M, Schnölzer M, Garnica J, Fischle W, Emiliani S, Rackwitz H-R, Verdin E. 1999. Acetylation of the HIV-1 tat protein by p300 is important for its transcriptional activity. Curr Biol 9:1489–1492.
98. McCaffrey PG, Jain J, Jamieson C, Sen R, Rao A. 1992. A T cell Nuclear factor ressembling NF-AT binds to an NF-κB site and to the conserved lymphokine promoter sequence ‘Cytokine 1’. J Biol Chem 267:1864–1871.
99. Cron RQ, Bartz SR, Clausell A, Bort SJ, Klebanoff SJ, Lewis DB. 2000. NFAT1 enhances HIV-1 gene expression in primary human CD4 T cells. Clin Immunol 94:179–191.
100. Garcia-Rodriguez C, Rao A. 1998. Nuclear Factor of Activated T Cells (NFAT)-dependent Transactivation Regulated by the Coactivators p300/CREB-binding Protein (CBP). J Exp Med 187:2031–2036.
101. Duverger A, Wolschendorf F, Zhang M, Wagner F, Hatcher B, Jones J, Cron

RQ, van der Sluis RM, Jeeninga RE, Berkhout B, Kutsch O. 2013. An AP-1 binding site in the enhancer/core element of the HIV-1 promoter controls the ability of HIV-1 to establish latent infection. J Virol 87:2264–77.

102. Tesmer VM, Rajadhyaksha A, Babin J, Bina M. 1993. NF-IL6-Mediated Transcriptional Activation of the Long Terminal Repeat of the Human Immunodeficiency Virus Type 1. Proc Natl Acad Sci United States Am 90:7298–7302.
103. Zacharova V, Becker MLB, Zachar V, Ebbesen P, Goustin AS. 1997. DNA sequence analysis of the long terminal repeat of the C subtype of human immunodeficiency virus type 1 from Southern Africa reveals a dichotomy between B subtype and African subtypes on the basis of upstream NF-IL6 motif. AIDS Res Hum Retroviruses 13:719–724.
104. Liu Y, Nonnemacher MR, Stauff DL, Li L, Banerjee A, Irish B, Kilareski E, Rajagopalan N, Suchitra JB, Khan ZK, Ranga U, Wigdahl B. 2010. Structural and functional studies of CCAAT/enhancer binding sites within the human immunodeficiency virus type 1 subtype C LTR. Biomed Pharmacother 64:672–680.
105. Ruocco MR, Chen X, Ambrosino C, Dragonetti E, Liu W, Mallardo M, de Falco G, Palmieri C, Franzoso G, Quinto I, Venuta S, Scala G. 1996. Regulation of HIV-1 Long Terminal Repeats by Interaction of C/EBP (NF-IL6) and NF-κB/Rel Transcription Factors. J Biol Chem 271:22479–22486.
106. Henderson AJ, Zou X, Calame KL. 1995. C/EBP Proteins Activate Transcription from the Human Immunodeficiency Virus Type 1 Long Terminal Repeat in Macrophages/Monocytes. J Virol 69:5337–5344.
107. Henderson AJ, Calame KL. 1997. CCAAT/enhancer binding protein (C/EBP) sites are required for HIV-1 replication in primary macrophages but not CD4(+) T cells. Proc Natl Acad Sci U S A 94:8714–8719.
108. Mukerjee R, Sawaya BE, Khalili K, Amini S. 2007. Association of p65 and C/EBPβ with HIV-1 LTR modulates transcription of the viral promoter. J Cell Biochem 100:1210–1216.
109. Sheridan PL, Sheline CT, Cannon K, Voz ML, Pazin MJ, Kadonaga JT, Jones KA. 1995. Activation of the HIV-1 enhancer by the LEF-1 HMG protein on nucleosome-assembled DNA in vitro. Genes Dev 9:2090–2104.
110. Giese K, Kingsley C, Kirshner JR, Grosschedl R. 1995. Assembly and function

of a TCRα enhancer complex is dependent on LEF-1 induced DNA bending and multiple protein-protein interactions. Genes Dev 9:995–1008.

111. Sieweke MH, Tekotte H, Jarosch U, Graf T. 1998. Cooperative interaction of Ets-1 with USF-1 required for HIV-1 enhancer activity in T cells. EMBO J 17:1728–1739.
112. Jones KA, Peterlin BM. 1994. Control Of RNA Elongation At The HIV-1 Promoter. Annu Rev Biochem 63:717–743.
113. Estable MC, Bell B, Merzouki A, Montaner JSG, O'Shaughnessy M V., Sadowski IJ. 1996. Human Immunodeficiency Virus Type 1 Long Terminal Repeat Variants from 42 Patients Representing All Stages of Infection Display a Wide Range of Sequence Polymorphism and Transcription Activity. J Virol 70:4053–4062.
114. d'Adda di Fagagna FD, Marzio G, Gutierrez MI, Kang LY, Falaschi A, Giacca M. 1995. Molecular and functional interactions of transcription factor USF with the long terminal repeat of human immunodeficiency virus type 1. J Virol 69:2765–2775.
115. Bentley K, Deacon N, Sonza S, Zeichner S, Churchill M. 2004. Mutational analysis of the HIV-1 LTR as a promoter of negative sense transcription. Arch Virol 149:2277–2294.
116. Yang H-C, Shen L, Siliciano RF, Pomerantz JL. 2009. Isolation of a cellular factor that can reactivate latent HIV-1 without T cell activation. Proc Natl Acad Sci 106:6321–6326.
117. Chen J, Malcolm T, Estable MC, Roeder RG, Sadowski I. 2005. TFII-I Regulates Induction of Chromosomally Integrated Human Immunodeficiency Virus Type 1 Long Terminal Repeat in Cooperation with USF. J Virol 79:4396–4406.
118. Estable MC, Hirst M, Bell B, O'Shaughnessy M V., Sadowski I. 1999. Purification of RBF-2, a transcription factor with specificity for the most conserved cis-element of naturally occurring HIV-1 LTRs. J Biomed Sci 6:320–332.
119. Malcolm T, Kam J, Pour PS, Sadowski I. 2008. Specific interaction of TFII-I with an upstream element on the HIV-1 LTR regulates induction of latent provirus. FEBS Lett 582:3903–3908.
120. Sadowski I, Mitchell DA. 2005. TFII-I and USF (RBF-2) regulate Ras/MAPK-responsive HIV-1 transcription in T cells. Eur J Cancer 41:2528–2536.
121. Shaw JP, Utz PJ, Durand DB, Toole JJ, Emmel EA, Crabtree GR. 1998.

Identification of a putative regulator of early T cell activation genes. Science 241:202–205.

122. Tsytsykova AV, Goldfeld AE. 2000. Nuclear Factor of Activated T Cells Transcription Factor NFATp Controls Superantigen-induced Lethal Shock. J Exp Med 192:581–586.
123. Markovitz DM, Hannibal MC, Smith MJ, Cossman R, Nabel GJ. 1992. Activation of the Human Immunodeficiency Virus Type 1 Enhancer Is Not Dependent on NFAT-1. J Virol 66:3961–3965.
124. Ghosh D. 1992. Glucocorticoid Receptor-Binding Site in the Human Immunodeficiency Virus Long Terminal Repeat. J Virol 66:586–590.
125. Mitra D, Sikder SK, Laurence J. 1995. Role of glucocorticoid receptor binding sites in the human immunodeficiency virus type 1 long terminal repeat in steroid-mediated suppression of HIV gene expression. Virology 214:512–521.
126. Dasgupta P, Saikumar P, Reddy CD, Reddy EP. 1990. Myb protein binds to human immunodeficiency virus 1 long terminal repeat (LTR) sequences and transactivates LTR-mediated transcription. Proc Natl Acad Sci United States Am 87:8090–8094.
127. Churchill MJ, Ramsay RG, Rhodes DI, Deacon NJ. 2001. c-Myb influences HIV type 1 gene expression and virus production. AIDS Res Hum Retroviruses 17:1481–1488.
128. Stern JB, Smith KA. 1986. Interleukin-2 induction of T-cell G1 progression and c-myb expression. Science 233:203–206.
129. Pereira LA, Churchill MJ, Elefanty AG, Gouskos T, Lambert PF, Ramsay RG, Deacon NJ. 2002. Characterization of interactions between transcription factors and a regulatory region spanning nt -320 to -281 of the HIV-1 LTR in T-lymphoid and non-T-lymphoid cells. J Biomed Sci 9:68–81.
130. Sawaya BE, Rohr O, Aunis D, Schaeffer E. 1996. Regulation of human immunodeficiency virus type 1 gene transcription by nuclear receptors in human brain cells. J Biol Chem 271:22895–22900.
131. Canonne-Hergaux F, Aunis D, Schaeffer E. 1995. Interactions of the transcription factor AP-1 with the long terminal repeat of different human immunodeficiency virus type 1 strains in Jurkat, glial, and neuronal cells. J Virol 69:6634–42.
132. Schwartz C, Catez P, Rohr O, Lecestre D, Aunis D, Schaeffer E. 2000.

Functional interactions between C/EBP, Sp1, and COUP-TF regulate human immunodeficiency virus type 1 gene transcription in human brain cells. J Virol 74:65–73.

133. Rohr O, Aunis D, Schaeffer E. 1997. COUP-TF and Sp1 interact and cooperate in the transcriptional activation of the human immunodeficiency virus type 1 long terminal repeat in human microglial cells. J Biol Chem 272:31149–31155.
134. Romerio F, Gabriel MN, Margolis DM. 1997. Repression of Human Immunodeficiency Virus Type 1 through the Novel Cooperation of Human Factors YY1 and LSF. J Virol 71:9375–9382.
135. Coull JJ, Romerio F, Sun J-M, Volker JL, Galvin KM, Davie JR, Shi Y, Hansen U, Margolis DM. 2000. The Human Factors YY1 and LSF Repress the Human Immunodeficiency Virus Type 1 Long Terminal Repeat via Recruitment of Histone Deacetylase 1. J Virol 74:6790–6799.
136. He G, Margolis DM. 2002. Counterregulation of chromatin deacetylation and histone deacetylase occupancy at the integrated promoter of human immunodeficiency virus type 1 (HIV-1) by the HIV-1 repressor YY1 and HIV-1 activator Tat. Mol Cell Biol 22:2965–2973.
137. El Kharroubi A, Verdin E. 1994. Protein-DNA interactions within DNase I-hypersensitive sites located downstream of the HIV-1 promoter. J Biol Chem 269:19916–19924.
138. Dahiya S, Liu Y, Nonnemacher MR, Dampier W, Wigdahl B. 2014. CCAAT enhancer binding protein and nuclear factor of activated T cells regulate HIV-1 LTR via a novel conserved downstream site in cells of the monocyte-macrophage lineage. PLoS One 9:e88116.
139. El Kharroubi A, Martin MA. 1996. cis-Acting Sequences Located Downstream of the Human Immunodeficiency Virus Type 1 Promoter Affect Its Chromatin Structure and Transcriptional Activity. Mol Cell Biol 16:2958–2966.
140. Van Lint C, Amella CA, Emiliani S, John M, Jie T, Verdin E. 1997. Transcription factor binding sites downstream of the human immunodeficiency virus type 1 transcription start site are important for virus infectivity. J Virol 71:6113–6127.
141. Romanchikova N, Ivanova V, Scheller C, Jankevics E, Jassoy C, Serfling E. 2003. NFAT transcription factors control HIV-1 expression through a binding site downstream of TAR region. Immunobiology 208:361–365.
142. Folks T, Powell DM, Lightfoote MM, Benn S, Martin MA, Fauci AS. 1986.

Induction of HTLV-III/LAV from a Nonvirus-Producing T-Cell Line: Implications for Latency. Science 231:600–602.

143. Stevenson M, Stanwick TL, Dempsey MP, Lamonica CA. 1990. HIV-1 replication is controlled at the level of T cell activation and proviral integration. EMBO J 9:1551–1560.
144. Ho DD, U. NA, Perelson AS, Chen W, Leonard JM, Markowitz M. 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature 373:123–126.
145. Wei X, Ghosh SK, Taylor ME, Johnson VA, Emini EA, Deutsch P, Lifson JD, Bonhoeffer S, Nowak MA, Hahn BH, Saag MS, Shaw GM. 1995. Viral dynamics in human immunodeficiency virus type 1 infection. Nature 373:117–122.
146. Chun TW, Finzi D, Margolick J, Chadwick K, Schwartz D, Siliciano RF. 1995. In vivo fate of HIV-1-infected T cells: Quantitative analysis of the transition to stable latency. Nat Med 1:1284–1290.
147. Siliciano RF, Greene WC. 2011. HIV Latency. Cold Spring Harb Perspect Med 1:1–19.
148. Vatakis DN, Kim S, Kim N, Chow SA, Zack JA. 2009. Human Immunodeficiency Virus Integration Efficiency and Site Selection in Quiescent CD4+ T Cells. J Virol 83:6222–6233.
149. Cameron P. U., Saleh S., Sallmann G., Solomon A., Wightman F., Evans V. A., Boucher G., Haddad E. K., Sekaly R. P., Harman A. N., Anderson J. L., Jones K. L., Mak J., Cunningham A. L., Jaworowski A., Lewin S. R. 2010. Establishment of HIV-1 latency in resting CD4+ T cells depends on chemokine-induced changes in the actin cytoskeleton. Proc Natl Acad Sci 107:16934–16939.
150. Pace MJ, Graf EH, Agosto LM, Mexas AM, Male F, Brady T, Bushman FD, O'Doherty U. 2012. Directly infected resting CD4+T cells can produce HIV Gag without spreading infection in a model of HIV latency. PLoS Pathog 8:e1002818.
151. Chavez L, Calvanese V, Verdin E. 2015. HIV Latency Is Established Directly and Early in Both Resting and Activated Primary CD4 T Cells. PLoS Pathog 11:1–21.
152. Chun T-W, Stuyver L, Mizell SB, Ehler LA, Mican JAM, Baseler M, Lloyd AL, Nowak MA, Fauci AS, Sciences M. 1997. Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. Proc Natl Acad Sci U S A

94:13193–13197.

153. Chun T-W, Engel D, Berrey MM, Shea T, Corey L, Fauci AS. 1998. Early establishment of a pool of latently infected, resting CD4+ T cells during primary HIV-1 infection. Proc Natl Acad Sci U S A 95:8869–8873.
154. Whitney JB, Hill AL, Sanisetty S, Penaloza-MacMaster P, Liu J, Shetty M, Parenteau L, Cabral C, Shields J, Blackmore S, Smith JY, Brinkman AL, Peter LE, Mathew SI, Smith KM, Borducchi EN, Rosenbloom DIS, Lewis MG, Hattersley J, Li B, Hesselgesser J, Geleziunas R, Robb ML, Kim JH, Michael NL, Barouch DH. 2014. Rapid seeding of the viral reservoir prior to SIV viraemia in rhesus monkeys. Nature 512:1–4.
155. Zack JA, Arrigo SJ, Weitsman SR, Go AS, Haislip A, Chen ISY. 1990. HIV-1 entry into quiescent primary lymphocytes: Molecular analysis reveals a labile, latent viral structure. Cell 61:213–222.
156. Dahabieh MS, Ooms M, Simon V, Sadowski I. 2013. A Doubly Fluorescent HIV-1 Reporter Shows that the Majority of Integrated HIV-1 Is Latent Shortly after Infection. J Virol 87:4716–4727.
157. Matsuda Y, Kobayashi-Ishihara M, Fujikawa D, Ishida T, Watanabe T, Yamagishi M. 2015. Epigenetic Heterogeneity in HIV-1 Latency Establishment. Sci Rep 5:1–11.
158. Ho Y-C, Siliciano JD. 2015. Efforts to eliminate the latent reservoir in resting CD4+ T cells: strategies for curing HIV-1 infection. J Virus Erad 1:229.
159. Hermankova M, Siliciano JD, Zhou Y, Monie D, Chadwick K, Margolick JB, Quinn TC, Siliciano RF. 2003. Analysis of Human Immunodeficiency Virus Type 1 Gene Expression in Latently Infected Resting CD4+ T Lymphocytes In Vivo. J Virol 77:7383–7392.
160. Siliciano JD, Kajdas J, Finzi D, Quinn TC, Chadwick K, Margolick JB, Kovacs C, Gange SJ, Siliciano RF, Cd T, Siliciano JD, Kajdas J, Finzi D, Quinn TC, Chadwick K, Margolick JB, Kovacs C, Stephen J, Siliciano RF, Hiv- L. 2003. Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. Nat Med 9:727–728.
161. Finzi D, Blankson JN, Siliciano JD, Margolick JB, Chadwick K, Pierson T, Smith K, Lisziewicz J, Rosenberg E, Walker B, Gange S, Gallant J, Siliciano RF. 1999. Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. Nat Med 5:512–517.

162. Harrigan PR, Whaley M, Montaner JSG. 1999. Rate of HIV-1 RNA rebound upon stopping antiretroviral therapy. AIDS 13:F59–F62.
163. Stöhr W, Fidler S, McClure M, Weber J, Cooper D, Ramjee G, Kaleebu P, Tambussi G, Schechter M, Babiker A, Phillips RE, Porter K, Frater J. 2013. Duration of HIV-1 Viral Suppression on Cessation of Antiretroviral Therapy in Primary Infection Correlates with Time on Therapy. PLoS One 8:e78287.
164. Ho Y, Shan L, Hosmane NN, Wang J, Laskey SB, Daniel IS, Lai J, Blankson JN, Siliciano JD, Siliciano RF. 2013. Replication-competent non-induced proviruses in the latent reservoir increase barrier to HIV-1 cure. Cell 155:540–551.
165. Mitchell RS, Beitzel BF, Schroder ARW, Shinn P, Chen H, Berry CC, Ecker JR, Bushman FD. 2004. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS Biol 2:1127–1137.
166. Lewinski MK, Yamashita M, Emerman M, Ciuffi A, Marshall H, Crawford G, Collins F, Shinn P, Leipzig J, Hannenhalli S, Berry CC, Ecker JR, Bushman FD. 2006. Retroviral DNA integration: Viral and cellular determinants of target-site selection. PLoS Pathog 2:0611–0622.
167. Han Y, Lassen K, Monie D, Sedaghat AR, Shimoji S, Liu X, Pierson TC, Margolick JB, Siliciano RF, Siliciano JD. 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 78:6122–6133.
168. Han Y, Lin YB, An W, Xu J, Yang H, Connell KO, Dordai D, Boeke JD, Siliciano JD, Siliciano RF. 2008. Orientation-dependent Regulation of Integrated HIV-1 Expression by Host Gene Transcriptional Readthrough. Cell Host Microbe 4:134–146.
169. Shan L, Yang HH-C, Rabi SA, Bravo HC, Shroff NS, Irizarry RA, Zhang H, Margolick JB, Siliciano JD, Siliciano RF. 2011. Influence of Host Gene Transcription Level and Orientation on HIV-1 Latency in a Primary-Cell Model. J Virol 85:5384–5393.
170. Greger IH, Demarchi F, Giacca M, Proudfoot NJ. 1998. Transcriptional interference perturbs the binding of Sp1 to the HIV-1 promoter. Nucleic Acids Res 26:1294–1300.
171. Crampton N, Bonass WA, Kirkham J, Rivetti C, Thomson NH. 2006. Collision events between RNA polymerases in convergent transcription studied by atomic

force microscopy. Nucleic Acids Res 34:5416–5425.
172. Hobson DJ, Wei W, Steinmetz LM, Svejstrup JQ. 2012. RNA Polymerase II Collision Interrupts Convergent Transcription. Mol Cell 48:365–374.
173. Martinez J, Patkaniowska A, Urlaub H, Lührmann R, Tuschl T, Lu R, Tuschl T. 2002. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell 110:563–574.
174. Morris K V., Chan SW, Jacobsen SE. 2004. Small Interfering RNA-Induced Transcriptional Gene Silencing in Human Cells. Science 305:1289–1293.
175. Hu WY, Bushman FD, Siva AC. 2004. RNA interference against retroviruses. Virus Res 102:59–64.
176. Suzuki K, Shijuuku T, Fukamachi T, Zaunders J, Guillemin G, Cooper D, Kelleher A. 2005. Prolonged transcriptional silencing and CpG methylation induced by siRNAs targeted to the HIV-1 promoter region. J RNAi Gene Silenc 1:66–78.
177. Jacque J-M, Triques K, Stevenson M. 2002. Control of HIV-1 replication by RNA interference. Nature 418:435–438.
178. Bisgrove D, Lewinski M, Bushman F, Verdin E. 2005. Molecular mechanisms of HIV-1 proviral latency. Expert Rev Anti Infect Ther 3:805–814.
179. Darcis G, Van Driessche B, Van Lint C. 2017. HIV Latency: Should We Shock or Lock? Trends Immunol 38:217–228.
180. Perkins KJ, Proudfoot NJ. 2008. An Ungracious Host for an Unwelcome Guest. Cell Host Microbe 4:89–91.
181. Eissenberg JC, Louis SS, Louis SS, Elgin S. 2014. Heterochromatin and Euchromatin. eLS 1–7.
182. Tamaru H. 2010. Confining euchromatin/heterochromatin territory: jumonji crosses the line. Genes Dev 24:1465–1478.
183. Felsenfeld G, Groudine M. 2003. Controlling the double helix. Nature 421:448–453.
184. Verdin E, Paras P, Van Lint C. 1993. Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. EMBO J 12:3249–3259.
185. Van Lint C, Emiliani S, Ott M, Verdin E. 1996. Transcriptional activation and chromatin remodeling of the HIV-1 promoter in response to histone acetylation. EMBO J 15:1112–1120.

186. Yang X-J, Seto E. 2007. HATs and HDACs: From structure, function and regulation to novel strategies for therapy and prevention. Oncogene 26:5310–5318.
187. Imai K, Togami H, Okamoto T. 2010. Involvement of histone H3 lysine 9 (H3K9) methyltransferase G9a in the maintenance of HIV-1 latency and its reactivation by BIX01294. J Biol Chem 285:16538–16545.
188. Williams SA, Chen LF, Kwon H, Ruiz-Jarabo CM, Verdin E, Greene WC. 2006. NF-κB p50 promotes HIV latency through HDAC recruitment and repression of transcriptional initiation. EMBO J 25:139–149.
189. Tyagi M, Karn J. 2007. CBF-1 promotes transcriptional silencing during the establishment of HIV-1 latency. EMBO J 26:4985–4995.
190. Friedman J, Cho W-K, Chu CK, Keedy KS, Archin NM, Margolis DM, Karn J. 2011. Epigenetic Silencing of HIV-1 by the Histone H3 Lysine 27 Methyltransferase Enhancer of Zeste 2. J Virol 85:9078–9089.
191. du Chéné I, Basyuk E, Lin YL, Triboulet R, Knezevich A, Chable-Bessia C, Mettling C, Baillat V, Reynes J, Corbeau P, Bertrand E, Marcello A, Emiliani S, Kiernan R, Benkirane M. 2007. Suv39H1 and HP1γ are responsible for chromatin-mediated HIV-1 transcriptional silencing and post-integration latency. EMBO J 26:424–435.
192. Cheutin T, McNairn AJ, Jenuwein T, Gilbert DM, Singh PB, Misteli T. 2003. Maintenance of stable heterochromatin domains by dynamic HP1 binding. Science 299:721–725.
193. Blazkova J, Trejbalova K, Gondois-Rey F, Halfon P, Philibert P, Guiguen A, Verdin E, Olive D, Van Lint C, Hejnar J, Hirsch I. 2009. CpG methylation controls reactivation of HIV from latency. PLoS Pathog 5:e1000554.
194. Kauder SE, Bosque A, Lindqvist A, Planelles V, Verdin E. 2009. Epigenetic regulation of HIV-1 latency by cytosine methylation. PLoS Pathog 5:e1000495.
195. Cheng X, Blumenthal RM. 2008. Mammalian DNA Methyltransferases: A Structural Perspective. Structure 16:341–350.
196. Palacios JA, Perez-Pinar T, Toro C, Sanz-Minguela B, Moreno V, Valencia E, Gomez-Hernando C, Rodes B. 2012. Long-Term Nonprogressor and Elite Controller Patients Who Control Viremia Have a Higher Percentage of Methylation in Their HIV-1 Proviral Promoters than Aviremic Patients Receiving Highly Active Antiretroviral Therapy. J Virol 86:13081–13084.

197. Blazkova J, Murray D, Justement JS, Funk EK, Nelson AK, Moir S, Chun T-W, Fauci AS. 2012. Paucity of HIV DNA Methylation in Latently Infected, Resting CD4+ T Cells from Infected Individuals Receiving Antiretroviral Therapy. J Virol 86:5390–5392.
198. Coiras M, López-Huertas MR, Rullas J, Mittelbrunn M, Alcamí J. 2007. Basal shuttle of NF-κB/IκBα in resting T lymphocytes regulates HIV-1 LTR dependent expression. Retrovirology 4:1–13.
199. Karn J, Mbonye U. 2011. Control of HIV Latency by Epigenetic and Non-Epigenetic Mechanisms. Curr HIV Res 9:554–567.
200. Lacroix I, Lipcey C, Imbert J, Kahn-Perlès B. 2002. Sp1 transcriptional activity is up-regulated by phosphatase 2A in dividing T lymphocytes. J Biol Chem 277:9598–9605.
201. Barboric M, Matija Peterlin B. 2005. A new paradigm in eukaryotic biology: HIV Tat and the control of transcriptional elongation. PLoS Biol 3:0200–0203.
202. Adelman K, Lis JT. 2012. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat Rev Genet 13:720–731.
203. Ping YH, Rana TM. 2001. DSIF and NELF Interact with RNA Polymerase II Elongation Complex and HIV-1 Tat Stimulates P-TEFb-mediated Phosphorylation of RNA Polymerase II and DSIF during Transcription Elongation. J Biol Chem 276:12951–12958.
204. Adams M, Sharmeen L, Kimpton J, Romeo JM, Garcia JV, Peterlin BM, 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. Proc Natl Acad Sci U S A 91:3862–3866.
205. Toohey MG, Jones KA. 1989. In vitro formation of short RNA polymerase II transcripts that terminate within the HIV-1 and HIV-2 promoter-proximal downstream regions. Genes Dev 3:265–282.
206. Parada CA, Roeder RG. 1996. Enhanced processivity of RNA polymerase II triggered by Tat-induced phosphorylation of its carboxy-terminal domain. Nature 384:375–378.
207. Fujinaga K, Irwin D, Huang Y, Taube R, Kurosu T, Peterlin BM. 2004. Dynamics of Human Immunodeficiency Virus Transcription: P-TEFb Phosphorylates RD and Dissociates Negative Effectors from the Transactivation Response Element. Mol Cell Biol 24:787–795.

208. Benkirane M, Chun RF, Xiao H, Ogryzko V V., Howard BH, Nakatani Y, Jeang K. 1998. Activation of Integrated Provirus Requires Histone Acetyltransferase p300 and P/CAF are coactivators for HIV-1 Tat. J Biol Chem 273:24898–24905.
209. Brady J, Kashanchi F. 2005. Tat gets the 'green' light on transcription initiation. Retrovirology 2:1–8.
210. Karn J. 2011. The molecular biology of HIV latency: Breaking and restoring the Tat-dependent transcriptional circuit. Curr Opin HIV AIDS 6:4–11.
211. Chiang K, Sung T-L, Rice AP. 2012. Regulation of Cyclin T1 and HIV-1 Replication by MicroRNAs in Resting CD4+ T Lymphocytes. J Virol 86:3244–3252.
212. Budhiraja S, Famiglietti M, Bosque A, Planelles V, Rice AP. 2013. Cyclin T1 and CDK9 T-Loop Phosphorylation Are Downregulated during Establishment of HIV-1 Latency in Primary Resting Memory CD4+ T Cells. J Virol 87:1211–1220.
213. Nguyen VT, Kiss T, Michels AA, Bensaude O. 2001. 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. Nature 414:322–325.
214. Barboric M, Yik JHN, Czudnochowski N, Yang Z, Chen R, Contreras X, Geyer M, Peterlin BM, Zhou Q. 2007. Tat competes with HEXIM1 to increase the active pool of P-TEFb for HIV-1 transcription. Nucleic Acids Res 35:2003–2012.
215. Bisgrove DA, Mahmoudi T, Henklein P, Verdin E. 2007. Conserved P-TEFb-interacting domain of BRD4 inhibits HIV transcription. Proc Natl Acad Sci 104:13690–13695.
216. Darcis G, Van Driessche B, Bouchat S, Kirchhoff F, Van Lint C, Driessche B Van, Bouchat S, Kirchhoff F. 2017. Molecular Control of HIV and SIV Latency. Curr Top Microbiol Immunol 6:23–27.
217. Sluis RM Van Der, Pollakis G, Gerven ML Van, Berkhout B, Jeeninga RE, van der Sluis RM, Pollakis G, van Gerven ML, Berkhout B, Jeeninga RE. 2011. Latency profiles of full length HIV-1 molecular clone variants with a subtype specific promoter. Retrovirology 8:73.
218. Shen A, Zink MC, Mankowski JL, Chadwick K, Margolick JB, Carruth LM, Li M, Clements JE, Siliciano RF. 2003. Resting CD4+ T Lymphocytes but Not Thymocytes Provide a Latent Viral Reservoir in a Simian Immunodeficiency Virus- Macaca nemestrina Model of Human Immunodeficiency Virus Type 1-Infected Patients on Highly Active Antiretroviral Therapy. J Virol 77:4938–4949.

219. Shen A, Yang H-C, Zhou Y, Chase AJ, Boyer JD, Zhang H, Margolick JB, Zink MC, Clements JE, Siliciano RF. 2007. Novel pathway for induction of latent virus from resting CD4(+) T cells in the simian immunodeficiency virus/macaque model of human immunodeficiency virus type 1 latency. J Virol 81:1660–1670.
220. North TW, Higgins J, Deere JD, Hayes TL, Villalobos A, Adamson L, Shacklett BL, Schinazi RF, Luciw PA. 2010. Viral Sanctuaries during Highly Active Antiretroviral Therapy in a Nonhuman Primate Model for AIDS. J Virol 84:2913–2922.
221. Dinoso JB, Rabi SA, Blankson JN, Gama L, Mankowski JL, Siliciano RF, Zink MC, Clements JE. 2009. A Simian Immunodeficiency Virus-Infected Macaque Model To Study Viral Reservoirs That Persist during Highly Active Antiretroviral Therapy. J Virol 83:9247–9257.
222. Pace MJ, Agosto L, Graf EH, O'Doherty U. 2011. HIV Reservoirs and Latency Models. Virology 411:344–354.
223. Evans DT, Silvestri G. 2013. Non-Human Primate Models in AIDS Research. Curr Opin HIV AIDS 8:255–261.
224. McCune JM. 1996. Development and applications of the SCID-hu mouse model. Semin Immunol 8:187–196.
225. Honeycutt JB, Wahl A, Archin N, Choudhary S, Margolis D, Garcia JV. 2013. HIV-1 infection, response to treatment and establishment of viral latency in a novel humanized T cell-only mouse (TOM) model. Retrovirology 10:1–12.
226. Melkus MW, Estes JD, Padgett-thomas A, Gatlin J, Denton PW, Othieno FA, Wege AK, Haase AT, Garcia JV. 2006. Humanized mice mount specific adaptive and innate immune responses to EBV and TSST-1. Nat Med 12:1316–1322.
227. Horwitz JA, Halper-Stromberg A, Mouquet H, Gitlin AD, Tretiakova A, Ploss A, Bjorkman PJ, Klein F, Nussenzweig MC. 2013. HIV-1 suppression and durable control by combining single broadly neutralizing antibodies and antiretroviral drugs in humanized mice. Proc Natl Acad Sci 110:16538–16543.
228. Denton PW, Othieno F, Martinez-Torres F, Zou W, Krisko JF, Fleming E, Zein S, Powell DA, Wahl A, Kwak YT, Welch BD, Kay MS, Payne DA, Gallay P, Appella E, Estes JD, Lu M, Garcia JV. 2011. One Percent Tenofovir Applied Topically to Humanized BLT Mice and Used According to the CAPRISA 004 Experimental Design Demonstrates Partial Protection from Vaginal HIV

Infection , Validating the BLT Model for Evaluation of New Microbicide Candidates. J Virol 85:7582–7593.

229. Denton PW, Long JM, Wietgrefe SW, Sykes C, Spagnuolo RA, Snyder OD, Perkey K, Archin NM, Choudhary SK, Yang K, Hudgens MG, Pastan I, Haase AT, Kashuba AD, Berger EA, Margolis DM, Garcia JV. 2014. Targeted Cytotoxic Therapy Kills Persisting HIV Infected Cells During ART. PLoS Pathog 10:1–9.
230. Garcia JV. 2016. In vivo platforms for analysis of HIV persistence and eradication. J Clin Invest 126:424–431.
231. Greenblatt MB, Vbranac V, Tivey T, Tsang K, Tager AM, Aliprantis AO. 2012. Graft versus Host Disease in the Bone Marrow , Liver and Thymus Humanized Mouse Model. PLoS One 7:1–9.
232. Folks TM, Justement J, Kinter A, Dinarello CA, Fauci AS. 1987. Cytokine-Induced Expression of HIV-1 in a Chronically Infected Promonocyte Cell Line. Science 238:800–802.
233. Folks TM, Clouset KA, Justement J, Rabson AB, Duho E, Kehrl JH, Fauci AS. 1989. Tumor necrosis factor a induces expression of human immunodeficiency virus in a chronically infected T-cell clone. Proc Natl Acad Sci U S A 86:2365–2368.
234. Antoni BA, Rabson AB, Kinter A, Bodkin M, Poli G. 1994. NF-κB-Dependent and Independent Pathways of HIV Activation in a Chronically Infected T Cell Line. Virology.
235. Pomerantz RJ, Trono D, Feinberg MB, Baltimore D. 1990. Cells nonproductively infected with HIV-I exhibit an abherrant pattern of viral RNA expression: a molecular model for latency. Cell 61:1271–1276.
236. Krishnan V, Zeichner SL. 2004. Host Cell Gene Expression during Human Immunodeficiency Virus Type 1 Latency and Reactivation and Effects of Targeting Genes That Are Differentially Expressed in Viral Latency. J Virol 78:9458–9473.
237. Symons J, Chopra A, Malantinkova E, De Spiegelaere W, Leary S, Cooper D, Abana CO, Rhodes A, Rezaei SD, Vandekerckhove L, Mallal S, Lewin SR, Cameron PU. 2017. HIV integration sites in latently infected cell lines: evidence of ongoing replication. Retrovirology 14:2.
238. Emiliani S, Van Lint C, Fischle W, Paras P, Orr M, Bradyt J, Verdin E. 1996. A point mutation in the HIV-1 Tat responsive element is associated with

postintegration latency. Proc Natl Acad Sci U S A 93:6377–81.
239. Emiliani S, Fischle W, Ott M, Van Lint C, Amella CA, Verdin E. 1998. Mutations in the tat gene are responsible for human immunodeficiency virus type 1 postintegration latency in the U1 cell line. J Virol 72:1666–70.
240. Jordan A, Bisgrove D, Verdin E. 2003. HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. EMBO J 22:1868–1877.
241. Mahmoudi T, Parra M, Vries RGJ, Kauder SE, Verrijzer CP, Ott M, Verdin E. 2006. The SWI/SNF chromatin-remodeling complex is a cofactor for Tat transactivation of the HIV promoter. J Biol Chem 281:19960–19968.
242. Limsirichai P, Gaj T, Schaffer D V. 2016. CRISPR-mediated activation of latent HIV-1 expression. Mol Ther 24:499–507.
243. Besnard E, Hakre S, Kampmann M, Lim HW, Hosmane NN, Martin A, Bassik MC, Verschueren E, Battivelli E, Chan J, Svensson JP, Gramatica A, Conrad RJ, Ott M, Greene WC, Krogan NJ, Siliciano RF, Weissman JS, Verdin E. 2016. The mTOR Complex Controls HIV Latency. Cell Host Microbe 20:785–797.
244. Lorenzo-Redondo R, Fryer HR, Bedford T, Kim E-Y, Archer J, Pond SLK, Timothy W, Chung Y-S, Penugonda S, Chipman JG, Fletcher C V., Schacker TW, Malim MH, Rambaut A, Haase AT, McLean AR, Wolinsky SM. 2016. Persistent HIV-1 replication maintains the tissue reservoir during therapy. Nature 530:51–56.
245. Sedaghat AR, Siliciano RF, Wilke CO. 2008. Low-level HIV-1 replication and the dynamics of the resting CD4+ T cell reservoir for HIV-1 in the setting of HAART. BMC Infect Dis 8:1–14.
246. Martinez-Picado J, Deeks SG. 2016. Persistent HIV-1 replication during antiretroviral therapy. Curr Opin HIV AIDS 11:417–423.
247. Sahu GK, Lee K, Ji J, Braciale V, Baron S, Cloyd MW. 2006. A novel in vitro system to generate and study latently HIV-infected long-lived normal CD4+ T-lymphocytes. Virology 355:127–137.
248. Tyagi M, Pearson RJ, Karn J. 2010. Establishment of HIV Latency in Primary CD4+ Cells Is due to Epigenetic Transcriptional Silencing and P-TEFb Restriction. J Virol 84:6425–6437.
249. Marini A, Harper JM, Romerio F. 2008. An In Vitro System to Model the Establishment and Reactivation of HIV-1 Latency. J Immunol 181:7713–7720.
250. Bosque A, Planelles V. 2009. Induction of HIV-1 latency and reactivation in

primary memory CD4+ T cells. Blood 113:58–65.

251. Yang H-C, Xing S, Shan L, O'Connell K, Dinoso J, Shen A, Zhou Y, Shrum CK, Han Y, Liu JO, Zhang H, Margolick JB, Siliciano RF. 2009. Small-molecule screening using a human primary cell model of HIV latency identifies compounds that reverse latency without cellular activation. J Clin Invest 119:3473–3486.
252. Hakre S, Chavez L, Shirakawa K, Verdin E. 2012. HIV latency : experimental systems and molecular models. FEMS Microbiol Rev 36:706–716.
253. Swiggard WJ, Baytop C, Yu JJ, Li C, Schretzenmair R, Theodosopoulos T, Doherty UO. 2005. Human Immunodeficiency Virus Type 1 Can Establish Latent Infection in Resting CD4 + T Cells in the Absence of Activating Stimuli. J Virol 79:14179–14188.
254. Saleh S, Solomon A, Wightman F, Xhilagal M, Cameron PU, Lewin SR. 2007. The CCR7 ligands CCL19 and CCL21 increase permissiveness of resting CD4+ T cells to HIV infection. Blood 110:4161–4164.
255. Burke B, Brown HJ, Marsden MD, Bristol G, Vatakis DN, Zack JA. 2007. Primary cell model for activation-inducible human immunodeficiency virus. J Virol 81:7424–7434.
256. Calvanese V, Chavez L, Laurent T, Ding S, Verdin E. 2013. Dual-Color HIV Reporters Trace a Population of Latently Infected Cells and Enable Their Purification. Virology 446:283–292.
257. Peden K, Emerman M, Montagnier L. 1991. Changes in growth properties on passage in tissue culture of viruses derived from infectious molecular clones of HIV-1LAI, HIV-1MAL, and HIV-1ELI. Virology 185:661–672.
258. Hübner W, Chen P, Del Portillo A, Liu Y, Gordon RE, Chen BK. 2007. Sequence of Human Immunodeficiency Virus Type 1 (HIV-1) Gag Localization and Oligomerization Monitored with Live Confocal Imaging of a Replication-Competent, Fluorescently Tagged HIV-1. J Virol 81:12596–12607.
259. Shaner NC, Campbell RE, Steinbach PA, Giepmans BNG, Palmer AE, Tsien RY. 2004. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nat Biotechnol 22:1567–1572.
260. Boshart M, Weber F, John G, Dorsch-Hasler K, Fleckenstein B, Schaffner W, Jahn G, Dorsch-hiisler K, Fleckenstein B. 1985. A very strong enhancer element

is located upstream of an immediate early gene of human cytomegalovirus. Cell 41:5212–5530.

261. Wilkinson GWGG, Akrigg A. 1992. Constitutive and enhanced expression from the CMV major IE promoter in a defective adenovirus vector. Nucleic Acids Res 20:2233–2239.
262. Rasmussen TA, Lewin SR. 2016. Shocking HIV out of hiding: where are we with clinical trials of latency reversing agents? Curr Opin HIV AIDS 11:394–401.
263. Ylisastigui L, Archin NM, Lehrman G, Bosch RJ, Margolis DM. 2004. Coaxing HIV-1 from resting CD4 T cells: Histone deacetylase inhibition allows latent viral expression. AIDS 18:1101–1108.
264. Contreras X, Schweneker M, Chen C-S, McCune JM, Deeks SG, Martin J, Peterlin BM. 2009. Suberoylanilide Hydroxamic Acid Reactivates HIV from Latently Infected Cells. J Biol Chem 284:6782–6789.
265. Li Z, Zhu W-G. 2014. Targeting Histone Deacetylases for Cancer Therapy: From Molecular Mechanisms to Clinical Implications. Int J Biol Sci 10:757–770.
266. Cillo AR, Sobolewski MD, Bosch RJ, Fyne E, Piatak M, Coffin JM, Mellors JW. 2014. Quantification of HIV-1 latency reversal in resting CD4+ T cells from patients on suppressive antiretroviral therapy. Proc Natl Acad Sci 111:7078–7083.
267. Bernhard W, Barreto K, Saunders A, Dahabieh MS, Johnson P, Sadowski I. 2011. The Suv39H1 methyltransferase inhibitor chaetocin causes induction of integrated HIV-1 without producing a T cell response. FEBS Lett 585:3549–3554.
268. Fernandez G, Zeichner SL. 2010. Cell line-dependent variability in HIV activation employing DNMT inhibitors. Virol J 7:1–10.
269. Meichle A, Schutze S, Hensel G, Brunsing D, Kronke M. 1990. Protein kinase C-independent activation of nuclear factor κB by tumor necrosis factor. J Biol Chem 265:8339–8343.
270. Coudronniere N, Villalba M, Englund N, Altman A. 2000. NF-κB activation induced by T cell receptor/CD28 costimulation is mediated by protein kinase C-θ. Proc Natl Acad Sci 97:3394–9.
271. Lamph WW, Wamsley P, Sassone-Corsi P, Verma IM. 1988. Induction of proto-oncogene JUN/AP-1 by serum and TPA. Nature 334:629–631.
272. Griffin GE, Leung K, Folks TM, Kunkel S, Nabel GJ. 1989. Activation of HIV

gene expression during monocyte differentiation by induction of NF-κB. Nature.
273. Williams SA, Chen LF, Kwon H, Fenard D, Bisgrove D, Verdin E, Greene WC. 2004. Prostratin antagonizes HIV latency by activating NF-κB. J Biol Chem 279:42008–42017.
274. Bocklandt S, Blumberg PM, Hamer DH. 2003. Activation of latent HIV-1 expression by the potent anti-tumor promoter 12-deoxyphorbol 13-phenylacetate. Antiviral Res 59:89–98.
275. Díaz L, Martínez-Bonet M, Sánchez J, Fernández-Pineda A, Jiménez JL, Muñoz E, Moreno S, Álvarez S, Muñoz-Fernández MÁ. 2015. Bryostatin activates HIV-1 latent expression in human astrocytes through a PKC and NF-κB-dependent mechanism. Sci Rep 5:1–12.
276. Lebkowski JS, Mcnally MA, Okarma TB, Lerch LB. 1987. Inducible gene expression from multiple promotes by the tumor-promoting agent, PMA. Nucleic Acids Res 15:9043–9055.
277. Kulkosky J, Culnan DM, Roman J, Dornadula G, Schnell M, Boyd MR, Pomerantz RJ. 2001. Prostratin: activation of latent HIV-1 expression suggests a potential inductive adjuvant therapy for HAART. Blood 98:3006–3015.
278. Hamer DH, Bocklandt S, McHugh L, Chun T-W, Blumberg PM, Sigano DM, Marquez VE. 2003. Rational design of drugs that induce human immunodeficiency virus replication. J Virol 77:10227–10236.
279. Contreras X, Barboric M, Lenasi T, Peterlin BM. 2007. HMBA releases P-TEFb from HEXIM1 and 7SK snRNA via PI3K/Akt and activates HIV transcription. PLoS Pathog 3:1459–1469.
280. Jones RB, Mueller S, O'Connor R, Rimpel K, Sloan D, Karel D, Wong HC, Jeng EK, Thomas AS, Whitney JB, Lim S-Y, Kovacs C, Benko E, Karandish S, Huang S-H, Buzon MJ, Lichterfeld M, Irrinki A, Murry JP, Tsai A, Yu H, Geleziunas R, Trocha A, Ostrowski MA, Irvine DJ, Walker BD. 2016. A Subset of Latency-Reversing Agents Expose HIV-Infected Resting CD4 + T-Cells to Recognition by Cytotoxic T-Lymphocytes. PLoS Pathog 12:1–25.
281. Banerjee C, Archin N, Michaels D, Belkina AC, Denis G V., Bradner J, Sebastiani P, Margolis DM, Montano M. 2012. BET bromodomain inhibition as a novel strategy for reactivation of HIV-1. J Leukoc Biol 92:1147–1154.
282. Laird GM, Bullen CK, Rosenbloom DIS, Martin AR, Hill AL, Durand CM, Siliciano JD, Siliciano RF. 2015. Ex vivo analysis identifies effective HIV-1 latency –

reversing drug combinations. J Clin Invest 125:1901–1912.
283. Darcis G, Bouchat S, Kula A, Van Driessche B, Delacourt N, Vanhulle C, Avettand-Fenoel V, De Wit S, Rohr O, Rouzioux C, Van Lint C. 2017. Reactivation capacity by latency-reversing agents ex vivo correlates with the size of the HIV-1 reservoir. AIDS 31:181–189.
284. Münch J, Rajan D, Rücker E, Wildum S, Adam N, Kirchhoff F. 2005. The role of upstream U3 sequences in HIV-1 replication and CD4+ T cell depletion in human lymphoid tissue ex vivo. Virology 341:313–320.
285. Selhorst P, Combrinck C, Ndabambi N, Ismail SD, Abrahams M-R, Lacerda M, Samsunder N, Garrett N, Abdool Karim Q, Abdool Karim SS, Williamson C. 2017. Replication Capacity of Viruses from Acute Infection Drives HIV-1 Disease Progression. J Virol 91:1–15.
286. Chopera DR, Mann JK, Mwimanzi P, Omarjee S, Kuang XT, Ndabambi N, Goodier S, Martin E, Naranbhai V, Abdool Karim SS, Abdool Karim Q, Brumme ZL, Ndung'u T, Williamson C, Brockman MA. 2013. No Evidence for Selection of HIV-1 with Enhanced Gag-Protease or Nef Function among Breakthrough Infections in the CAPRISA 004 Tenofovir Microbicide Trial. PLoS One 8:2–6.
287. Abdool Karim Q, Abdool Karim SS, Frohlich JA, Grobler AC, Baxter C, Mansoor LE, Kharsany AB, Sibeko S, Mlisana KP, Omar Z, Gengiah TN, Maarschalk S, Arulappan N, Mlotshwa M, Morris L, Taylor D. 2010. Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. Science 329:1168–1174.
288. Valley-Omar Z, Sibeko S, Anderson J, Goodier S, Werner L, Arney L, Naranbhai V, Treurnicht F, Abrahams M-R, Bandawe G, Swanstrom R, Karim QA, Karim SSA, Williamson C. 2012. CAPRISA 004 tenofovir microbicide trial: No impact of tenofovir gel on the HIV transmission bottleneck. J Infect Dis 206:35–40.
289. Hübner W, Mcnerney GP, Chen P, Dale BM, Gordon RE, Chuang FYS, Li X, Asmuth DM, Huser T, Chen BK. 2009. Quantitative 3D Video Microscopy of HIV Transfer Across T Cell Virological Synapses. Science 323:1743–1747.
290. Chang LJ, Urlacher V, Iwakuma T, Cui Y, Zucali J. 1999. Efficacy and safety analyses of a recombinant human immunodeficiency virus type 1 derived vector system. Gene Ther 6:715–728.
291. Iwakuma T, Cui Y, Chang LJ. 1999. Self-Inactivating lentiviral vectors with U3 and U5 modifications. Virology 261:120–132.

292. Cui Y, Iwakuma T, Chang L-J. 1999. Contributions of Viral Splice Sites and cis-Regulatory Elements to Lentivirus Vector Function. J Virol 73:6171–6176.
293. Zolotukhin S, Potter M, Hauswirth WW, Guy J, Muzyczka N. 1996. A 'Humanized' Green Fluorescent Protein cDNA Adapted for High-Level Expression in Mammalian Cells. J Virol 70:4646–4654.
294. Wang Y, Prosen DE, Mei L, Sullivan JC, Finney M, Vander Horn PB. 2004. A novel strategy to engineer DNA polymerases for enhanced processivity and improved performance in vitro. Nucleic Acids Res 32:1197–1207.
295. Van Den Ent F, Löwe J. 2006. RF cloning: A restriction-free method for inserting target genes into plasmids. J Biochem Biophys Methods 67:67–74.
296. Bond SR, Naus CC. 2012. RF-Cloning.org: An online tool for the design of restriction-free cloning projects. Nucleic Acids Res 40:1–5.
297. Platt EJ, Bilska M, Kozak SL, Kabat D, Montefiori DC. 2009. Evidence that Ecotropic Murine Leukemia Virus Contamination in TZM-bl Cells Does Not Affect the Outcome of Neutralizing Antibody Assays with Human Immunodeficiency Virus Type 1. J Virol 83:8289–8292.
298. Takeuchi Y, McClure MO, Pizzato M. 2008. Identification of Gammaretroviruses Constitutively Released from Cell Lines Used for Human Immunodeficiency Virus Research. J Virol 82:12585–12588.
299. Wei X, Decker JM, Liu H, Zhang Z, Arani RB, Kilby JM, Saag MS, Wu X, Shaw GM, Kappes JC. 2002. Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy. Antimicrob Agents Chemother 46:1896–1905.
300. Derdeyn CA, Decker JM, Sfakianos JN, Wu X, O'Brien WA, Ratner L, Kappes JC, Shaw GM, Hunter E. 2000. Sensitivity of human immunodeficiency virus type 1 to the fusion inhibitor T-20 is modulated by coreceptor specificity defined by the V3 loop of gp120. J Virol 74:8358–67.
301. Platt EJ, Wehrly K, Kuhmann SE, Chesebro B, Kabat D. 1998. Effects of CCR5 and CD4 Cell Surface Concentrations on Infections by Macrophagetropic Isolates of Human Immunodeficiency Virus Type 1. J Virol 72:2855–2864.
302. Scherer WF, Syverton JT, Gey GO. 1953. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix. J Exp Med 97:695–710.
303. Weiss A, Wiskocil RL, Stobo JD. 1984. The role of T3 surface molecules in the

activation of human T cells: a two-stimulus requirement for IL 2 production reflects events occurring at a pre-translational level. J Immunol 133:123–128.

304. Howell DN, Andreotti PE, Dawson JR, Cresswell P. 1985. Natural killing target antigens as inducers of interferon: studies with an immunoselected, natural killing-resistant human T lymphoblastoid cell line. J Immunol 134:971–6.
305. Lyerly HK, Reed DL, Matthews TJ, Langlois AJ, Ahearne PA, Petteway SRJ, Weinhold KJ. 1987. Anti-GP 120 antibodies from HIV seropositive individuals mediate broadly reactive anti-HIV ADCC. AIDS Res Hum Retroviruses 3:409–422.
306. Trkola A, Matthews J, Gordon C, Ketas T, Moore JP. 1999. A cell line-based neutralization assay for primary human immunodeficiency virus type 1 isolates that use either the CCR5 or the CXCR4 coreceptor. J Virol 73:8966–74.
307. Chesnoy S, Huang L. 2000. Structure and Function og Lipid-DNA Complexes for Gene Delivery. Annu Rev Biochem 69:1005–1075.
308. Hirko A, Tang F, Hughes J. 2003. Cationic Lipid Vectors for Plasmid DNA Delivery. Curr Med Chem 10:1185–1193.
309. Liu D, Ren T, Gao X. 2003. Cationic transfection lipids. Curr Med Chem 10:1307–1315.
310. Dalby B, Cates S, Harris A, Ohki EC, Tilkins ML, Price PJ, Ciccarone VC. 2004. Advanced transfection with Lipofectamine 2000 reagent: Primary neurons, siRNA, and high-throughput applications. Methods 33:95–103.
311. Reed LJ, Muench H. 1938. A Simple Method of Estimating Fifty Percent Endpoints. Am J Hyg 27:493–497.
312. Davis HE, Morgan JR, Yarmush ML. 2002. Polybrene increases retrovirus gene transfer efficiency by enhancing receptor-independent virus adsorption on target cell membranes. Biophys Chem 97:159–172.
313. van Loggerenberg F, Mlisana K, Williamson C, Auld SC, Morris L, Gray CM, Abdool Karim Q, Grobler A, Barnabas N, Iriogbe I, Abdool Karim SS, Acute C, Study I. 2008. Establishing a Cohort at High Risk of HIV Infection in South Africa : Challenges and Experiences of the CAPRISA 002 Acute Infection Study. PLoS One 3:1–8.
314. Li Y, Svehla K, Mathy NL, Voss G, Mascola JR, Wyatt R. 2006. Characterization of Antibody Responses Elicited by Human Immunodeficiency Virus Type 1 Primary Isolate Trimeric and Monomeric Envelope Glycoproteins in Selected

Adjuvants † 80:1414–1426.
315. Tamura K, Nei M. 1993. Estimation of the Number of Nucleotide Substitutions in the Control Region of Mitochondrial DNA in Humans and Chimpanzees. Mol Biol Evol 10:512–526.
316. Markowitz M, Louie M, Hurley A, Sun E, Di Mascio M, Perelson AS, Ho DD. 2003. A novel antiviral intervention results in more accurate assessment of human immunodeficiency virus type 1 replication dynamics and T-cell decay in vivo. J Virol 77:5037–8.
317. Hahn BH, Shaw GM, Taylor ME, Redfield RR, Markham PD, Salahuddin SZ, Wong-Staal F, Gallo RC, Parks ES, Parks WP. 1986. Genetic Variation in HTLV-III/LAV Over Time in Patients with AIDS or at Risk for AIDS. Science 232:1548–1553.
318. Tamura K, Nei M, Kumar S. 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci U S A 101:11030–11035.
319. Procopio FA, Fromentin R, Kulpa DA, Brehm JH, Bebin AG, Strain MC, Richman DD, O'Doherty U, Palmer S, Hecht FM, Hoh R, Barnard RJO, Miller MD, Hazuda DJ, Deeks SG, Sékaly RP, Chomont N. 2015. A Novel Assay to Measure the Magnitude of the Inducible Viral Reservoir in HIV-infected Individuals. EBioMedicine 2:874–883.
320. Mbonye U, Karn J. 2017. The Molecular Basis for Human Immunodeficiency Virus Latency. Annu Rev Virol 4:264–285.
321. Yamamoto Y, Pahwa R, Pahwa S. 2006. S-nitrosoglutathione modulates CXCR4 and ICOS expression. Cell Mol Biol Lett 11:30–36.
322. Wills MR, Poole E, Lau B, Krishna B, Sinclair JH. 2015. The immunology of human cytomegalovirus latency: Could latent infection be cleared by novel immunotherapeutic strategies? Cell Mol Immunol 12:128–138.
323. Sinclair J. 2008. Human cytomegalovirus: Latency and reactivation in the myeloid lineage. J Clin Virol 41:180–185.
324. Sambucetti LC, Cherrington JM, Wilkinson GW, Mocarski ES. 1989. NF-kappa B activation of the cytomegalovirus enhancer is mediated by a viral transactivator and by T cell stimulation. EMBO J 8:4251–8.
325. Hunninghake GW, Monick MM, Liu B. 1989. The Promoter-Regulatory Region of the Major Immediate-Early Gene of Human Cytomegalovirus Responds to T-

Lymphocyte Stimulation and Contains Functional Cyclic AMP-Response Elements 63:3026–3033.

326. Norrman K, Fischer Y, Bonnamy B, Sand FW, Ravassard P, Semb H. 2010. Quantitative comparison of constitutive promoters in human ES cells. PLoS One 5.
327. Deng K, Pertea M, Rongvaux A, Wang L, Durand CM, Ghiaur G, Lai J, McHugh HL, Hao H, Zhang H, Margolick JB, Gurer C, Murphy AJ, Valenzuela DM, Yancopoulos GD, Deeks SG, Strowig T, Kumar P, Siliciano JD, Salzberg SL, Flavell RA, Shan L, Siliciano RF. 2015. Broad CTL response is required to clear latent HIV-1 due to dominance of escape mutations. Nature 517:381–385.
328. Sambrook J, Russell DW. 2001. Molecular Cloning: A Laboratory ManualThird Edit. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York.