The role of clonally expanded HIV-1 infected cells in maintaining HIV reservoirs in adults and children on antiretroviral treatment

by

Johannes Christiaan Botha

Dissertation presented for the degree of

Doctor of Philosophy in Medical Virology

in the Faculty of Medicine and Health Sciences

at Stellenbosch University

Supervisor: Prof. Gert U van Zyl

Co-supervisor: Prof. Susan Engelbrecht

December 2020

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

December 2020

Abstract

The HIV-1 pandemic remains a major public health dilemma. Treatment with combination antiretroviral therapy (cART) can, in the majority of patients, sufficiently suppress HIV viral replication. However, cART is not a cure as it does not affect long-lived reservoirs. HIV proviruses are harboured in the genome of long-lived CD4 helper cells that are maintained by having long half-lives and by being replenished through cellular proliferation. HIV infected cells that proliferate can be identified as clonal populations by all having the same integration site. A small proportion of proviruses are intact from where HIV rebounds after therapy interruption and is the barrier to achieving cure or remission. Therefore, many molecular assays have been designed to characterise HIV proviruses, with a focus on intact proviruses. From this the proviral landscape is estimated to consist of ~2% intact proviruses and up to 98% defective proviruses comprising deletions and hypermutations. However, highly deleted proviruses such as solo-long terminal repeats (LTRs) remain largely uncharacterised. Recently, it has become apparent that some proviral clones are able to express infectious HIV and cause clonal viraemia in patients on cART, which are not indicative of ongoing cycles of viral replication. This has further confirmed the importance of proviral clones in maintaining the HIV reservoir. Subsequently, aspects of proviral clones were investigated in this study.

A novel assay capable of characterising severely deleted proviruses by targeting the unique integration sites was developed. The characterisation of proviral clones in paediatric patients with this novel assay revealed that severely deleted solo-LTR proviruses may be more prevalent than previously expected.

Further investigation of these solo-LTR proviral clones was performed with another novel assay capable of quantifying proviruses based on unique integration sites. The longitudinal waxing and waning of solo-LTR proviral clones could be observed. Although solo-LTR proviruses do not contribute to the true reservoir, it suggests that current proviral landscape proportions require adjustment to account for all HIV integration events in cells.

Clonal viraemia was studied in HIV positive adult patients on long term cART presenting with non-suppressible HIV viraemia. Three clusters of monotypic plasma viraemia and cell associated DNA HIV sequences were identified in one patient. These monotypic proviruses were shown to be replication competent by viral outgrowth. This provides evidence that

clonally proliferated cells harbour at least a proportion of intact proviruses and therefore constitute an important component of the true HIV reservoir. Two additional patients were identified with uncommon cases of sustained viraemia. Firstly, a probable large cell clone harbouring non-infectious virus was found to leak proviral nucleic acid into plasma, resulting in apparent treatment failure. Secondly, possible clonal viremia was observed in a patient with very slow decaying viral load, despite objective evidence of adherence and initial undetectability of treatment-relevant drug resistance over a period of 2 years. We conclude that clonal viraemia may be underreported, and that cases like these show that clonal viraemia should be considered as an explanation of non-suppressed viraemia or apparent therapy failure in the management of HIV infected patients on cART.

Opsomming

Die MIV-1 pandemie is steeds ‘n geweldige gesondheids probleem. Behandeling met kombinasie antiretrovirale terapie (kART) kan, in die meerderheid van pasiënte, voldoende onderdrukking van die MIV virale lading behaal. kART is egter nie ‘n geneesmiddel nie omdat dit geen effek op die lang-lewende reservoir het nie. MIV provirusse word gehuisves in die genoom van lang-lewende CD4 helper selle wat voortleef weens ‘n lang halfleeftyd en aangevul word deur sellulêre vermeerdering. MIV geïnfekteerde selle wat vermenigvuldig kan geïdentifiseer word as klonale populasies met identiese integrasie liggings. ‘n Klein proporsie provirusse is intakt en is die bron van MIV wat terugbons na terapie gestaak word en dit verhinder genesing of remissie. Daarom is etlike molekulêre toetse ontwerp om MIV provirusse te karakteriseer, met ‘n fokus op intakte provirusse. Bykans 2% van die provirale landskap is intakte provirusse terwyl soveel as 98% foutief is. Hoogs-gebrekkige provirusse, soos die wat alleen bestaan uit lang terminale herhalings (LTRs), is egter hoofsaaklik on-gekarakteriseer. Dit het onlangs geblyk dat sekere provirale klone infektiewe MIV kan produseer en klonale viremie veroorsaak in pasiënte op kART, sonder om aanduidend te wees van volgehoue replikasie. Hierdie bevindings bevestig die belangrikheid van provirale klone in die onderhouding van MIV reservoirs. In hierdie studie is aspekte van provirale klone dus verder bestudeer.

‘n Nuwe toets is ontwikkel wat hoogs-gebrekkige provirusse kan karakteriseer deur die unieke integrasie ligging te teiken. Karakterisering van provirale klone in pediatriese pasiënte met hierdie nuwe toets het onthul dat hoogs-gebrekkige LTR-alleen provirusse meer algemeen mag wees as wat voorheen verwag is.

Verdere ondersoek van die LTR-alleen provirusse is gedoen deur nog ‘n nuwe toets wat provirusse kan kwantifiseer deur gebruik te maak van die unieke integrasie ligging. Longitudinale toename en kwyning van LTR-alleen provirale klone kon waargeneem word. Hoewel LTR-alleen provirusse nie self tot die reservoir bydra nie, blyk dit hieruit dat daar aanpassing benodig word in die huidig-aanvaarde provirale landskap proporsies om al die integrasie voorvalle in te sluit.

Klonale viremie is bestudeer in MIV positiewe volwasse pasiënte op kART van wie viremie nie-onderdukbaar was nie. Drie groeperings monotipiese plasma viremie en sellulêr

geassosieerde DNA MIV nukleotied volgorde is in een pasiënt geïdentifiseer. Hierdie monotipiese provirusse het replikasie-vermoë soos bewys is met virale uitgroei toetse. Dit verleen ondersteuning daaraan dat klonaal-vermeerderde selle ‘n gedeelte van die intakte provirusse mag huisves en dus ‘n aansienlike komponent van die ware MIV reservoir uitmaak. Twee addisionele pasiënte met ongewone volhoubare viremie is geïdentifiseer. Eerstens, is ‘n waarskynlike groot sel-kloon wat ’n gebrekkige provirus huisves en provirale nukleïensuur in plasma in lek, geïdentifiseer, wat tot klaarblyklike terapie-faling lei. Tweedens, is moontlike klonale viremie waargeneem in ‘n pasiënt met baie stadige afname van die virale lading oor twee jaar van studie, ten spyte van terapie en geen terapie-verwante weerstandigheid nie. Ons samevatting is dat klonale viremie meer algemeen mag wees as verwag, en dat gevalle soos hier bespreek, daarop dui dat klonale viremie ’n oorweging moet wees in die hantering van MIV geïnfekteerde pasiënte op kART met volhoubare viremie of oënskynlike terapie-faling.

Acknowledgements

IN EXCELSIS DEO

I would like to thank and acknowledge the following people:

My supervisor Prof Gert U van Zyl, thank you for this opportunity. Your continued advice, guidance and support is truly appreciated. It has been an honour to work with you.

My co-supervisor Prof Susan Engelbrecht, thank you for your advice and support throughout this project.

Prof John W Mellors and his team, thank you for the opportunity to work with you, the invaluable training and advancing my abilities as a researcher.

Dr Mary Kearney, thank you for your expert advice and guidance.

Philip Martin Webb, you are the most important person in my world. Thank you for your love and encouragement, you make everything better.

The entire Reservoir team and Medical Virology colleagues, thank you for constant encouragement and always being there for a laugh.

Thank you to my family and friends for their support and love throughout this time.

A special thank you to the study participants, without your contributions none of this would have been possible.

Thank you to the Poliomyelitis Research Foundation, the National Research Foundation and the National Institute of Health for financial support and making this research possible.

Research Outputs

Conference participation:

Abstracts and posters included in Addendum A.

Botha, J.C., Steegen, K., Hans L., Karstaedt A., Carmona S., Reddy D., van Zyl G.U., 2019, "HIV Pseudo viremia from large defective cell population," Annual Academic Year Day, Stellenbosch University, 21 August 2019. (Poster)

Halvas, E.K., Joseph, K., Brandt, L.D., **Botha, J.C**., Sobolewski, M., Jacobs, J.L., Keele, B.F., Kearney, M.F., Coffin, J.M., Rausch, J.W., Guo, S., Wu, X., Hughes, S.H. & Mellors, J.W., 2019, "Nonsuppressible viremia on ART from large cell clones carrying intact proviruses," *CROI 2019,* Abstract 23. (Presentation)

Botha, J.C., Steegan, K., Hans, L., Karstaedt, A., Carmona, S., Reddy, D. & van Zyl, G.U., 2019, "HIV Pseudo viremia from large defective cell population," *28th International Workshop on HIV Drug Resistance and Treatment Strategies*. (Poster and Presentation)

Botha, J.C., Steegan, K., Hans, L., Karstaedt, A., Carmona, S., Reddy, D., Kearney, M., Mellors, J. & van Zyl, G.U., 2020, "'False ART failure' from identical hypermutated HIV nucleic acid in plasma," *CROI 2020,* Abstract 373. (Poster)

Table of Contents

List of figures

List of tables

List of abbreviations

µg	Microgram
µl	Microlitre
µM	Micromolar
3TC	Lamivudine
ACTG	AIDS clinical trials group
AIDS	Acquired immunodeficiency syndrome
ALLINIs	Allosteric integrase inhibitors
APOBEC	Apolipoprotein B mRNA editing enzyme
ART	Antiretroviral therapy
Asp	Antisense protein
AZT	Azidothymidine
BC	Buffy coat
BLAT	BLAST-like alignment tool
BSA	Bovine serum albumin
CAD	Cell associated DNA
CAR	Cell associated RNA
cART	Combination antiretroviral therapy
cDNA	Complimentary DNA
CHER	Children with HIV early antiretroviral
Chr	Chromosome
CI	Confidence intervals

CRF	Circulating recombinant forms
Ct	Cycle threshold
DMSO	Dimethyl sulfoxide
dNTPs	Deoxynucleotide triphosphates
DPO	Dual priming oligonucleotide
DTG	Dolutegravir
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EFV	Efavirenz
Exo	Exonuclease
FBS	Foetal bovine serum
FTC	Emtricitabine
gp	Glycoprotein
GuSCN	Guanidinium thiocyanate
HFH	Host to full-length provirus to host amplification
Hg19	Human genome 19
HIV	Human immunodeficiency virus
InSTI	Integrase strand transfer inhibitors
IS	Integration site
ISA	Integration site analysis
ISSAQ	Integration site specific proviral absolute quantification
ISSPA	Integration site specific proviral amplification
IUPM	Infectious units per million
kb	Kilobase

LPV/R	Lopinavir / Ritonavir
LTR	Long terminal repeats
MDA	Multiple displacement amplification
MGB	Minor groove binding
min	Minute
mL	Millilitre
mM	Millimolar
NCI	National Cancer Institute
NEB	New England Biolabs Inc.
NFL	Near full-length
ng	Nanogram
NNRTI	non-nucleoside reverse transcription inhibitors
NTC	No template control
NTRI	nucleoside reverse transcription inhibitors
OARAC	Office of AIDS research advisory council
P6PrRT	p6, Protease and Reverse transcriptase
PBMC	Peripheral blood mononuclear cell
PBS	Primer binding site
PBS	Phosphate buffered saline
PHA	Phytohaemagglutinin
pmol	Picomole
Pol	Polymerase
PPT	Polypurine tracts
QVOA	Quantitative viral outgrowth assay

rSAP	Shrimp Alkaline Phosphatase
rt	*real-time*
sec	Second
SGS	Single genome sequencing
SNP	Single nucleotide polymorphism
SPRI	Solid phase reversible immobilisation
STCM	Super T cell media
TAR	Transactivation response
TDF	Tenofovir
TLD	Tenofovir, lamivudine and dolutegravir
Tm	Melting temperatures
TP	Time points
UK	United Kingdom
USA	United States of America
VL	Viral load

Chapter 1

Literature review

1.1 General introduction

Human immunodeficiency virus (HIV) remains a major threat to public health with 38 million individuals currently diagnosed, 1.7 million in just the last year. Due to the scale of the problem, a staggering 12.6 million of these people are still awaiting treatment (UNAIDS 2020). The discovery and subsequent advancements of antiretroviral drugs revolutionised HIV treatment, suppressing the virus and allowing patients a relatively normal life (Chun et al. 1997; Broder 2010). When HIV replication is not fully suppressed HIV drug resistance development is common, especially when using low genetic barrier regimens. Low genetic barrier drugs lose clinical effectiveness, in the presence of single drug resistance mutations, which often occur without a significant impact on viral fitness (Tang and Shafer, 2012; AIDSinfo 2020). In addition, it is very difficult to objectively assess drug adherence and self-reported adherence is inaccurate due to social desirability bias (Nachega et al. 2011). Therefore, when drug resistance is not detected, treatment failure is usually ascribed to inadequate adherence. Alternative pathways to drug resistance that occur in genes such as *gag* and *env* and which are not included in resistance scoring algorithms have also been proposed, but it remains uncertain whether these mechnanisms are clinically relevant (Dam et al. 2009; Özen et al. 2014; Malet et al. 2017; van Duyne et al. 2019). Thresholds for treatment failure are often arbitrarily defined and there is no threshold with near-perfect receiver operator characteristics, and thresholds are therefore either not very sensitive or not very specific (Joya et al. 2019). Low-level and transient viraemia therefore complicates patient care (AIDSinfo 2020). With the use of modern tolerable and high genetic barrier regimens, the large majority of patients can achieve virological suppression. Nevertheless, these therapies do not constitute a cure as viral rebound occurs soon after treatment interruption (Chun et al. 1997). The main barrier to HIV cure is long-lived reservoirs. HIV is maintained as proviruses in the genome of these long-lived cells throughout the body, surviving for decades (Finzi et al. 1997; Blankson et al. 2000; Besson et al. 2014). The identification of monotypic residual viraemia (plasma viraemia consisting of multiple identical

sequences) led to the discovery that proviral clones, rather than ongoing replication, are responsible for HIV survival (Dinoso et al. 2009; Anderson et al. 2011; Wagner et al. 2013; Kearney et al. 2014; Reeves et al. 2018). Due to the size of the human genome and relatively random HIV integration, cells that harbour the same proviral integration site represent proviral clones (Simonetti et al. 2016; Cohn & Nussenzweig 2017). Cells harbouring latent HIV-1 proviruses are not only long-lived but can self-renew through clonal proliferation and constitute the major barrier to cure (Maldarelli et al. 2014; Wagner et al. 2014). Moreover, the size of these clones and their ability to wax and wane over time complicate proviral studies (Wang et al. 2018; Simonetti et al. 2020). Significant advances have been made in HIV reservoir research, with a multitude of assays designed for reservoir characterization, including viral outgrowth and proviral induction, viral and proviral quantification and nucleotide sequencing assays (Finzi et al. 1997; Palmer et al. 2005; Li et al. 2007; Cillo et al. 2013; Hatano et al. 2013; Bruner et al. 2019). Additionally, the study of proviral integration sites offers substantial insights into clonal proliferation and reservoir survival (Chomont et al. 2009; Maldarelli et al. 2014; Patro et al. 2019). The fact that HIV survives in long-lived cells with the potential for self-renewal, rather than low-level ongoing replication (Wagner et al. 2013; Maldarelli et al. 2014), directly affects which curative approaches are most feasible. Several curative strategies have been suggested. “Shock/kick and kill” is based on reversing proviral latency followed by the killing of HIV infected cells (Deeks 2012). In contrast the “block and lock” approach aims to enforce latency to prevent proviral reactivation (Suzuki et al. 2008; Ahlenstiel et al. 2015). Gene editing to induce host resistance or deactivate HIV proviruses have also been suggested (Ebina et al. 2013; Wang et al. 2014). However, since it may not be feasible to target all HIV infected cells, more emphasis has been placed on functional cure approaches. Functional cure (or remission) can be defined as an approach that, in the absence of antiretroviral therapy (ART), achieves control of viral replication, thereby preventing disease progression and rendering the individual non-infectious without ridding the body of all reservoirs. Therapeutic vaccines to achieve HIV remission are currently being investigated (Davenport et al. 2019). However, substantial advances are still required to achieve scalable HIV cures (Thomas et al. 2020).

1.2 Virus characteristics

1.2.1 Classification

HIV belongs to the genus Lentivirus from the family *Retroviridae* and is divided into two types, HIV-1 and HIV-2. HIV-1 is further divided into four groups namely HIV-1 M (major), HIV-1 N (new), HIV-1 O (outlier), HIV-1 P (Simon et al. 1998; Robertson et al. 2000; Plantier et al. 2009; Los Alamos 2017). Group M, the predominant group, is further divided into subtypes A through K with over 51 circulating recombinant forms distributed across the globe (Hemelaar 2012).

1.2.2 Morphology

Mature HIV virions are round and consist of a conical capsid with a lipid membrane envelope, measuring 100 - 120 nm in diameter (Gelderblom 1991; Niedrig et al. 1994). The virus capsid contains two strands of positive-sense single-stranded RNA and viral proteins, including reverse transcriptase and integrase (Figure 1.1) (Luciw 1996).

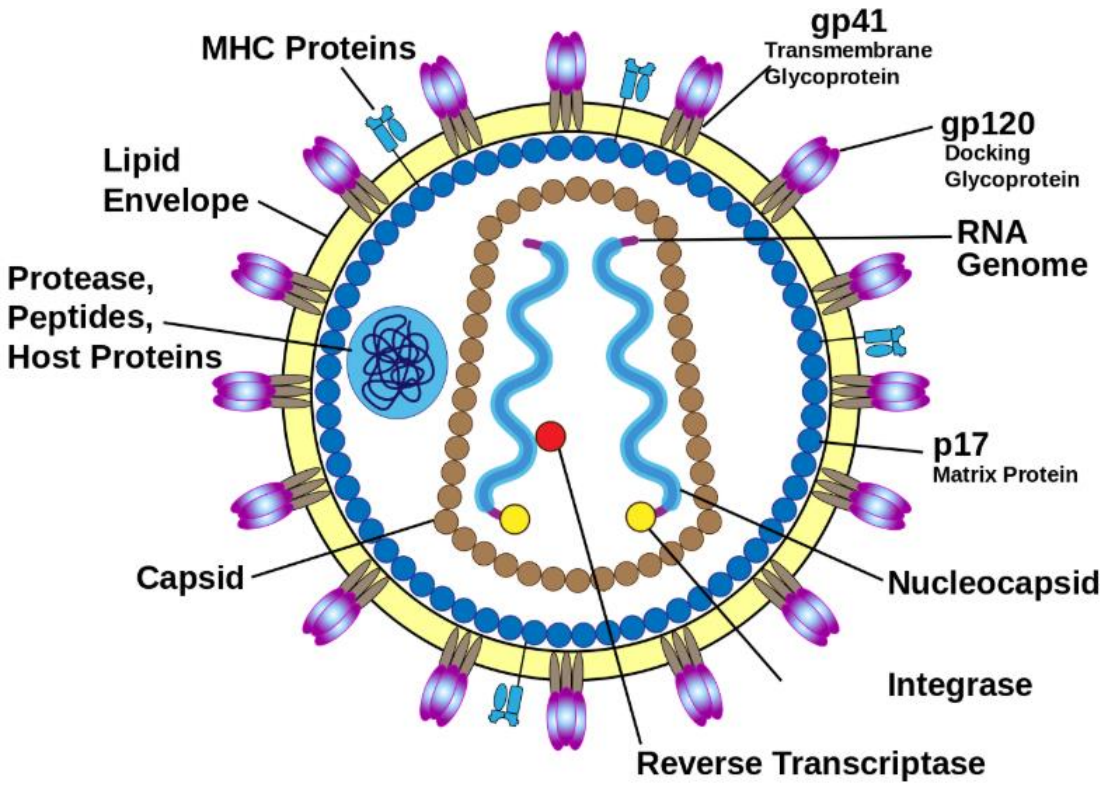


Figure 1.1: Schematic overview of the HIV-1 virion indicating the viral capsid and lipid membrane envelope, with structural and non-structural proteins that make up the mature virus particle. Image credit: NIAID.

1.2.3 Genome

The ~9 kilobase (kb) diploid RNA genome of HIV-1 consists of 10 genes, encoded over all three open reading frames and incomplete long terminal repeats (LTRs) at the 5' (R-U5) and 3' (U3-R) ends (Figure 1.2) (Ooms et al. 2007; Los Alamos 2017).

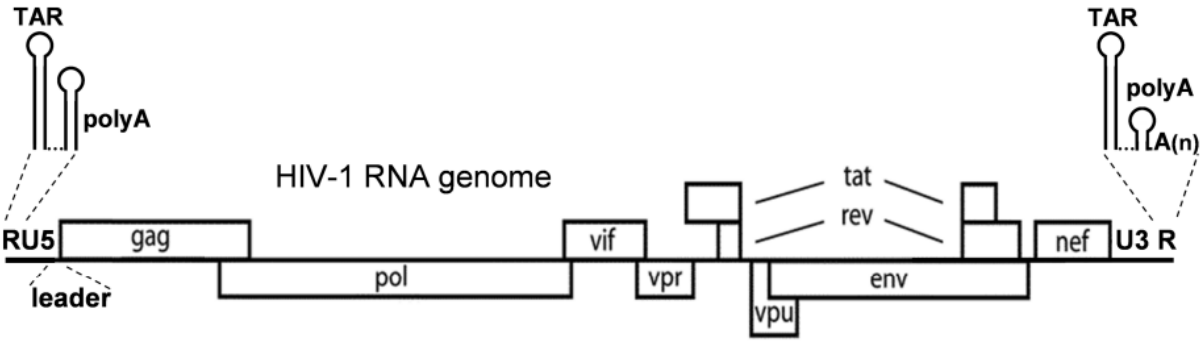


Figure 1.2: Representation of HIV-1 RNA genome indicating the various genes with 5' and 3' incomplete long terminal repeats (Ooms et al. 2007).

After infection, the reverse transcribed HIV genome integrates into the host genome with complete LTRs at the 5' and 3' ends (Figure 1.3) (Greene & Peterlin 2006; Los Alamos 2017). HIV LTRs contain multiple transcription factor binding sites and act as regulator for the expression and suppression of viral genes (Jordan, Defechereux & Verdin 2001; Schröder et al. 2002; Weinberger et al. 2005; Burnett et al. 2009). The HIV genome yields 17 proteins from 10 genes, including major gene proteins (*gag*, *polymerase* and envelope), regulatory proteins (Tat and Rev) and accessory proteins (Nef, Vpr, Vpu and Vif). The major genes give rise to several proteins including *gag* (matrix-p17, capsid-p24, nucleocapsid-p7 and p6), *polymerase* (protease, reverse transcriptase, RNase and integrase) and envelope (glycoprotein (gp) 120 and gp41) (Figure 1.3) (Hope & Trono 2000; Greene & Peterlin 2006; Los Alamos 2017). Additionally, an antisense protein (Asp) gene has been identified in frame 2 reverse (HXB2 position 7373 – 7942) of most HIV-1 group M subtypes (Cassan et al. 2016).

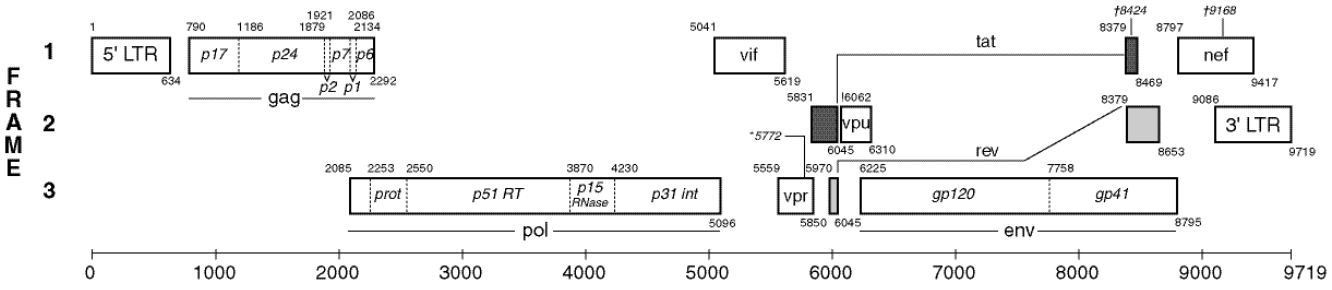


Figure 1.3: Representation of the HIV-1 genome. Open reading frames are indicated for each gene. Genome numbering is based on HXB2 (K03455) (Spach 2019).

1.2.4 Proteins and function

1.2.4.1 Major proteins

Each major gene (*gag*, *polymerase* and envelope) gives rise to a precursor protein, p55 myristoylated protein, gag-Pol polyprotein and gp160 respectively. Each polyprotein is proteolytically cleaved by the encoded viral protease in the process of viral maturation (Göttlinger, Sodroski & Haseltine 1989; Hope & Trono 2000).

The p55 gag polyprotein is cleaved into four functional proteins, matrix-p17, capsid-p24, nucleocapsid-p7 and p6. The matrix protein forms part of the N-terminal of the p55 polyprotein and remains associated with the viral envelope of mature virions. The functional matrix-p17 protein targets *gag* to the lipid membrane and assists in the incorporation of envelope gp120 and gp41 into immature virions. The matrix protein has also been associated with post-assembly steps in the replication cycle of HIV (Yu et al. 1992; Yuan et al. 1993; Dorfman et al. 1994; Zhou et al. 1994; Reil et al. 1998; Göttlinger 2010). The capsid-p24 protein is crucial for virion formation and carrying the genetic material into target cells after infection. The capsid protein is responsible for the assembly of the conical virion core, a distinguishable feature of lentiviruses that contains the viral RNA and proteins (reverse transcriptase and integrase). Reconstructed models of the viral core indicate a relatively open structure, permeable to nucleotide triphosphates, suggesting that reverse transcription of the viral RNA can occur within the core (Gelderblom 1991; S. Li et al. 2000; Göttlinger 2010). The nucleocapsid-p7 protein forms part of the capsid-p24 C-terminal and is essential for virus replication. The nucleocapsid protein is required for the specific packaging of two un-spliced copies of viral RNA into the core structure, in addition to coating the viral RNA in mature virions. Nucleocapsid proteins also assists in annealing reactions in the viral life cycle including, placement of transfer RNA primers, dimeric RNA maturation into compact thermostable forms and RNA strand transfer during reverse transcription (Berkowitz, Fisher & Goff 1996; Rein, Henderson & Levin 1998; Göttlinger 2010). The p6 domain is separated from the nucleocapsid by a peptide, encoded to direct translational frameshifting to the overlapping *polymerase* gene. As a result of this, the p6 domain is excluded from the gag-Pol polyprotein. The p6 domain lacks conventional secondary protein structure and primarily serves as flexible extension to provide binding sites for cellular factors. Although only the N-terminal of p6 appear to be essential for propagation in cell culture, small deletions in the p6 domain is associated with non-progressive HIV-1 infection in humans (Göttlinger et al. 1991; Henderson et al. 1992; Stys, Blaha & Strop 1993; Alexander et al. 2000).

Due to the p6 frameshifting activity, protease, *reverse transcriptase*, RNase and *integrase* domains are expressed as a gag-Pol polyprotein (Jacks et al. 1988). This frame shift to the *polymerase* reading frame is only successful approximately 5% of the time, resulting in a *gag* precursor to gag-Pol polyprotein expression ratio of about 20:1 (Parkin, Chamorro & Varmus 1992; Hope & Trono 2000). The viral protease cleaves the *gag* and Pol peptides from each other during virion maturation, followed by further digestion to separate the protease, reverse transcriptase, RNase and integrase proteins. The Pol peptide cleavage is ineffective with about 50% of the reverse transcriptase and RNase proteins remaining as polyproteins (Hope & Trono 2000). The viral protease is classified as an aspartyl protease and performs as a dimer. The protease is essential for polyprotein cleavage, as discussed. Determining the three-dimensional structure of the HIV-1 protease resulted in drug formulation to inhibit its activity (Miller et al. 1989; Navia et al. 1989; Ashorn et al. 1990; Hope & Trono 2000). The reverse transcriptase protein forms part of the polymerase activities. During reverse transcription, double-stranded DNA is synthesised from the single-stranded diploid HIV-1 RNA. The RNase protein digests the RNA template from the first strand DNA to allow for double-stranded DNA synthesis. Transactivation response (TAR) elements at the 5' ends of viral RNA are required for reverse transcription initiation. Although HIV-1 DNA synthesis occurs relatively quickly after viral entry, viral DNA may remain unintegrated for prolonged periods of time. HIV polymerase is error-prone due to a lack of proof-reading activity (Zack et al. 1990a; Harrich, Ulich & Gaynor 1996; Hope & Trono 2000). Integrase is responsible for the insertion of the double-stranded HIV-1 DNA into the genome of the infected cell. The integrated HIV DNA is referred to as the provirus. Proviral integration will be discussed in detail as part of viral replication. All the *polymerase* gene products are included in the capsid of mature virions.

The gp160 polyprotein is synthesised in the endoplasmic reticulum followed by migration to the Golgi complex for glycosylation, required for infectivity. Polyprotein gp160 is then proteolytically cleaved to yield gp120 and gp41. The gp41 protein consists of three domains, extracellular, transmembrane and cytoplasmic, and is embedded in the viral envelope as trimers (Figure 1.4) (Bernstein et al. 1995; Garg et al. 2011). Protein gp120 is comprised of two domains, the inner and outer domains, and interact with gp41 as trimers. The inner domain forms a non-covalent interaction with gp41 while the larger, less conserved and highly glycosylated outer domain interacts with CD4 and co-receptors (CXCR4 and CCR5) (Landau, Warton & Littman 1988; Kwong et al. 1998; Blumenthal, Durell & Viard 2012).

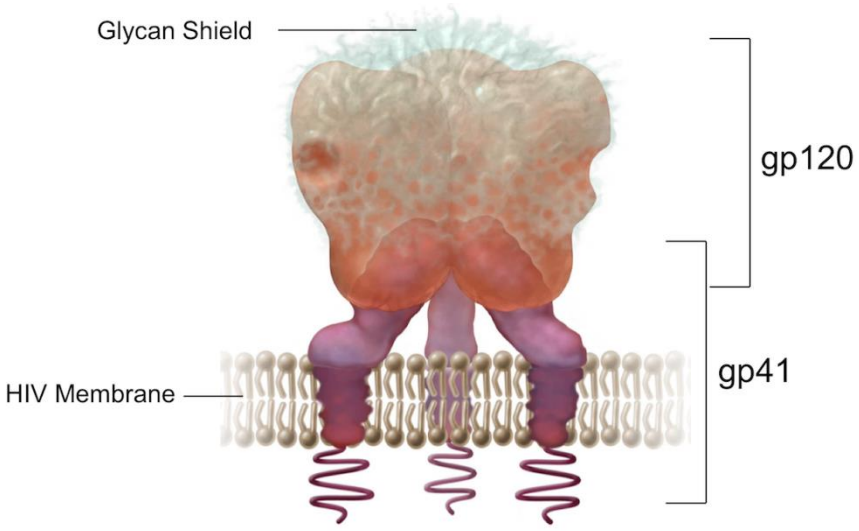


Figure 1.4: Representation of the HIV-1 envelope spike, indicating the gp120/gp41 heterodimer trimers, along with the immuno-protective glycan shield. (Spach 2019)

1.2.4.2 Regulatory proteins

Tat and Rev are essential regulatory proteins for HIV gene expression. Tat is a trans-activator that binds TAR at the 5' end of HIV RNA and activates transcription from LTR by over 1000-fold. Tat is primarily found in the nucleus and nucleolus of infected cells in two forms, a 72 amino acid minor form and an 86 amino acid major form, both of which are functional transcriptional activators (Feng & Holland 1988; Roy et al. 1990). Importantly Tat promotes the elongation phase of HIV-1 transcription for the production of full-length viral transcripts. The absence of Tat results in the production of short transcripts (<100 nucleotides) (Feng & Holland 1988; Roy et al. 1990). Furthermore, Tat and cytokines have been shown to induce the expression of the *Tat* gene (Levy 2011; Seitz 2016).

The shift of early to late phase viral gene expression is induced by Rev. The Rev protein binds a region comprised of a complex RNA secondary structure, known as the Rev response element. This allows for the export of incomplete and un-spliced RNA from the nucleus into the cytoplasm, and results in a rapid cycle between the nucleus and cytoplasm (Kim et al. 1989; Malim et al. 1989; Zapp & Green 1989; Hope & Trono 2000; Los Alamos 2017).

1.2.4.3 Accessory proteins

In addition to the major and regulatory proteins, the HIV-1 genome contains additional genes that encodes accessory proteins (Nef, Vpr, Vpu, Vif and Asp). Although proteins Nef, Vpr, Vpu and Asp are not essential for viral replication, they do represent critical virulence factors.

Vpu, Vpr and Vif proteins are only expressed during late phase replication from incompletely spliced RNA and are therefore Rev dependent.

The Nef protein is expressed during early phase viral replication and exhibits several activities. Nef downregulates the cell surface expression of CD4 by upregulating CD4 targeted endocytosis and lysosomal digestion. The Nef protein has also been shown to disturb T cell activation. Additionally, Nef stimulates virion infectivity, producing virions 10 times more infectious than viral particles generated in the absence of Nef (Kim et al. 1989; Luria, Chambers & Berg 1991; Garcia & Miller 1992; Aiken et al. 1994; Miller et al. 1994; Pandori et al. 1996).

About 100 copies of Vpr protein are incorporated in the HIV virion through interaction with p55 *gag*. Vpr assists in HIV infection of nondividing cells by targeting the nuclear import of the pre-integration complex (Cohen et al. 1990; Zack et al. 1990b; Heinzinger et al. 1994).

Vpu is expressed at rates 10-fold lower than envelope and integrates into internal membranes. Vpu facilitates the degradation of CD4 in the endoplasmic reticulum in addition to enhancing virion release from infected cells (Sato et al. 1990; Schwartz et al. 1990; Schubert et al. 1996).

Althought HIV can replicate in particular tumour cell-lines without a functional Vif protein, it is essential for HIV replication in the natural host, with its absence resulting in defective virions (Strebel et al. 1987). Vif prevents the encapsidation activity of “apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like” (APOBEC) 3G and APOBEC3F, two potent human anti-retroviral cytidine deaminases. APOBEC targets viral mRNA and induces C→U mutations, resulting in DNA G→A mutations. This effect is more commonly known as G→A hypermutations, resulting in the disruption of the HIV genome (Jarmuz et al. 2002; Harris et al. 2003; Mangeat et al. 2003; Bishop et al. 2004; Liddament et al. 2004). The Vif protein directly counteracts APOBEC3G and APOBEC3F by binding these enzymes for targeted degradation. Although this is effective, small amounts of APOBEC enzymes can still be encapsulated in HIV-1 virions (Lecossier et al. 2003; Mariani et al. 2003; Marin et al. 2003).

Finally, the existence of the Asp protein was only recently confirmed, with indications of selective pressure on the *Asp* gene. Although early investigations suggest the Asp protein plays a role in co-receptor selection for HIV entry, this has not been confirmed (Torresilla, Mesnard & Barbeau 2015; Cassan et al. 2016; Dimonte 2017).

1.3 HIV replication cycle

The gp120 trimers of mature HIV virions bind CD4 receptors of targeted cells and in turn allows for the binding of chemokine receptors (CCR5 and CXCR4). This binding results in a conformational change in gp41 trimers which mediates fusion of the viral envelope and cell membrane, leading to the release of the viral core into the cytoplasm (Figure 1.5) (Chan & Kim 1998; Kwong et al. 1998; Doms & Trono 2000). Some individuals exhibit naturally occurring mutant CCR5 receptors, as a result of a 32 base-pair deletion in the *CCR5* gene. Individuals with homozygous mutant *CCR5* genes display significant resistance to HIV infection, emphasizing the importance of CCR5 receptors for HIV infection (Liu et al. 1996; Martinson et al. 1997).

Uncoating of the viral core, mediated by Nef and Vif, initiates during entry into the cytoplasm. The HIV core releases the reverse transcription complex consisting of diploid viral RNA, lysine transfer RNA (reverse transcription primer), the HIV reverse transcriptase, viral integrase, Vpr, matrix protein and nucleocapsid protein (Lu et al. 1998; Öhagen & Gabuzda 2000; Schaeffer, Geleziunas & Greene 2001). Viral reverse transcription from single-stranded RNA to double-stranded DNA is referred to as retro-transcription (Figure 1.6). Lysine transfer RNA attaches to the complimentary primer binding site in the *gag*-leader region. Reverse transcription yields complimentary negative strand DNA of the 5' R and U5 region. The 5' R and U5 RNA region is hydrolysed, releasing the single-stranded DNA section, followed by strand transfer to the complimentary R region at the 3' RNA end. Complimentary negative strand DNA synthesis proceeds for the entire genome, with segmented RNase hydrolysis. Two specific RNase cleavage resistant RNA sections (central and 3' polypurine tracts (PPT)) are left intact and remain attached to the negative strand DNA strand, facilitating positive strand DNA synthesis initiation. Positive strand DNA synthesis coincides with RNase cleavage of the PPT sites and the lysine transfer RNA, exposing the primer binding site. Complimentary primer binding sites of the positive and negative DNA strands anneal and is followed by strand displacement synthesis to yield double-stranded HIV DNA with a complete LTR at each end (Huber & Richardsong 1990; Rausch & le Grice 2004; Basu et al. 2008; Esposito, Corona & Tramontano 2012). Reverse transcription often yields highly deleted double-stranded HIV DNA that can be integrated into the human genome (Li & Craigie 2009; Craigie & Bushman 2012; Maldarelli 2016).

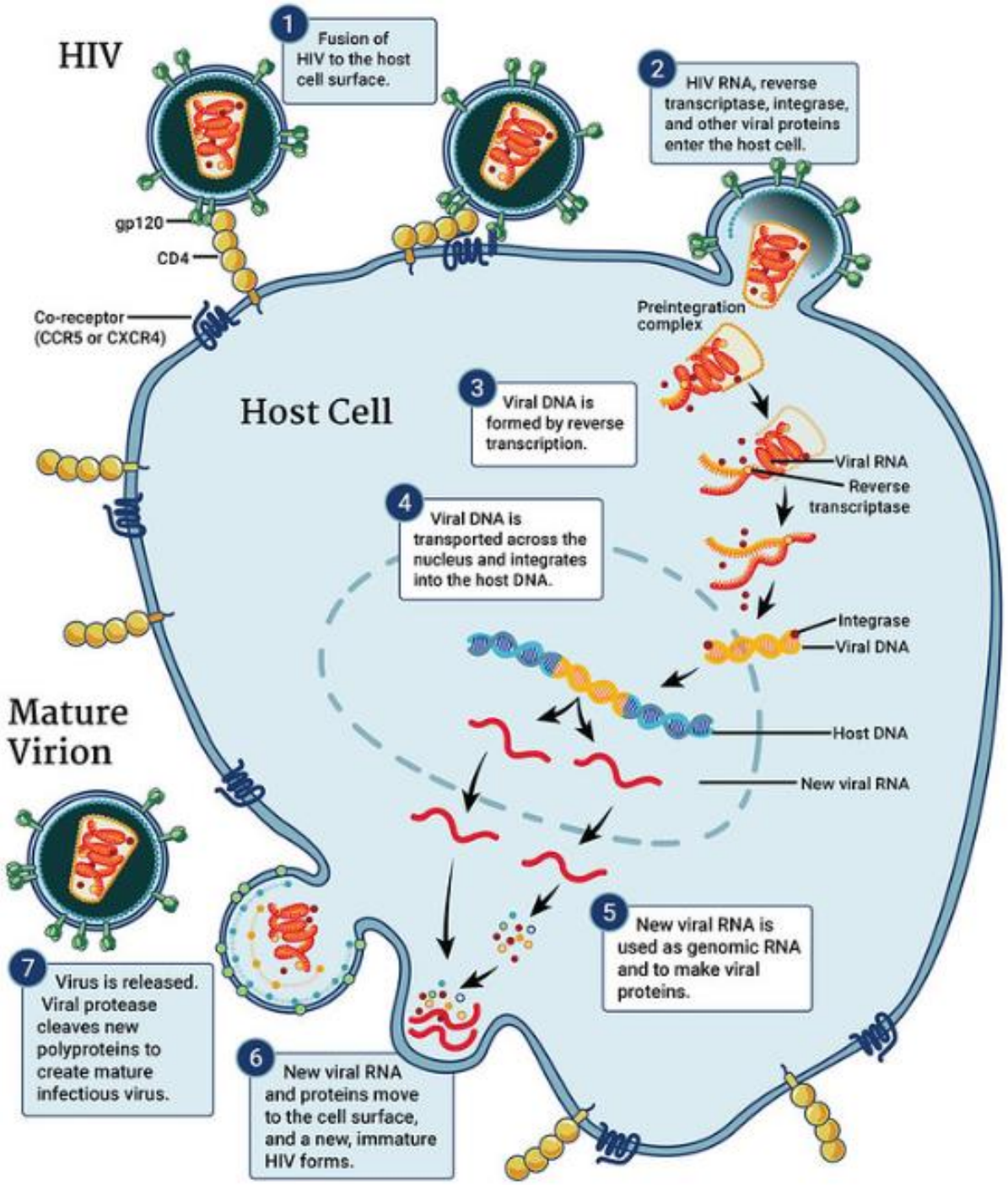


Figure 1.5: Illustration of the HIV replication cycle, indicating the various steps from virion binding to release. Important host and viral factors are also indicated. (NIAID 2018)

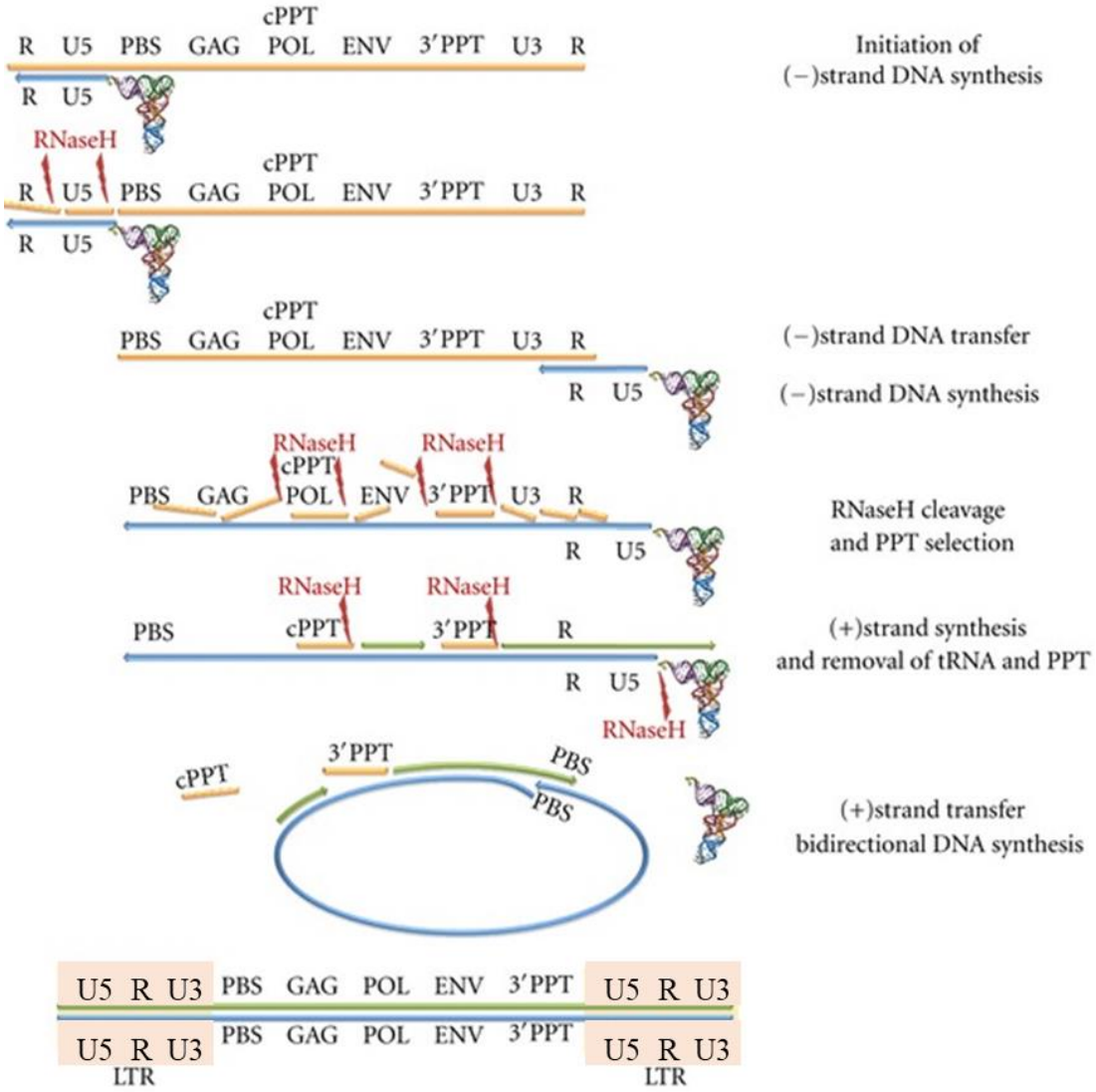


Figure 1.6: Representation of the nine step HIV retro-transcription to yield double-stranded DNA with complete LTR at each end. The HIV genes along with the primer binding site (PBS) and central and 3' polypurine tracts (PPT) are indicated. Image modified from (Esposito et al. 2012).

The double-stranded HIV DNA together with *integrase*, *reverse transcriptase*, matrix protein, Vpr and host DNA-binding protein form the pre-integration complex, and moves toward the nucleus (Miller, Farnet & Bushman 1997; Greene & Peterlin 2006). The viral pre-integration complex enters through the nuclear pores, with the matrix, Vpr and *integrase* proteins assisting in this process (Bukrinsky et al. 1993; Heinzinger et al. 1994; Gallay et al. 1997). In the nucleus, proviral integration proceed into relatively random locations of the host genome, assisted by *integrase*, high-mobility group proteins, barrier-to-autointegration factor and other host factors. Additionally, the *integrase* protein mediates the DNA ligation to establish the provirus in the host genome (Farnet & Bushman 1997; Miller et al. 1997; Chen & Engelman 1998; Lee &

Craigie 1998; L. Li et al. 2000; Lin & Engelman 2003). Proviral integration in transcriptionally active cells is favoured in promoter rich areas, while proviruses in resting cells are often associated with suboptimal gene expression regions (Jordan et al. 2001; Wang et al. 2007; Brady et al. 2009).

Proviral integration can result in transcriptionally active or latent infection. Proviruses can remain latent for years and are not affected by combination antiretroviral therapy (cART) (Adams et al. 1994; Jordan et al. 2001; Greene & Peterlin 2006). The proviral LTR contains promoter elements, including the initiator, TATA-box and three Sp1 transcription factors, which helps position human RNA polymerase II for transcription initiation. Nuclear factor kappa-B, nuclear factor of activated T cells and erythroblast transformation specific transcription factors are also involved in HIV transcription initiation (Jones & Peterlin 1994; Taube et al. 1999; Karin & Ben-Neriah 2000; Barboric et al. 2001). HIV transcription proceeds and is greatly enhanced by Tat to yield HIV mRNA and the various proteins as discussed previously.

Viral particle assembly occurs at the cellular plasma membrane, with each virion constructed with ~1500 p55 *gag* polyproteins, ~100 *gag*-Pol polyproteins, Vpr and two copies of HIV RNA (Freed 1998; Wilk et al. 2001). Virion assembly is further assisted by human proteins and HIV Nef and envelope proteins (gp41 and gp120) (Pandori et al. 1996; Zimmerman et al. 2002; Ganser-Pornillos, Yeager & Sundquist 2008; Sundquist & Krausslich 2012). The *gag* polyprotein associates with the cell membrane due to cholesterol and glycolipid tropism. Therefore, virus budding results in cholesterol rich virion envelopes which favours virion release, stability and fusion of targeted cells (Göttlinger et al. 1989; Campbell, Crowe & Mak 2001; Ono & Freed 2001; Ono 2009).

Virion budding employs the host "endosomal sorting complexes required for transport" pathway, recruited by two late domain motifs of HIV p6 protein. This results in successful virion release from the infected cell (Garnier et al. 1999; Strack et al. 2000; Usami et al. 2009; Peel et al. 2011). Virion maturation initiates during the budding process, driven by the proteolytic cleavage of the gag and gag-Pol polyproteins, resulting in fully formed infectious HIV virions (Briggs 2003; Benjamin et al. 2005; Hill, Tachedjian & Mak 2005; Briggs et al. 2006).

1.4 Transmission and Epidemiology

The HIV pathogen can be detected in blood, semen, rectal and vaginal secretions, and breast milk of infected individuals. As a result, the most common routes of HIV transmission are sexual contact (heterosexual and homosexual), injection drug users and mother to child transmission (Shaw & Hunter 2012; WHO 2020).

Analyses of samples collected in 1959 and the 1960s indicate HIV genetic diversification, suggesting that HIV was already circulating as a "silent" epidemic prior to 1930 (Worobey et al. 2008). However, it was only after discovering an epidemic of acquired immunodeficiency in 1981, predominantly among young men who have sex with men in the USA and Europe, with the subsequent cell culture isolation of HIV in 1983, that the scientific community became aware of HIV as the cause of acquired immunodeficiency syndrome (AIDS) (Gottlieb et al. 1981; Levy 1993). Today the pandemic continues to grow worldwide despite substantial treatment and prevention strategies, with about 38 million people living with HIV (UNAIDS 2020). South Africa remains the country with the biggest HIV epidemic. Current data indicates ~200 000 new infections were reported over the last year, accruing in a total of ~7.5 million HIV positive individual (UNAIDS 2020). In order to drastically reduce the HIV pandemic a 90-90-90 plan was put into place with the aim to diagnose 90% of HIV positive individuals, ensure treatment for 90% of these individuals and to achieve viral suppression in 90% of treated patients by end 2020. Even though over 25 million HIV positive people are currently on treatment, the 90-90-90 objectives are still not achieved, with current data indicating 79% diagnosed, 78% treatment coverage and 86% viral suppression (Avert 2020; UNAIDS 2020). The 2030 goal of eliminating AIDS is therefore off track and significant efforts are required if this is to be achieved (UNAIDS 2020).

1.4.1 Global distribution of HIV

HIV-1 group M is responsible for the global HIV pandemic, with subtype C accounting for more than 50% of all cases, followed by subtypes A and B at ~12% each. Several recombinant strains including CRF01_AE, CRF02_AG and CRF07_BC have also been detected in Asia and West Africa and represent about 5% of cases each (Hemelaar et al. 2011; Bbosa, Kaleebu & Ssemwanga 2019). HIV subtype A predominates in East Africa, Russia and Asia, with CRF02_AG and other recombinant types widely detected in Asia (Figure 1.7) (Saksena et al. 2011; Gounder et al. 2017; Aibekova et al. 2018). Subtype B predominates in North and South

America, Europe, Australia, the Middle East and North Africa (Castley et al. 2017; Daw et al. 2017; Lima et al. 2017; Oster et al. 2017; Sallam, Şahin, et al. 2017; Sallam, Esbjörnsson, et al. 2017; Alexiev et al. 2018; Hebberecht et al. 2018; Hernandez-Sanchez et al. 2018; Tumiotto et al. 2018; Volz et al. 2018). Subtype C is predominantly detected in Southern Africa and India (Ngcapu et al. 2017; Sharma et al. 2017; Sivay et al. 2018). Subtype D is not a predominant strain but circulates in East Africa along with subtypes A and C (Billings et al. 2017).

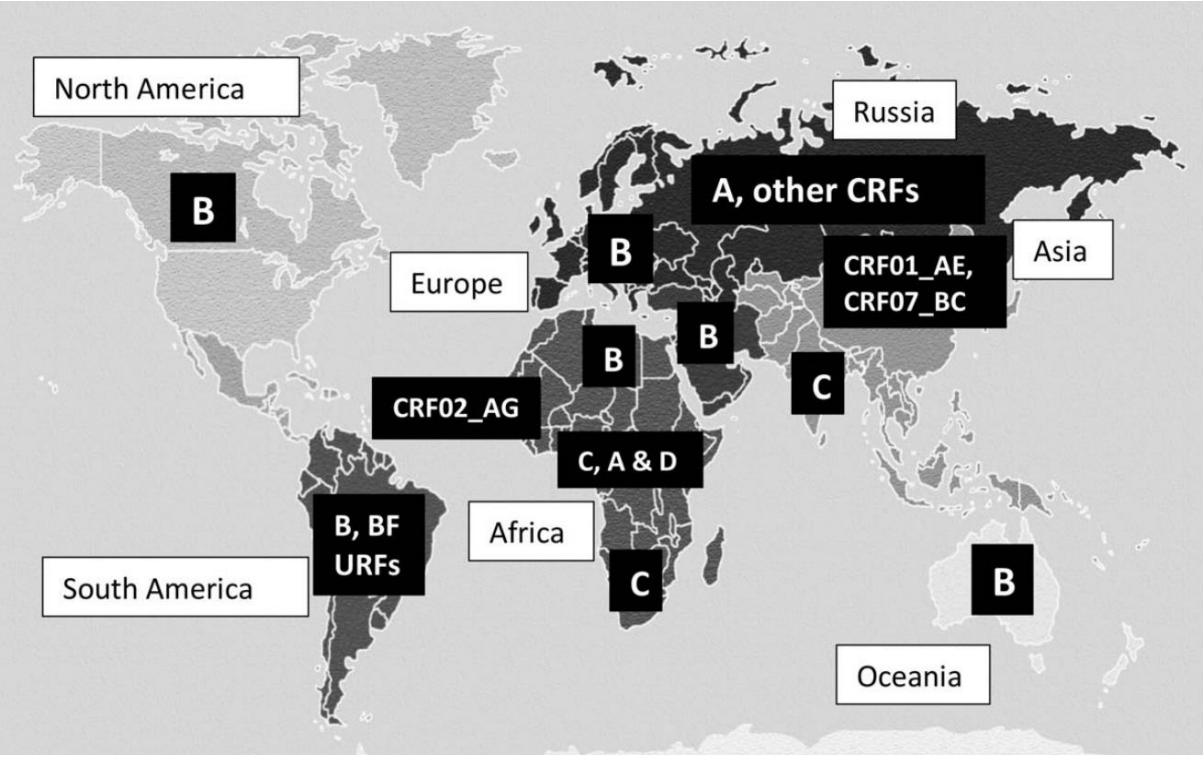


Figure 1.7: World map indicating the global distribution of major HIV-1 group M subtypes and circulating recombinant forms (CRF) responsible for the pandemic. (Bbosa et al. 2019)

1.5 Diagnosis and treatment

Although there is a higher HIV testing coverage today than ever, there is still some way to go in order to reach the UNAIDS goal of 90% of people living with HIV knowing their status (UNAIDS 2020). HIV testing is mostly done by two approaches, serology assays and nucleic acid detection. HIV serological assays can be laboratory-based enzyme immunoassays or point of care rapid lateral flow assays. The latter is referred to as the HIV rapid serological assay (CDC 2020). These rapid tests offer sensitivity above 95% and specificity above 99% and are often used for community level diagnosis (Boadu et al. 2016). Laboratory-based assays have developed through several generations, with current fourth-generation enzyme immuno-assays

detecting HIV p24 antigen in addition to antibody, in order to shorten the diagnostic window period to 15-20 days (Ananworanich et al. 2013). Rapid assays have also been adapted to detect HIV p24 antigen, but in general are less sensitive for antigen than laboratory based fourth generation HIV serology (CDC 2020). Qualitative and quantitative nucleic acid assays are used to confirm HIV infection in children before the age of 18 months and are used to diagnose acute HIV in adults as these assays detect HIV before antigen or antibodies are detectable (Cohen et al. 2010). Quantitative HIV-1 RNA assays are used to monitor the HIV viral load during infection and treatment. Molecular based assays can target HIV DNA or RNA, where DNA will detect HIV provirus and RNA will detect HIV cellular mRNA and virion RNA. For this reason, detection of active circulating virus by HIV RNA assays, are performed on plasma samples devoid of cellular material (Templer et al. 2016; Ochodo et al. 2018). The window period following infection before HIV can be detected with molecular assays is at least 10 days (Cohen et al. 2010; CDC 2020).

Since the discovery of antiretroviral drugs in 1987, followed by the adaptation to dual drug therapy in 1992 and finally triple drug regimens in 1996, ART development has been revolutionised and is more accessible than ever before (Arts & Hazuda 2012; Kemnic & Gulick 2020). HIV treatment with cART substantially improves patient survival and quality of life by suppressing viral replication and preventing disease progression, but it is not a cure (Chun et al. 1997; Broder 2010). Several groups of antiretrovirals, with multiple drugs each, have been developed. The most common groups of drugs are nucleoside/non-nucleoside reverse transcription inhibitors (NRTIs and NNRTIs), protease inhibitors, *integrase* inhibitors, fusion inhibitors, co-receptor antagonists and pharmacokinetic enhancers. Each of these drug groups target a specific step of the HIV replication cycle as stated in the name, and activity can be augmented by the pharmacokinetic enhancers (Figure 1.8) (Engelman & Cherepanov 2012; Laskey & Siliciano 2014). Although cART suppresses the viral load and prevents new rounds of cellular infection, it does not directly affect the HIV provirus (Kearney et al. 2017; Rosenbloom et al. 2017; van Zyl et al. 2017; McManus et al. 2019). These latently infected $CD4^+$ T cells can survive for decades in patients on HIV suppressive therapy (Chun et al. 1997; Vanhamel, Bruggemans & Debyser 2019).

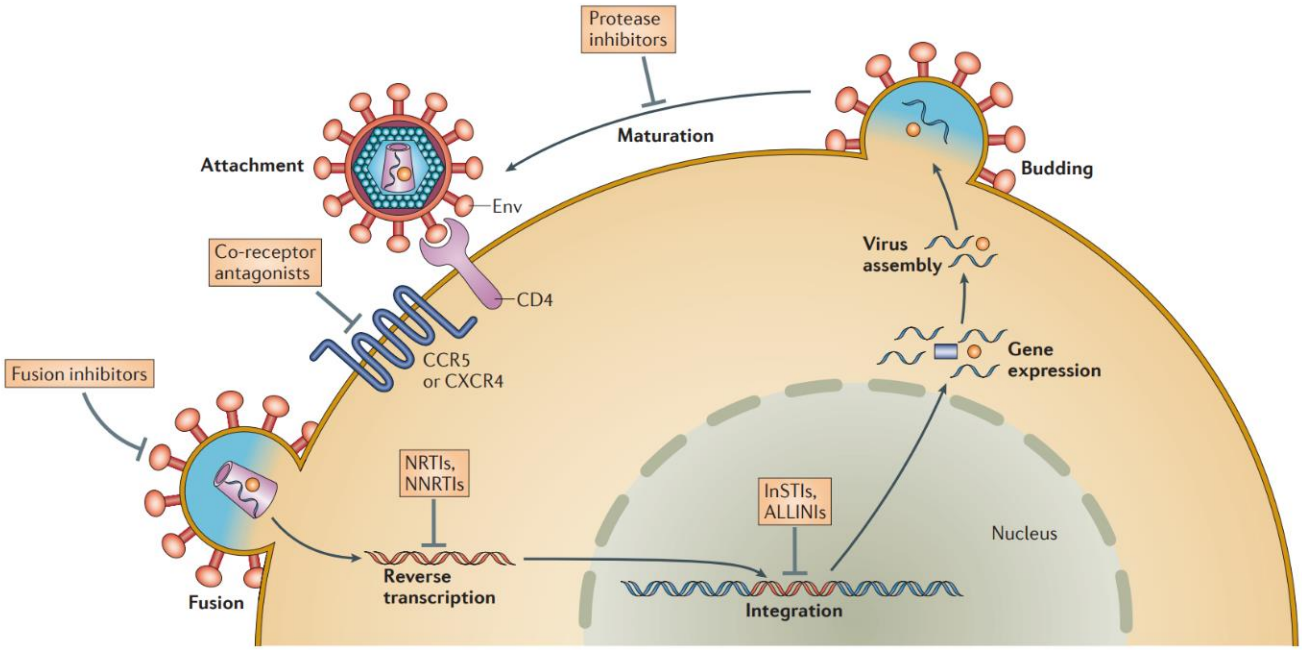


Figure 1.8: Representation of the HIV replication cycle, indicating the various targets for antiretroviral drugs including, nucleoside/non-nucleoside reverse transcription inhibitors (NRTIs and NNRTIs), protease inhibitors, integrase inhibitors (integrase strand transfer inhibitors (InSTIs) and allosteric integrase inhibitors (ALLINIs)), fusion inhibitors and co-receptor antagonists. (Laskey & Siliciano 2014)

1.5.1 Low-level viraemia and virologic failure

Several factors affect HIV-1 viral load thresholds: 1) The threshold of commercial assays; 2) clinical impact of elevated viral loads on health and development of drug resistance; 3) whether drug resistance testing is feasible at a particular threshold. Initial commercial HIV-1 RNA assays had thresholds around 400-500 copies/mL (O'Shea et al. 2000); and drug resistance assays had initially been validated for viral loads > 1000 copies/mL (Mitsuya et al. 2006), with inhouse assays validated for lower viral loads (Yang et al. 2010). Thereafter "ultrasensitive" HIV-1 viral load assays became available with a lower limit of detection of 50 copies/mL, but the impact on clinical practice was not pronounced as most individuals failed with relative higher viral loads (Muir et al. 2000). Since then, HIV-1 viral load has been considered suppressed when quantitative RNA assays indicate the plasma viral load as <50 copies/mL for over 6 months (EACS 2018; AIDSinfo 2020). Viral suppression is generally achieved in 8 – 24 weeks after cART initiation, in treatment adherent individuals (AIDSinfo 2020). With viral load assays becoming more sensitive, an increase of sustained and episodic low-level (50 – 1000 copies/mL) and very low-level or residual (<50 copies/mL) viraemia has been observed (Greub et al. 2002; Palmer et al. 2008). The implications of prolonged detection of low-level

viraemia require further investigation, while some studies indicate an increased risk of subsequent virologic failure, this association was not observed in other studies (Damond et al. 2007; Gatanaga et al. 2009; Willig et al. 2010; Laprise et al. 2013; Hermans et al. 2018; AIDSinfo 2020). The office of AIDS research advisory council (OARAC) and AIDS clinical trials group (ACTG) therefore adapted their guidelines on HIV treatment to define a confirmed viral load of over 200 copies/mL as virologic failure (Aldous & Haubrich 2009; AIDSinfo 2020). However, other groups define virologic failure as a viral load above 50 copies/mL, to avoid acquisition of HIV drug resistance mutations (Nettles et al. 2005; Mitsuya et al. 2006; Doyle et al. 2012; Henrich, Wood & Kuritzkes 2012). In contrast, low- and middle-income countries, including South Africa, adopted the WHO treatment guidelines which defines treatment failure at a viral load of greater than 1000 copies/mL (World Health Organisation 2016). Although there are no clear guidelines of low-level viraemia treatment in clinical settings, some studies recommend immediate treatment adjustment to fully active regimens to avoid negative outcomes (Nasta, Castelnuovo & Paraninfo 2005; Cohen 2009; Ryscavage et al. 2014). These discordant definitions of low-level viraemia and virologic failure also complicate comparative analyses on treatment failure and outcomes between different patient subsets.

HIV virologic/treatment failure occurs when the cART regimen no longer controls HIV viraemia. The main contributing factors to virologic failure are inadequate drug exposure, predominantly due to inadequate adherence and treatment relevant drug resistance (ie. drug resistance mutations against the administered treatment). Poorly tolerable regimens, and drug toxicity impact adherence and drug exposure, whereas inadequate drug exposure could contribute to the acquisition of drug resistance. HIV variants that harbour drug resistance mutations can also be transmitted, which increases the risk of subsequent treatment failure in newly infected individuals (AIDSinfo 2020). HIV reverse transcription lacks proofreading, which results in the emergence of multiple variants, some of which by chance may harbour drug resistance mutations. In the absence of prior treatment exposure or transmitted drug resistance, these drug resistant variants mostly occur as single mutations in a particular variant and some are very rare considering the fitness price of particular resistance mutations (Coffin 1995; O'Neil et al. 2002). Therefore, in the absence of prior drug exposure or transmission of drug resistant variants, cART effectively suppresses HIV replication and evolution, thereby preventing the further emergence of drug resistance (Maggiolo et al. 2005; Bangsberg 2006; Nachega et al. 2007; Martin et al. 2008). In contrast, poor or non-adherence to treatment could

create an environment where HIV could replicate in the presence of non-suppressive drug levels resulting in the emergence of drug resistant HIV strains and disease progression (Bangsberg et al. 2001; Harrigan et al. 2005; Nachega et al. 2011; Hassan et al. 2014; Ahmed, Merga & Jarso 2019; Endalamaw et al. 2020). Acquired drug resistance due to poor adherence is therefore one of the most important predictors of virologic failure. Individuals presenting with viraemia indicating treatment failure require immediate adherence counselling and HIV drug resistance sequence analyses to assist in the selection of an active treatment regimen (Kiweewa et al. 2019; AIDSinfo 2020). Some antiretroviral drugs require multiple mutations before phenotypic drug resistance and clinical treatment response are sufficiently impacted. These are referred to as high genetic barrier regimens and are very valuable in treatment-experienced patients (Clotet et al. 2007; Banhegyi et al. 2012; Eron et al. 2013). Recently, with the introduction of first-line regimens comprising high genetic barrier second generation integrase inhibitors such as the widely used combination of tenofovir with lamivudine and dolutegravir (TLD), the risk of drug resistance and treatment failure are likely to decrease substantially (Brenner & Wainberg 2017; Phillips et al. 2019).

1.5.2 Transient viraemia

Transient viraemic episodes, also referred to as viraemic blips, are quite common in patients on suppressive cART. However, definitions of viral blips are discordant, including a single viral load above 200 copies/mL (AIDSinfo 2020), viral loads between 50 and 500 copies/mL for up to six weeks (Sörstedt et al. 2016) or any viral load above 50 copies/mL followed by suppression to below 50 copies/mL (Havlir 2001). The WHO guidelines, followed in South Africa, define single viraemic episode up to 1000 copies/mL as a viral blip (World Health Organisation 2016). The relevance of viral blips is also disputed; while some studies suggest limited or no clinical significance and suggest no treatment change (Havlir et al. 2001; Lee et al. 2006; Garcia-Gasco et al. 2008), other studies observed associations between viral blips and virologic failure and suggest prompt action (Young et al. 2015; Sörstedt et al. 2016). Several factors have been associated with transient viraemic episodes including, high baseline viral loads, poor adherence to cART, opportunistic infections and vaccinations (Günthard et al. 2000; Jones & Perelson 2002; Kolber et al. 2002; Jones & Perelson 2005; Fung et al. 2012; Sörstedt et al. 2016). Further investigation of viral blips is essential to standardise its definition and clarify the potential clinical implications.

1.6 Proviral latency and persistence

HIV proviruses can remain latent for decades while the viral load is suppressed by cART (Adams et al. 1994; Chun et al. 1997; Siliciano et al. 2003; Strain et al. 2003; Vanhamel et al. 2019). The genomic location of the proviral integration site has been associated with latency. While integration is favoured in promoter rich areas of transcriptionally active cells, proviruses are often associated with suboptimal gene expression regions in resting cells (Jordan et al. 2001; Wang et al. 2007; Brady et al. 2009). Several molecular mechanisms have been implicated in maintaining proviral latency including, epigenetic changes in chromatin (Jordan, Bisgrove & Verdin 2003), DNA methylation (Blazkova et al. 2009; Kauder et al. 2009) and transcriptional interference (Lenasi, Contreras & Peterlin 2008; De Marco et al. 2008; Shan et al. 2011). Resting cells further maintain latent proviruses by blocking HIV transcript elongation (Adams et al. 1999; Lin et al. 2003). This blockage is reversed after T cell activation, as observed by an increased ratio of elongated HIV transcripts (Kaiser et al. 2017; Yukl et al. 2018). The underlying mechanism of proviral latency is therefore thought to be multifactorial (Lassen et al. 2004; Siliciano & Greene 2011). Although many proviruses of various levels of intactness can latently infect any of the above-mentioned cell types and persist during cART, the replication competent proviruses that yield infectious virions are defined as the reservoir (Siliciano & Greene 2011).

Human $CD4^+$T cells are the primary HIV-1 reservoir, with long-lived memory T cells likely the most important in harbouring latent provirus (Ahmed et al. 2009; Restifo & Gattinoni 2013; Doitsh et al. 2014). The large memory T cell subset, central memory T cells, is considered one of the most significant reservoirs by persisting through T cell survival and cellular proliferation in addition to being very long lived (Chomont et al. 2009; Soriano-Sarabia et al. 2014; Barton, Winckelmann & Palmer 2016; Pallikkuth et al. 2016). HIV infected memory T cells have been identified in the gut-associated lymphoid tissue and lymph nodes during suppressive therapy (Chun et al. 2008). Elevated levels of HIV-1 DNA in the gut-associated lymphoid tissue and lymph nodes have been observed when compared to the peripheral blood (Chun et al. 2008; Yukl et al. 2010; Banga et al. 2016; Khoury et al. 2017). However, the penetration of cART drugs into the gut-associated lymphoid tissue was found to be sufficient to prevent viral replication, as observed by a lack of HIV proviral DNA diversification (Josefsson, von Stockenstrom, et al. 2013; von Stockenstrom et al. 2015). Other cells and tissues have also been implicated in harbouring HIV. HIV-DNA has been detected in latently infected bone marrow $CD133^+$ and mast cells of patients on cART (Sundstrom et al. 2007; McNamara et al. 2013).

Ex vivo analyses on purified liver Kupffer cells from patients on suppressive therapy yielded infectious virus (Kandathil et al. 2015). The presence of HIV in the cerebrospinal fluid and brains of some infected individuals has been well described (Ho et al. 1985; Levy et al. 1985), with multiple studies suggesting compartmentalised central nervous system infection (Koyanagi et al. 1987; Cheng-Mayer & Levy 1988; Chiodi et al. 1989; Korber et al. 1994; Wong et al. 1997; Strain et al. 2005; Harrington et al. 2009; Schnell et al. 2009). HIV DNA and RNA has been detected in the lungs of treated and untreated HIV positive patients. Lung $CD8^+$T cells and macrophages were found to harbour HIV (Sierra-Madero et al. 1994; Semenzato et al. 1995; Jambo et al. 2014) and $CD4^+$T cells obtained from the lungs showed 13 fold higher levels of HIV DNA than peripheral blood (Costiniuk et al. 2018). Additionally, in one study, some level of compartmentalisation was detected in HIV envelope sequences from lung tissues of AIDS patients (van't Wout et al. 1998). HIV virions, DNA and RNA have also been detected in kidney podocytes and tubular cells, and urine of cART treated patients (Winston et al. 2001; Marras et al. 2002; Chakrabarti et al. 2009; Canaud et al. 2014). The male and female reproductive tracts have also been found to harbour HIV while on suppressive therapy, in the prostate, testes, semen (Lecatsas et al. 1985; Borzy, Connell & Kiessling 1988; Silva et al. 1990; Baccetti et al. 1991; Pudney & Anderson 1991; Wolff et al. 1992; Bagasra et al. 1994; Muciaccia, Uccini, et al. 1998; Muciaccia, Filippini, et al. 1998; Shevchuk, Nuovo & Khalife 1998; Lafeuillade et al. 2002; Marcelin et al. 2008; Sheth et al. 2009; Halfon et al. 2010; Politch et al. 2012) and female genital secretions (Kovacs et al. 2001; Nunnari et al. 2005; Launay et al. 2011; Wahl et al. 2011; Fiscus et al. 2013), with evidence of compartmentalisation (Overbaugh et al. 1996; van't Wout et al. 1998; Paranjpe et al. 2002; Smith et al. 2004; Tirado et al. 2004; Coombs et al. 2006; Chomont et al. 2007; Bull et al. 2009; Kelley et al. 2010). In addition to all these sources, limited traces of HIV have been identified in the thymus (Deere et al. 2014), adipose tissues (Couturier et al. 2015; Damouche et al. 2015) and CCR6 tropic rectal $CD4^+$ T cells (Anderson et al. 2020) of treated individuals.

In peripheral blood of chronically infected individuals, significant decay (~86% decline) of HIV DNA is observed in the first 12 months after treatment initiation, followed by further decay (~23% per year) up to year four, after which proviral DNA levels stabilise (Besson et al. 2014). In contrast, treatment initiation during acute infection resulted in substantially sharper HIV DNA decay and the decline continued past four years on treatment (Hocqueloux et al. 2013; Ananworanich et al. 2016). In fact, the sooner after infection cART is initiated, the more dramatic the HIV DNA decay (Laanani et al. 2015). Similar observations made in children

associated delayed cART initiation with slower decay of HIV DNA, and treatment initiation in the first week after birth with a significant reduction of long-lived infected cells (Veldsman et al. 2019). However, after the initial years of decay the levels of HIV DNA in peripheral blood remains remarkably stable (Tierney et al. 2003; Viard et al. 2004; Besson et al. 2014). When cART is interrupted, long-lived latently infected cells are the source of rebounding virus, again confirming the importance of these cells in maintaining the HIV reservoir (Chun et al. 1997). The levels of HIV DNA in peripheral blood is a predictor of the time to viral rebound after cART interruption, while higher levels are associated with accelerated rebound (Piketty et al. 2010), low DNA levels have been associated with a higher probability of maintained virologic control (Williams et al. 2014; Assoumou et al. 2015). However, a recent study observed the emergence of large, intact HIV RNA populations after cART was ceased, and the re-seeding of tissue sites throughout the body (Chaillon et al. 2020). This finding suggests that treatment interruption studies could have serious consequences in terms of the repopulation of the proviral reservoir although the clinical implications remain unclear (Julg et al. 2019).

1.6.1 Proviral clones

Phylogenetic analyses indicate a lack of HIV-1 proviral genomic evolution over time, suggesting that ongoing viral replication is not the main mechanism for reservoir persistence (Josefsson, von Stockenstrom, et al. 2013; Wagner et al. 2013; von Stockenstrom et al. 2015; Kearney et al. 2016; Kearney et al. 2017). Due to the size of the human genome (3.0×10^9 base pairs) and relatively random HIV integration, cells that harbour the same integration site represent cellular clones (Simonetti et al. 2016; Cohn & Nussenzweig 2017). The appearance of monotypic HIV proviral sequences in patients on long-term therapy suggests that $CD4^+$ T cell clonal expansion, rather than ongoing replication, is responsible for reservoir maintenance, as proven by integration site analyses (Maldarelli et al. 2014; Reeves et al. 2018). Clonal expansion is a major mechanism for maintaining the reservoir, as these proliferated cells can harbour replication competent proviruses (Lorenzi et al. 2016; Simonetti et al. 2016; Reeves et al. 2018; de Scheerder et al. 2019). The rarity of intact proviruses is also reflected in the clonal population, as most of the identified proviral clones are defective (Imamichi et al. 2014; Cohn et al. 2015). Latently infected cells can stably proliferate without virion expression and infection of further cells (Hosmane et al. 2017). The majority of these long-lived clonally expanded cells contain only one provirus (Josefsson et al. 2011; Josefsson, Palmer, et al. 2013)

however, ex vivo experiments indicated that a single proliferated clone can be responsible for a large proportion of expressed infectious virions (Bui et al. 2017). Furthermore, Hosmane et al highlighted that in vivo proliferation of cells containing intact provirus could occur without producing virus and about 57% of the isolated replication competent proviruses in a single patient were clones (Hosmane et al. 2017). A recent study suggested that the HIV reservoir consists of a few large clones and many small clones (Reeves et al. 2018). Proviral landscape dynamics are further complicated, as intact and defective proviruses containing elements that favour protein expression decrease over time as a result of immune clearance (Pinzone et al. 2019). Recent reports on clonal waxing and waning suggest that antigen-driven and cytokine-induced proliferation, rather than continuous cell-autonomous proliferation, contribute to clonal expansion and the stability of the latent reservoir (Wang et al. 2018; Simonetti et al. 2020).

Understanding the mechanisms of HIV persistence is pertinent in designing future approaches to HIV cure or achieving remission. Clonal proliferation is increasingly recognised as an important mechanism of persistence. Intact expanded proviral clones present a barrier to HIV cure. In addition, cells harbouring defective proviruses may produce viral RNA and proteins, which drive chronic immune activation and hamper the clearance of cells harbouring the intact reservoir (Jones et al. 2017; Pollack et al. 2017; Pinzone et al. 2019).

1.6.2 Clonal viraemia

Residual viraemia, defined as a viral load below 50 copies/mL, has been shown to often constitute clonal viraemia, evident by monotypic plasma viraemia. This suggests that virus produced by clonally proliferated cells are drug sensitive and show no evidence of ongoing replication and evolution (Nettles et al. 2005; Tobin et al. 2005; Bailey et al. 2006; Anderson et al. 2011). The majority of CD4 clones contain replication defective proviruses however, they could become reactivated, express protein, and some defective viruses are able to produce and package viral RNA (Cohn et al. 2015; Rassler et al. 2016; Pollack et al. 2017). In contrast to residual viraemia, viraemia detectable by commercial viral load assays has until recently been considered a reflection of poor adherence or drug resistance. Nevertheless, a recent case of HIV-1 plasma RNA viraemia between 100 and 200 copies/mL represented virus release from a replication competent infectious clone for over 13 years, until the patient passed away from metastatic squamous cell carcinoma (Simonetti et al. 2015). Another case with transient

viraemia at two time points (viral loads of 297 and 100 copies/mL) represented viruses released from a defective clone, followed by successful viral suppression (Rassler et al. 2016). Cells harbouring hypermutated proviruses are unlikely to produce viral particles if these hypermutations occur in *gag*, as a functionally intact *gag* is essential for capsid assembly and RNA packaging (Delchambre et al. 1989; Pollack et al. 2017). However, it is possible that hypermutated proviruses may have intact *gag* or alternatively, defective proviral genomes may be packaged when complemented by another provirus in the same cell, which has an intact *gag* gene (Iwabu et al. 2008).

With the advent of potent and tolerable antiretroviral regimens, the most common causes of failure, including inadequate adherence and drug resistance may diminish and rarer cases of unexplained viraemia such as clonal viraemia may become more apparent. These cases require proper characterization (Jacobs et al. 2019). Further understanding of clonal viraemia would provide invaluable guidance for clinical management of patients with evidence of good adherence, in the absence of drug resistance, but with persisting HIV-1 viraemia.

1.7 The proviral landscape

The replication competent HIV reservoir is maintained during suppressive cART and rebounds after treatment interruption (Ho et al. 2013). However, the reservoir is vastly outnumbered by other proviruses of various intactness. In fact, only about 2% of the proviral landscape is represented by intact replication competent virus (Figure 1.9) (Bruner et al. 2016).

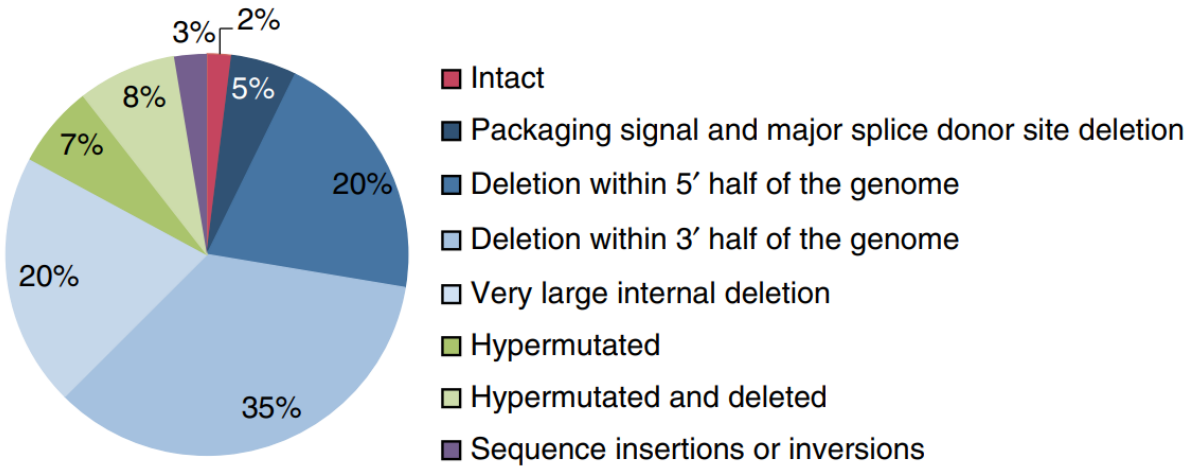


Figure 1.9: Representation of the proviral landscape characterised from 152 near full-length HIV sequences of 10 patients. (Bruner et al. 2016)

The remaining 98% of proviruses are defective in one or several ways. HIV proviruses with deletions within either the 5' or 3' genome halves or very large internal deletions made up about 75% of the 152 near full-length sequences from 10 patients (Figure 1.9) (Bruner et al. 2016). Approximately 5% of the investigated sequences had deletions within the packaging signal and major splice sites and about 3% of sequences contained insertions or inversions (Ho et al. 2013; Bruner et al. 2016). Hypermutations as a result of APOBEC3G and APOBEC3F activity is responsible for rendering between 15% and 32% of proviruses as defective (Ho et al. 2013; Bruner et al. 2016).

Importantly, due to the location of the primers used for near full-length amplification, severely deleted proviruses consisting of only a single LTR or partial LTR would be completely overlooked and therefore remains largely uncharacterised in HIV-1 infection (Bruner et al. 2016; Hiener et al. 2017). A recent study identified a predominant solo-LTR proviral clone comprising 25% of the proviral landscape (Anderson et al. 2020). Solo-LTR and partial-LTR proviruses are results of ineffective viral reverse transcription, yielding severely deleted genomes, limited to LTR sequences, and subsequently integrated into the host genome (Li & Craigie 2009; Craigie & Bushman 2012; Maldarelli 2016). Solo-LTR proviruses are well described in other retroviruses (Katzourakis et al. 2007) and comprise up to 90% of human endogenous retroviruses (Buzdin et al. 2006; Contreras-Galindo et al. 2013; Thomas, Perron & Feschotte 2018). Further investigation of the distribution and possible activity of solo-LTR and partial-LTR proviruses is required, which could significantly influence the proposed proportions of the proviral landscape.

1.7.1 Assays to characterise the proviral landscape

A multitude of assays to quantify the HIV reservoir and characterise the proviral landscape in patients on suppressive cART have been developed, each with its own benefits and limitations (Massanella & Richman 2016; Sharaf & Li 2017).

1.7.1.1 Cell based assays

Quantitative viral outgrowth assays (QVOA) to determine the infectious units per million cells have been described as the gold standard to quantify the proviral reservoir. With QVOA a limiting dilution cell culture format is performed after one round of stimulation to activate resting cells, and HIV p24 antigen is measured after several days of culture (Finzi et al. 1997;

Siliciano & Siliciano 2005). This assay is useful to demonstrate the stability of the proviral reservoir in HIV suppressed patients, with the main benefit that only replication competent provirus is observed (Siliciano et al. 2003; Crooks et al. 2015). However, a major limitation of QVOA is that only a fraction of replication competent proviruses is induced, resulting in a large proportion of the reservoir being overlooked (Ho et al. 2013; Bruner et al. 2016). Additionally, QVOA is resourse intensive and could yield variable results (Ho et al. 2013). Recently QVOA was performed with multiple rounds of stimulation, increasing the number of inducible intact proviruses and yielding increased quantities of infectious virions from more proviruses. Although the multiple stimulation QVOA approach yields more inducible proviruses, it still underestimates the total reservoir, in addition to being even more resource and labour intensive (Hosmane et al. 2017).

Quantifying RNA from inducible proviruses is considered a viable alternative to the resource and labour-intensive viral outgrowth assay and offers a more accurate representation of the inducible reservoir (Hermankova et al. 2003; Wei et al. 2014). However, RNA transcripts from defective proviruses could lead to an overestimation of the reservoir (Kearney et al. 2016).

1.7.1.2 Nucleic acid quantification

Several HIV DNA assays have also been developed to quantify various subsets of the proviral landscape. Proviral DNA quantification assays targeting various HIV regions including *gag* (Nottet et al. 2009; Klatt et al. 2014), *reverse transcriptase* (Strain et al. 2013), LTR (Hatano et al. 2013; Yukl et al. 2013; Yukl et al. 2014) and *integrase* (Cillo et al. 2013; Hong et al. 2016) have been developed. Due to the high levels of genetic variation in the *gag* and *reverse transcriptase* regions, the *integrase* targeted HIV DNA assay was found to be superior, offering single-copy sensitivity (Cillo et al. 2014). Proviral LTR quantification found 10-fold higher DNA levels than the other assays, due to the detection of highly deleted proviruses (Rozera et al. 2010). Some studies suggest that the total DNA proviral load reflect the size of the proviral reservoir and can predict time to viral rebound after cART interruption (Yerly et al. 2004; Williams et al. 2014). The HIV DNA targeted assays can quantify many, but not all defective proviral variants, resulting in an under-estimation of the total proviral load and a severe over-estimation of the reservoir (Eriksson et al. 2013; Ho et al. 2013; Bruner et al. 2016). Although the quantification of proviral DNA has been heavily criticised for over-estimating the reservoir, even defective proviruses can produce viral proteins and thereby participate in HIV pathogenesis (Imamichi et al. 2016). It is therefore essential to measure the entire proviral

landscape, as each proviral subset contributes to the complex pathophysiology of HIV (Rouzioux & Richman 2013; Rouzioux & Avettand-Fenoël 2018). To bridge this gap, an assay to simultaneously quantify and distinguish between intact and defective proviruses was recently developed. This assay targets the HIV packaging signal and Rev response element regions, with intact proviruses being quantifiable in both regions. Hypermutated proviral genomes are also effectively discerned by allelic discrimination probes targeting the Rev response element (Bruner et al. 2019).

Quantification of the different RNA transcripts (spliced, incompletely spliced and un-spliced) can be performed by targeting various regions in the HIV genome. Total HIV RNA transcript can be quantified by targeting the LTR (Hatano et al. 2013; Yukl et al. 2013; Yukl et al. 2014). Spliced HIV RNA is quantified by targeting the Tat and Rev regions (Pasternak et al. 2008), while targeting the *gag* region will quantify un-spliced RNA, a predictor of time to rebound after treatment initiation (Li et al. 2015).

1.7.1.3 Sequencing based assays

Single proviral template amplification and nucleotide sequencing of the envelope and partial *gag-polymerase* regions can be used for proviral characterisation. These single template assays offer immense value in studying the inter-patient proviral variability (Palmer et al. 2005; McKinnon et al. 2011; Evering et al. 2014; Nolan et al. 2017).

Near full-length HIV amplification and sequencing of single template proviruses has also been performed and add immense value in studying intact and defective proviruses. These near full-length assays offer insights into the proviral landscape and can differentiate between the various proviral subsets (Li et al. 2007; Ho et al. 2013; Bruner et al. 2016). The identification of identical near full-length proviral sequences indicate a strong probability that the source of these sequences is in fact clonal (Hiener et al. 2017). However, the only true indicator of clonality is the unique integration sites (Bailey et al. 2006; Maldarelli et al. 2014; Wagner et al. 2014). The investigation of integration sites can shed light on the size of any specific clone and its relative proportion compared to other proviral clones detected in a patient.

Various approaches have been developed to obtain nucleotide sequences of proviral integration sites. One approach is to target the *Alu* repeat region of the human genome together with HIV *gag*. The nucleotide sequences should reveal partial HIV sequence along with the unique integration site and adjacent human genome sequence (Chomont et al. 2009; de Spiegelaere et al. 2014). However, a major disadvantage of this method is that integration sites can only be

determined for proviruses containing *gag*, which excludes severely deleted proviruses. Alternatively, LTR targeted assays can successfully determine the integration site of any provirus, irrespective of intactness but provide no or very limited proviral sequence data (Maldarelli et al. 2014; Simonetti et al. 2016). A recent development that uses limiting dilution genomic DNA input followed by whole genome amplification, with subsequent sequence-based characterization of proviruses and their linked integration sites allows for the identification and linkage of both, but is costly and cumbersome (Patro et al. 2019). This is important for an improved understanding of HIV reservoirs, and improvements to make this less cumbersome or costly would further facilitate growth of the field.

Each of these mentioned assays have some advantages and limitations as indicated in the table below (Table 1.1).

Table 1.1: Advantages and limitations of selected assays aimed at analysing the HIV reservoir.

Assay	Advantages	Limitations
Quantitative viral outgrowth assays (QVOA)	Isolate inducible, replication competent proviruses	Very expensive, high workload, unable to detect all replication competent proviruses
Inducible virus assay	Faster, less expensive and more sensitive than QVOA	Unable to differentiate inducible from replication comptent proviruses
Gene specific quantification	Very fast and sensitive	Only specific to proviruses consisting of intact regions of the targeted gene
Intact and defective provirus identification and quantification	Fast, sensitive, able to distinguish between intact and defective proviruses	Only specific to proviruses consisting of the targeted regions, based on sequences from a determined sample subset
Single copy sequencing (*gag*-Pol)	Sensitive, able to assess proviral divestity	Limited to proviruses with *gag*-Pol, does not account for variation in other genome regions
Single copy sequencing (near full length)	Sensitive, detect intact proviruses	Requires intact 3’ and 5’ proviral regions, will overlook severely deleted proviruses
Integration site assays	Able to investigate the role of clonal proliferation	Yields no proviral sequence data
Combined near full length and integration site assay	Able to assess clonal proliferation of intact proviruses	Expensive, high workload, only focussed on intact or mostly intact proviruses

1.8 HIV cure

As cART has been shown to have great benefit, especially in early stages of infection, and could prevent onward transmission, all HIV infected individuals currently qualify for therapy. This results in a substantial burden on health care systems as patients require life-long treatment and frequent visits to facilities to monitor therapeutic success (UNAIDS 2020). Moreover, non-adherent individuals may acquire drug resistant HIV strains, compromising the durability of treatments. Novel approaches that would either eliminate HIV from the body, i.e. an eradication cure, or achieve sustained remission in the absence of therapy, i.e. a functional cure, may therefore have huge benefit for affected individuals and the healthcare systems (Ananworanich & Fauci 2015). Although two patients have successfully been cured of HIV, both patients endured severe cancer treatments, including allogenic stem cell transplantations from CCR5Δ32 individuals (Hütter et al. 2009; Gupta et al. 2019). This treatment is not considered as a viable and scalable therapeutic strategy.

A few different approaches to HIV cure are currently being investigated. The "shock/kick and kill" approach is aimed at inducing viral expression from the latent reservoir with latency reversing agents, which can then be targeted and cleared by cytotoxic T lymphocytes (Deeks 2012). However, current latency reversing agents do not efficiently induce the entire reservoir, and viral proteins from defective proviruses are also targeted by cytotoxic T lymphocytes resulting in inadequate clearance of the reservoir (Pollack et al. 2017; Abner & Jordan 2019). In contrast, the novel "block and lock" strategy aims to reinforce latency with small interfering RNAs to inhibit viral rebound after treatment termination (Suzuki et al. 2008; Ahlenstiel et al. 2015; Mousseau et al. 2015; Méndez et al. 2018). However, this strategy is still in development and requires further investigation. Gene editing methods such as CRISPR-Cas9 and zinc-finger nucleases could be used to target human or viral genes and thereby induce resistance to HIV, bolster viral latency or deactivate proviruses (Ebina et al. 2013; Wang et al. 2014; Liao et al. 2015; Xu et al. 2017; Dash et al. 2019; Xiao, Guo & Chen 2019). Although some of these approaches yielded promising preclinical results, significant advances are required to achieve scalable cures. Therapeutic HIV vaccines aim to eliminate rebounding virus rather than targeting the elusive proviral reservoir. Various approaches for an HIV vaccine have been studied, including specific viral protein targets (gag and Tat) (Schooley et al. 2010; Pollard et al. 2014; Deng et al. 2015; Sgadari et al. 2019), multiple viral protein targets (Lévy et al. 2005; Lévy et al. 2006) and autologous dendritic cells to stimulate T cell responses (García et al. 2011; Gandhi et al. 2016; Gay et al. 2018). Some of these vaccines have demonstrated moderate

success in delaying viral rebound (Lévy et al. 2006) and continuous reductions in the proviral reservoir (Sgadari et al. 2019). Even though these findings suggest that therapeutic vaccines may improve outcomes after treatment interruption, sustained viral remission has not been achieved (Davenport et al. 2019).

1.9 Study motivation

Recent studies have indicated the significance of intact and defective proviral clones in the barrier to cure. Although several assays can successfully characterise proviruses, severely deleted proviruses such as solo-LTRs are not detectable by these assays. Recent data indicates that solo-LTR proviral clones may account for a significant proportion of the proviral landscape. Additional data on severely deleted proviruses are required to investigate the potential impact they may have on the barrier to HIV cure. An assay capable of amplifying any provirus regardless of intactness is required to characterise even highly deleted proviruses. This could be achieved by targeting the unique integration site and a region in the host genome specific to a selected integration site for amplification, and so excluding the requirement of proviral intactness. Additionally, integration-site specific quantification would allow the study of any particular clone and could assist in determining the true proportions of the proviral landscape.

The recent characterisation of clonal viraemia as a possible cause of low level and transient viraemia requires further investigation. Replication competent proviral clones are possibly not only the primary source of viral rebound after treatment interruption but could also be responsible for sustained low-level and transient viraemia detected in some patients. Therefore, the investigation of these cases of unusual viraemia is essential to understanding the dynamics and kinetics of proviral clones. Additionally, cases of low-level and transient viraemia require in depth investigation to determine the implications on patient care, and to provide a framework for the recognition and management of such events.

1.10 Aim

The aim of this doctoral project was to investigate the HIV proviral landscape by means of studying specific integration sites and clones in HIV positive patients on long term cART.

1.11 Objectives

1. Characterise the HIV proviral landscape by amplifying proviruses with integration-site specific PCR reactions, followed by full-length sequencing. (Chapter 2)
2. Identify persisting HIV infected cell clones longitudinally and investigate clonal proliferation with quantitative or semi-quantitative integration-site specific PCR. (Chapter 3)
3. Characterise non-suppressible viraemia and investigate whether proviral clones are the source of viraemia, in patients on long term cART. (Chapters 4 and 5)

Chapter 2

Integration site specific proviral characterisation

2.1 Introduction

Characterising the reservoir in patients receiving long-term cART is an essential step in understanding HIV persistence and working towards a cure. Several molecular techniques target various regions of HIV DNA in order to study the integrated proviruses and through this, define the reservoir. Although these methods are valuable in studying the kinetics of HIV reservoirs, it does not explain the mechanism of persistence. Recently, clonal proliferation of HIV infected cells, rather than ongoing replication, has emerged as an important mechanism of long-term HIV persistence, assays are therefore needed to characterise HIV proviral clones. Specifically, there was a need for assays that could investigate particular proviruses, identified by their unique integration sites in order to describe HIV persistence. Multiple molecular assays have been developed by various groups to identify HIV proviral integration sites and better understand the latent reservoir (Vandegraaff et al. 2001; Yu et al. 2008; Ciuffi & Barr 2011; Maldarelli et al. 2014; Sunshine et al. 2016). Although these assays can be used to identify integration sites, they offer little to no molecular sequence data of the provirus itself. Additional techniques, capable of identifying integration sites and proviral sequences, were developed to address this gap in the field (Einkauf et al. 2019; Patro et al. 2019). The primary goal of these methods is to identify intact HIV proviruses along with defective proviruses comprising at least partial HIV genomes. All the above-mentioned assays enable one to characterise the proviral landscape, in order to determine the proportion of the reservoir compared to defective proviruses. They have an important limitation however, relying on HIV specific primers to amplify the HIV provirus. It is known from studies using these assays that a large proportion of proviruses have large internal deletions. Therefore, if any large deletions occur in any of these primer binding areas, the proviral assays would be unable to amplify the particular provirus. These assays are therefore inherently biased towards somewhat more intact proviruses (which would at least allow binding of the internal HIV proviral primers). Proviruses consisting of only HIV LTR, or “solo LTR” integration events would therefore not

be detected, and subsequently overlooked in the proviral landscape proportions. Although such defective proviruses do not form part of the true or replication competent reservoir, LTRs contain promoters and may impact on adjacent gene transcription and cell survival and clonal proliferation of infected cells (Maldarelli et al.2014; Cesana et al. 2017; Anderson & Maldarelli 2018a). The high prevalence of solo LTRs among retroviral elements in the mammalian germline may suggest that solo LTRs may be a common outcome of retroviral integration (Katzourakis et al. 2007; Contreras-Galindo et al. 2013a). We therefore developed a different approach aimed at amplifying proviruses irrespective of intactness. This approach makes use of identified integration sites and designing primers that would amplify proviruses by targeting the human genome and unique human genome-proviral junctions, using a semi-nested PCR approach. This approach will not require any HIV genome region as target and as a result will be able to amplify even severely deleted proviruses.

2.2 Materials and methods

2.2.1 Specimen selection

Participants in this study were from the children with HIV early antiretroviral (CHER) study and post-CHER cohorts. The CHER study was conducted in HIV infected children in South Africa, diagnosed between 4-6 weeks of life and compared clinical and immunological outcomes between a) early elective treatment initiation continued for 40 or 96 weeks followed by a period of treatment interruption until meeting criteria for treatment initiation and b) delayed (treatment started when meeting clinical or immunological criteria) but not interrupted (Violari et al. 2008; Cotton et al. 2013). The post-CHER cohort is a subset of participants retained from the CHER trial to study neurocognitive outcomes and HIV-1 reservoirs. Longitudinal whole blood EDTA specimens were collected twice annually as part of the CHER and post-CHER studies. As part of post-CHER, HIV proviral integration sites were identified at the HIV Dynamics and Replication Program, National Cancer Institute (NCI) (Frederick, MD, USA) with integration site analysis (ISA), in order to investigate reservoirs in children after receiving early cART (Bale et al. manuscript in preparation). Integration sites were successfully identified in 11 study participants, and further investigated for clonality by identifying the most expanded proviral clones (Table 2.1). These 11 participnts were selected using a sampling for diversity approach to represent children with variable ages of treatment initiation and variable intra-participant viral diversity, including particpants with monotypic

and highly diverse populations. For this study, 30 of the most expanded clones from the 11 patients of which ISA data was available, were selected for full-length proviral analysis by a novel integration site specific assay. These 30 most expanded integration sites, as detected by ISA, accounted for 10 of the 11 study participants with ISA data.

Table 2.1: Integration site data generated for the 11 patients. The total number of integration sites detected by the ISA assay is indicated along with the number of clones identified. The number of integrase cell associated DNA (iCAD) copies per 1 million PBMCs, an indication of the total proviral load, is included.

PID	Age at treatment initiation	Time to viral suppression	Viral status over study period	iCAD value	Number of integration sites detected	Number of clones detected
338206	9.9 months	0.92 years	Suppressed	21.3	311	9
340116	9.3 months	2.29 years	Suppressed	181.5	301	6
332406	2.7 months	1.38 years	Suppressed	23.6	272	21
339266	9.0 months	0.44 years	Suppressed	46.7	110	16
334696	6.1 months	0.47 years	Suppressed	9.2	94	3
335106	1.8 months	0.46 years	One viraemic episode	4.5	128	3
337286	17.7 months	1.37 years	Suppressed	32.8	101	2
337916	1.9 months	0.46 years	Suppressed	1.5	136	14
335836	5.1 months	0.93 years	Suppressed	11.9	454	30
341862	2.2 months	0.44 years	One viraemic episode	42.3	116	10
360806	2 months	3.76 years	One viraemic episode	186.2	133	5

2.2.2 Peripheral blood mononuclear cell isolation

Peripheral blood mononuclear cells (PBMCs) were isolated from freshly collected whole blood EDTA specimens from selected CHER patients according to the hanc cross-network PBMC processing protocol (www.hanc.info/labs/labresources/procedures/Pages/pbmcSop.aspx). Approximately 34 mL whole blood was collected per patient at each visit. Plasma was separated by centrifugation in a Jouan BR4i centrifuge (Jouan SA, St-Herblain, France) at 1623 rcf for 10 min and stored at -80^{o}C in 1 mL aliquots. The volume of plasma removed was replaced with 1x phosphate buffered saline (PBS) (Lonza, Basel, Switzerland) and mixed until homogeneous. The blood mixture was carefully layered on 15 mL histopaque-1077 (Sigma-

Aldrich, MO, USA) as density gradient medium (Figure 2.1) and centrifuged 721 rcf for 30 min with centrifuge acceleration and break set to the lowest setting.

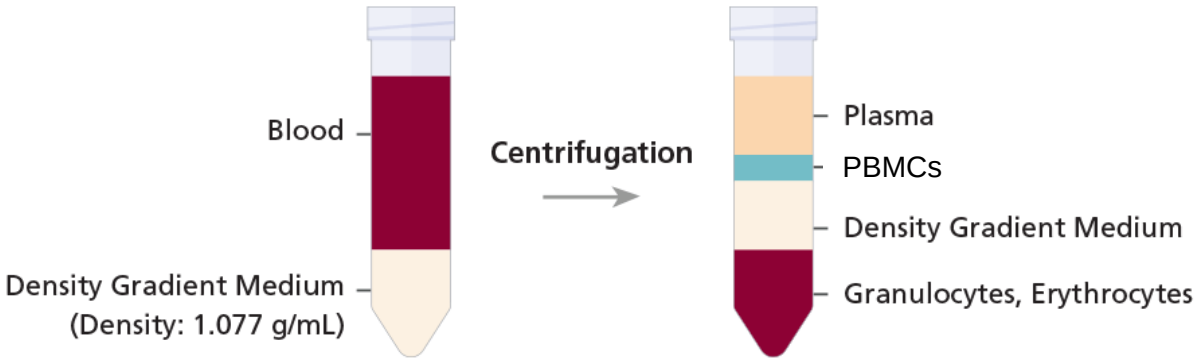


Figure 2.1: Diagram of whole blood separation into plasma, PBMCs, and red blood cells.

The PBMC layer (Figure 2.1) was carefully removed and washed twice in 45 mL 1x PBS with centrifugation at 462 rcf for 17 min. Purified PBMCs were resuspended in 20 mL 1x PBS. A 1:1 mixture of PBMCs (50 µl) and Trypan Blue Dye (50 µl) (Bio-Rad, Hercules, CA, USA) was used to determine cell counts with a TC20 automated cell counter (Bio-Rad) setting cell size gates to 6-17 µm. PBMCs were pelleted by centrifugation at 462 rcf for 10 min and PBS was removed, followed by suspension in 10 % dimethyl sulfoxide (DMSO) (Sigma-Aldrich) and 90 % heat inactivated foetal bovine serum (FBS) (Gibco, Waltham, MA, USA) to a concentration of 2.5x10E6 cells per mL. Purified PBMCs were stored in aliquots of 1 mL in a liquid nitrogen storage unit.

2.2.3 Nucleic acid isolation

Selected PBMC specimens were split to consist of ~1.25 million cells per isolation reaction. PBMCs were centrifuged in a Labnet Prism R refrigerated microcentrifuge (Labnet, Edison, NJ, USA) at 500 rcf for 15 min at 4°C. After removal of the supernatant, 600 µl custom lysis buffer consisting of 3.6 M Guanidinium thiocyanate (GuSCN), 0.13 M Dithiothreitol (DTT), 0.67 mg/mL Glycogen, 10 mg/mL N-Lauroylsarcosine, 10 mg/mL Sodium Citrate and 132 µl molecular grade water were added. Cell pellets were resuspended by pulse vortex mixing for 4 seconds and incubated at 25°C for 10 min. Reactions were inverted at least 30 times after the addition of 600 µl 100% Isopropanol and centrifuged at 17 200 rcf for 20 min at 10°C. The

supernatant was carefully removed without disturbing the nucleic acid pellet followed by the addition of 1mL ice cold 70% ethanol. Reactions were inverted and centrifuged at 17 200 rcf for 20 min at 10°C. The ethanol wash step was repeated twice more followed by the complete removal of all ethanol by centrifuging reactions twice at 17 200 rcf for 1 min and removing any supernatant. The nucleic acid pellets were dried for about 15 min at 25°C followed by resuspension in 200 µl cold 5 mM Tris-HCl (pH 8.0) (Sigma-Aldrich). Isolated nucleic acid was stored at -80°C until use.

2.2.4 Assay design and validation

2.2.4.1 Amplification strategy

We developed a **novel assay** to enable the amplification of specific proviruses from their known integration sites. A combined semi-nested PCR strategy was employed to amplify each provirus with human genome sequence at both 5' and 3' ends (Figure 2.2). The integration site specific proviral amplification (ISSPA) method consists of two **parallel** pre-nested reactions, reaction 1 primes in the human genome beyond the 3' end of the provirus, with the opposing primer being the HIV 5' U3 LTR human genome junction, while reaction 2 primes in the human genome beyond the 5' end of the provirus with the opposing primer being the HIV 3' U5 LTR human genome junction. Each of these PCR reactions are followed by 2 semi-nested reactions, using the same junction primers as in the pre-nested reactions but opposing internal human genome primers or alternatively junction primers: reaction 1.1 and 1.2, and 2.1 and 2.2 respectively (Figure 2.2). These four different **parallel** approaches to amplify proviruses by targeting unique integration site junctions and human genome positions were done as **confirmation** that the products generated represent the true amplification of proviruses and do not represent PCR artefacts.

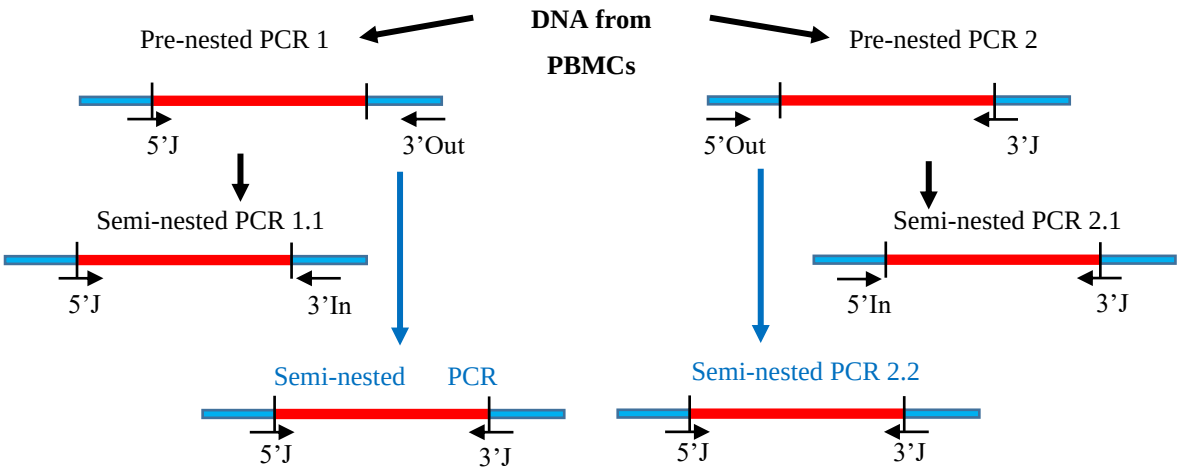


Figure 2.2: Representation of the integration site-specific amplification strategy with human genome in blue and provirus in red. Primer position and orientation are indicated by arrows.

Based on the the relative low total HIV-1 DNA levels and expected proportion that could be made up by individual clones according to literature, only a small fraction of the total DNA isolated from PBMCs is expected to contain any specific integration site. Therefore, in order to increase the probality of detecting these rare integrants ~60 000 PBMC equivalent DNA copies were tested for each integrant. To limit PCR inhibition by DNA overload each PCR step was performed in replicates of 11, with ~250 ng DNA per reaction. No template controls (NTCs) were included.

2.2.4.2 Primer design

Specific primers to amplify proviruses from each unique integration site were designed (Addendum B). Integration site data were obtained from ISA performed previously (Bale et al. 2019), mapped to human genome 19 (hg19) in BLAT together with LTR nucleotide sequences generated for each patient as part of a previous study. The semi-nested 5'In and 3'In human genome primers (semi-nested PCR 1.1 and 2.1 (Figure 2.2)) were designed at least 100 bp from the selected integration site. Junction primers were designed 6-8 bp into the HIV LTR to ensure integration site specificity (Figure 2.3). Primers were designed with caution to ensure proper

binding to template, palindromic sequences were avoided to prevent hairpin loops and corresponding primers were designed to have similar melting temperatures (Tm).

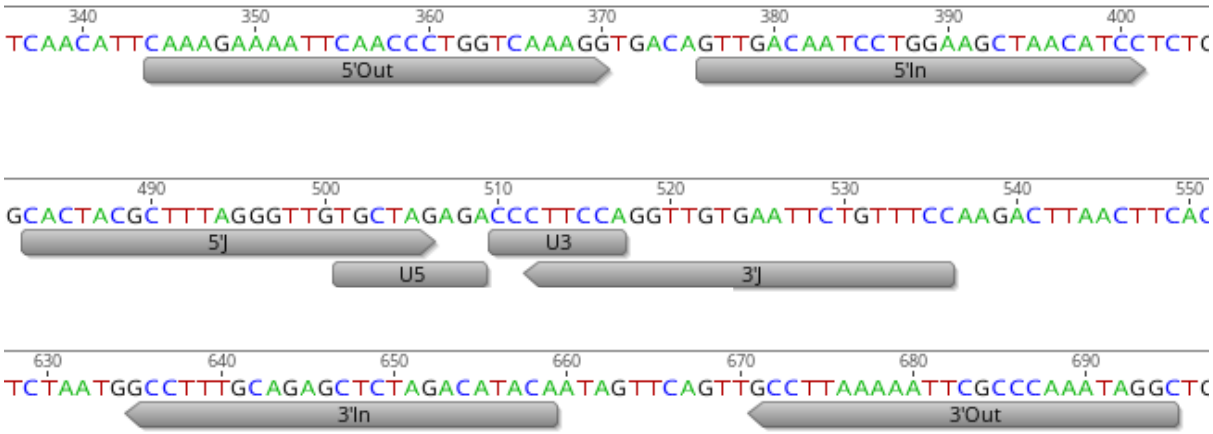


Figure 2.3: Representation of specific primer positioning for a selected integration site. Integration site provirus-human genome junctions with HIV LTR U3 and U5 ends are indicated along with the four human genome primers (5'Out, 5'In, 3'Out and 3'In) and two integration site junction primers (5'J and 3'J) required for ISSPA.

Standard primers were designed to target proviral junctions and the flanking human genome regions. Dual priming oligonucleotide (DPO) primers were designed targeting the same human genome regions as the standard primers in order to compare the efficacy of standard and DPO primers. The DPO primers were designed according to specifications from (Chun et al. 2007a; Chun et al. 2007b). All primers designed for ISSPA are included in Addendum B.

2.2.4.3 Validation with ACH2

Several assessments were performed in order to validate the ISSPA assay. The ACH2 cell line was selected as a control for this assay, due to the presence of one major clone (present in ~ 1/3 genome copies) and multiple minor clones contained in the genome. Standard and DPO primers designed to target the major ACH2 clone were used to attempt full-length (~9 kb) proviral amplification. Assay sensitivity was investigated by limiting dilution of ACH2 input DNA at concentrations of ~547, ~54 and ~5 copies of the major clone per PCR reaction. An assay specificity reaction was included for each of the primer sets, spiked with ~17 000 copies of HIV negative PBMC DNA per reaction. For the validation assay, a positive no dilution control was included containing >50 000 DNA copies of the major ACH2 clone.

A 400 bp region of HIV *integrase* was targeted to confirm full-length proviral (~9 kb) amplification after ISSPA was performed on the ACH2 control. Reaction mixes (25 µl) for the *integrase* PCR consisted of: 1x Ranger mix, 0,4 µM forward and reverse primer respectively and nuclease free water (Promega Corp., Madison, WI, USA). Amplification conditions were as follows: an initial denaturing step (94ºC, 2 min) was followed by 30 cycles of denaturation (94ºC, 30 sec), annealing (44ºC for 30 sec) and extension (68ºC, 30 sec). A final extension step (68ºC, 7 min) was included to ensure full-length amplicons.

2.2.5 Integration site specific proviral amplification (ISSPA)

Proviral amplification was attempted as mentioned above (2.2.4.1) in replicates of 11 reactions. Amplification was performed with Ranger mix (Bioline meridian bioscience, London, UK) and each reaction (50 µl) consisted of: 1x Ranger mix, 0.4 µM forward and reverse primer respectively and nuclease free water (Promega Corp.). Amplification conditions consisted of an initial denaturing step (95ºC, 1 min) followed by 30 cycles of denaturation (98ºC, 15 sec), annealing (first five cycles consisted of touchdown temperature starting at 63ºC down to 58ºC, with the remaining 25 cycles at 58ºC for 45 sec) and extension (68ºC, 11 min). A final extension step (68ºC, 15 min) was included to ensure full-length products. A second round semi-nested PCR reaction was performed at the same amplification conditions using 1 µl of a 1:4 dilution of the first-round product. If no successful proviral amplification was achieved, amplification was attempted on an additional 11 replicates, up to a maximum of 44 replicates.

2.2.6 Verification of proviral PCR products

When only a solo-LTR sequence was obtained after ISSPA, PCR recombination yielding artefactual solo-LTR products was excluded as a possibility by attempting amplification of a fragment that is slightly larger than LTR, by priming on the integration site junction and just beyond LTR. Amplification was attempted with two primer sets, namely: integration site junction – 642Rev and integration site junction – 788Rev, targeting the packaging signal and *gag* respectively (Table 2.2). The appropriate integration site-junction primers for solo-LTR verification were selected from patient ISSPA primers that yielded solo-LTR PCR products. Amplification was attempted with Ranger mix as described in section 2.2.5. Cycle conditions consisted of an initial denaturing step (95ºC, 1 min) followed by 30 cycles of denaturation

(98°C, 15 sec), annealing (first 10 cycles consisted of touchdown temperature starting at 65°C down to 55°C, with the remaining 20 cycles at 55°C for 30 sec) and extension (68°C, 1 min). A final extension step (68°C, 5 min) was included to ensure full-length products.

Table 2.2: Primers targeting the HIV packaging signal and *gag* used to verify solo-LTR amplification.

Primer name	Orientation	Sequence (5'-3')	Location in HXB2 genome
642Rev	Reverse	TTC GCT TTC **DV**G TCC CTG TTC GG	642 – 665
788Rev	Reverse	AGA TGG GTG CGA GAG CGT CA	788 – 808

2.2.7 Half-length integration site specific proviral amplification

If ISSPA failed to amplify any HIV products after several (22-55) replicates, we attempted to retrieve larger, at least partially intact, proviral sequences that could be intergrated at these particular integration sites by performing half-length integration site specific proviral amplification. In short, primers were chosen that targeted the 5'integration site provirus-human junction and *reverse transcriptase* (3410) for the 5' half and in the p6 region (1870) and 3'integration site provirus-human genome junction for the 3' half (Figure 2.4 and Table 2.3).

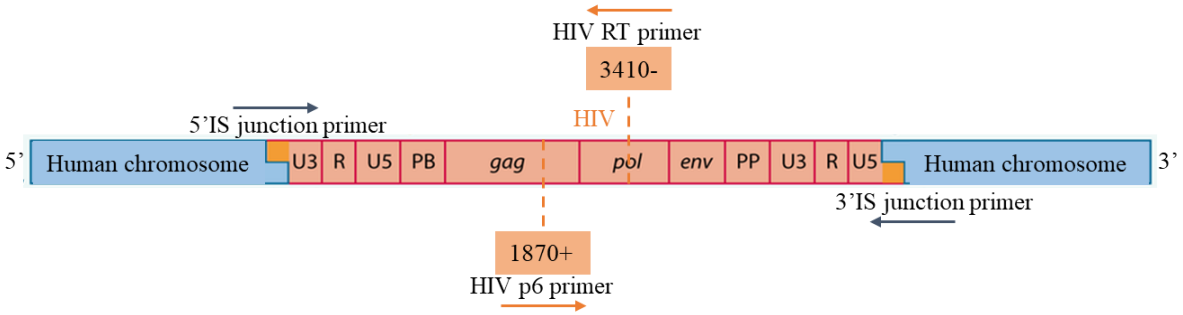


Figure 2.4: Representation of HIV provirus indicating primer locations for half-length integration site specific proviral amplification.

Table 2.3: Primers used for half-length integration site specific proviral amplification

Primer name	Orientation	Sequence (5'-3')	Location in HXB2 genome
5' integration site junction	Forward	Unique to targeted integration site	Integration Site Junction
HIV RT 3410	Reverse	CAG TTA GTC GTA CTA TGT CTG TTA GTG CTT	3409 – 3438
HIV p6 1870	Forward	GAG TGT TGG CTG AGG CAA TGA G	1871 – 1892
3' integration site junction	Reverse	Unique to targeted integration site	Integration Site Junction

Amplification was attempted with Ranger mix as described in section 2.2.5. Cycle conditions consisted of an initial denaturing step (95^{o}C, 1 min) followed by 30 cycles of denaturation (98^{o}C, 15 sec), annealing (55^{o}C for 45 sec) and extension (68^{o}C, 6 min). A final extension step (68^{o}C, 15 min) was included to ensure full-length products.

2.2.8 PCR product analysis and purification

PCR reactions were assessed for amplicon by mixing 5 µl PCR product with 15 µl EZ-Vision® Bluelight DNA Dye (Amresco, Solon, OH, USA) at a dilution of 1:10 000 in a 96 well plate. The UVprochemi II system (Uvitec, Cambridge, UK) was used to view plates in order to identify amplicon positive wells. Positive amplicons were further analysed with agarose gel electrophoresis. Gels consisting of 1% agarose (Lonza, Rockland, ME), 1x Sodium Boric acid buffer (0,4 mg/mL NaOH, 2,25 mg/mL Boric acid and distilled water) and 1x EZ-Vision Bluelight DNA Dye were used for analysis. A 1 kb DNA ladder (Promega Corp., Madison, WI) was used to determine amplicon size. The UVprochemi II system was used for gel analysis.

Selected PCR product bands were gel-separated, excised and purified using the Nucleotrap® kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's instructions and eluted in 30 µl elution-buffer. Purified PCR products were stored at -20^{o}C.

2.2.9 Nucleotide sequencing and analysis

Purified PCR products were used as template for Sanger sequencing. Nucleotide sequencing was performed with the ABI Prism BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) on an automated analyser. Sequencing reactions (20 µl) comprised of: Big Dye Terminator 1x sequencing buffer (Applied Biosystems), 1 µl reaction Terminator mix (Applied Biosystems), 2 pmol primer (Fermentas) and nuclease-free water (Promega Corp.). Thermal cycling conditions for sequence reactions were as follows: an initial denaturing step (95ºC, 3 min) was followed by 25 cycles of denaturation (94ºC, 30 sec), annealing (50ºC, 15 sec) and extension (60ºC, 4 min). Primers that yielded amplicon after ISSPA were used for the respective nucleotide sequencing reactions. Sequencing reactions were purified using the BigDye® XTerminator Purification Kit (Applied Biosystems) as per the manufacturer's instructions. Sequence data was collected with Genetic Analyser Data Collection V3.1 software after analysis on an automated ABI sequence analyser.

Nucleotide sequences were verified and translated using Sequencher™ 4.10.1 (Gene Codes Corp., Ann Arbor, MI, USA), Geneious Prime 2020.1 (Biomatters Ltd., Auckland, NZ) and BioEdit Sequence Alignment Editor (V7.2.5). Nucleotide sequences were identified as HIV by BLASTn (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&BLAST_SPEC=OGP__9606__9558&LINK_LOC=blasttab&LAST_PAGE=blastn). Proviral nucleotide sequences were identified by investigating nucleotide sequence ends for human genome sequence, specific to each integration site, and confirmed by Human BLAT Search (https://genome.ucsc.edu/cgi-bin/hgBlat). Proviral sequences were reconstructed with complimentary sequence data obtained from Sanger sequencing reactions targeting each specific integration site against the closest HIV sequence match in GenBank (AF321523).

2.2.10 Phylogenetic analysis

Proviral sequences were aligned with HIV consensus C using MAFFT Version 7 (https://mafft.cbrc.jp/alignment/server/). P-distance analysis was performed to construct phylogenetic trees in MEGA7 (Kumar, Stecher & Tamura 2016). Bootstrap analysis of 1000 replicates was perfomed and only clusters with >70% support were considered significant.

2.3 Results

2.3.1 Selected specimens

In total, 30 of the most expanded clones previously identified from 10 participants were selected for full-length proviral analysis. Integration site data such as location and orientation are provided in Table 2.4 along with human genome data such as gene of insertion. Orientation of the gene and whether the integration occurred in an intron or exon are also indicated as obtained from Human BLAT Search (genome.ucsc.edu/cgi-bin/hgBlat).

Table 2.4: Integration site data of selected patients with chromosome (Chr) and location indicated, along with integration site orientation and human gene of insertion with gene orientation. It is also indicated whether the integration occurred in the gene intron or exon.

Patient ID	Integration site location	Integration site orientation	Gene/Orientation	Intron/Exon
338206	Chr1_24294547	+	SRSF10/-	UTR3
	Chr19_1423908	+	DAZAP1/+	Intron
	Chr11_73469151	-	RAB6A/-	Intron
	Chr1_231081054	-	TTC13/-	Intron
	Chr1_153861271	-	GATAD2B/-	Intron
340116	Chr17_58684892	-	PPM1D/+	Intron
	Chr6_13641182	-	RANBP9/-	Intron
	Chr18_20529591	+	RBBP8/+	Exon
	Chr2_73620016	-	ALMS1/+	Intron
	Chr18_6467385	+	Nearest: LINC01387	Non-coding
	Chr16_17134060	+	Nearest: XYLT1/-	-
332406	Chr18_18609870	+	ROCK1/-	Intron
	Chr14_68759177	-	RAD51B/+	Intron
	Chr6_32786955	-	-	-
339266	Chr2_241353495	-	Nearest: GPC1/+	-
	Chr9_137214293	+	Nearest: RXRA/+	-
334696	Chr6_135525218	+	MYB/+	Intron
	Chr6_26227488	-	Nearest: HIST1H3E/+	-
335106	Chr6_45580207	-	Nearest: RUNX2/+	-
	Chr4_54281522	+	FIP1L1/+	Intron
337286	Chr17_40421775	-	STAT5B/-	Intron
	Chr19_53455758	+	ZNF816/-	Intron
337916	Chr16_47550997	-	PHKB/+	Intron
	Chr3_141257185	+	RASA2/+	Intron
	Chr8_43028875	-	HGSNAT/+	Exon
335836	Chr10_103786725	-	C10orf76	Intron
	Chr3_49116408	+	QRICH1/-	Intron
341862	Chr17_42550940	-	GPATCH8/-	Intron
	Chr11_116877612	+	SIK3/-	Intron
	Chr19_10287814	-	DNMT1/-	Intron

2.3.2 Assay validation outcomes

Full-length (~9kb) ACH2 proviral amplification attempted with human genome targeted DPO primers and standard primers targeting the integration site junction, was unsuccessful after multiple attempts (Figure 2.5).

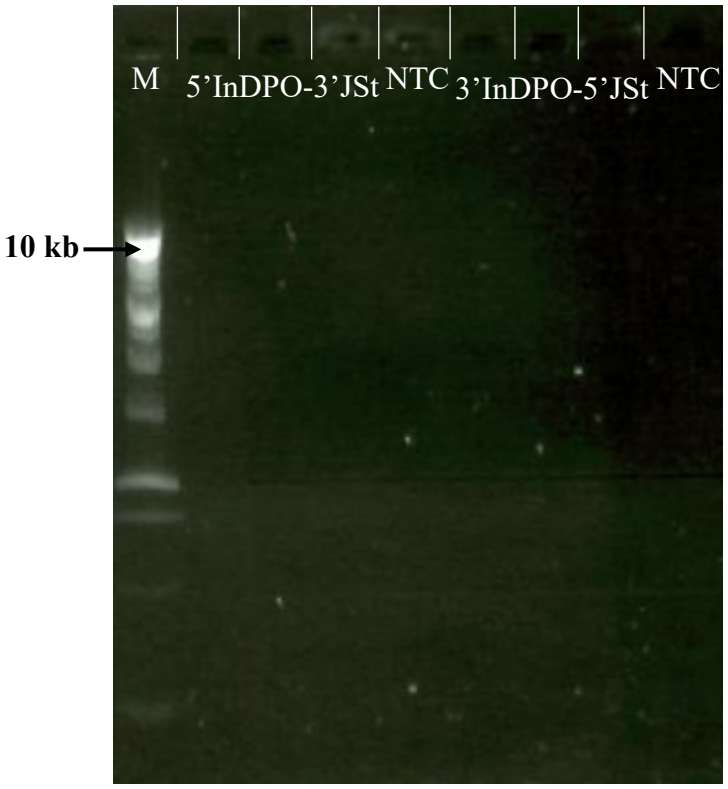


Figure 2.5: Gel electrophoresis of second round semi-nested PCR products after attempting amplification of full-length ACH2 proviral amplification with a combination of standard and DPO primers. Each primer set (5'human genome 'In' DPO primer-3' junction standard primer and 5' junction standard primer-3' human genome 'In' DPO primer) was attempted in triplicate. The 1 kb DNA ladder is indicated by M.

Proviral amplification of the major ACH2 clone was successful after performing ISSPA with standard primers targeting human genome and integration site. Amplification using standard primers with diluted input of ACH2 DNA (~547, ~54 and ~5 major clone proviral copies per reaction) yielded large (~9 kb) bands. No clear amplification bands were observed for HIV negative PBMC background PCR controls, except for a faint band in semi-nested PCR 2.2 (Figure 2.6).

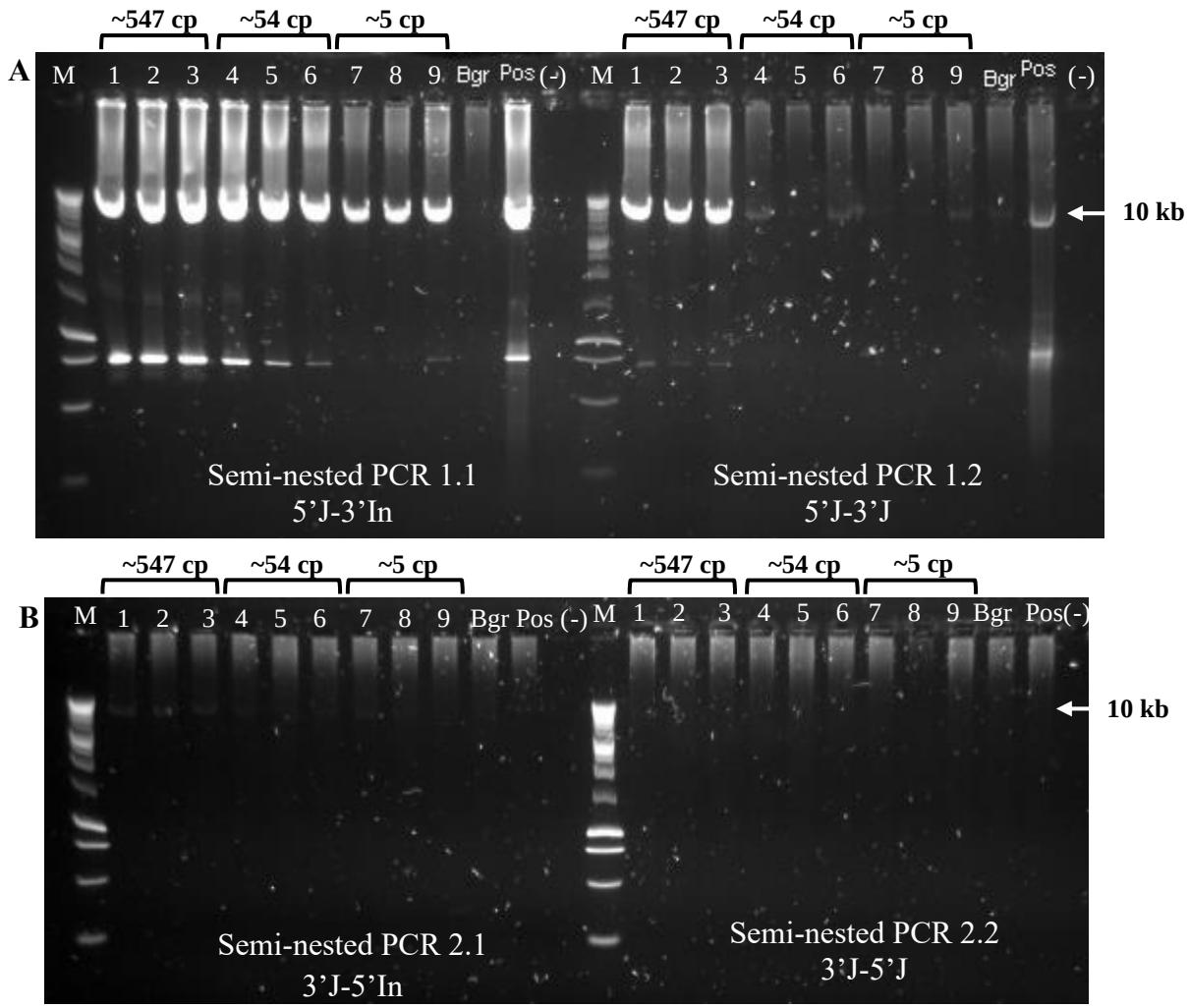


Figure 2.6: Gel electrophoresis of PCR products of A) Semi-nested PCR 1.1 and 1.2 and B) Semi-nested PCR 2.1 and 2.2. Proviral copy input per reaction is indicated for each well. Wells containing background only reactions are labelled with Bgr, positive controls are labelled as Pos and no template control wells are labelled as (-). The 1 kb DNA ladder is indicated by M. Lanes 1-3 (~547 copies (cp), lanes 4-6 (~54 cp) and lanes 7-9 (~5 cp).

Amplification targeting *integrase* performed on selected ACH2 ISSPA products from Figure 2.6 confirmed the presence of HIV (Figure 2.7). *Integrase* amplicons were observed in the most concentrated wells (well 1) of all ISSPA sets and in the least concentrated wells (well 9) of semi-nested PCR 2.1 and 2.2 (Figure 2.7).

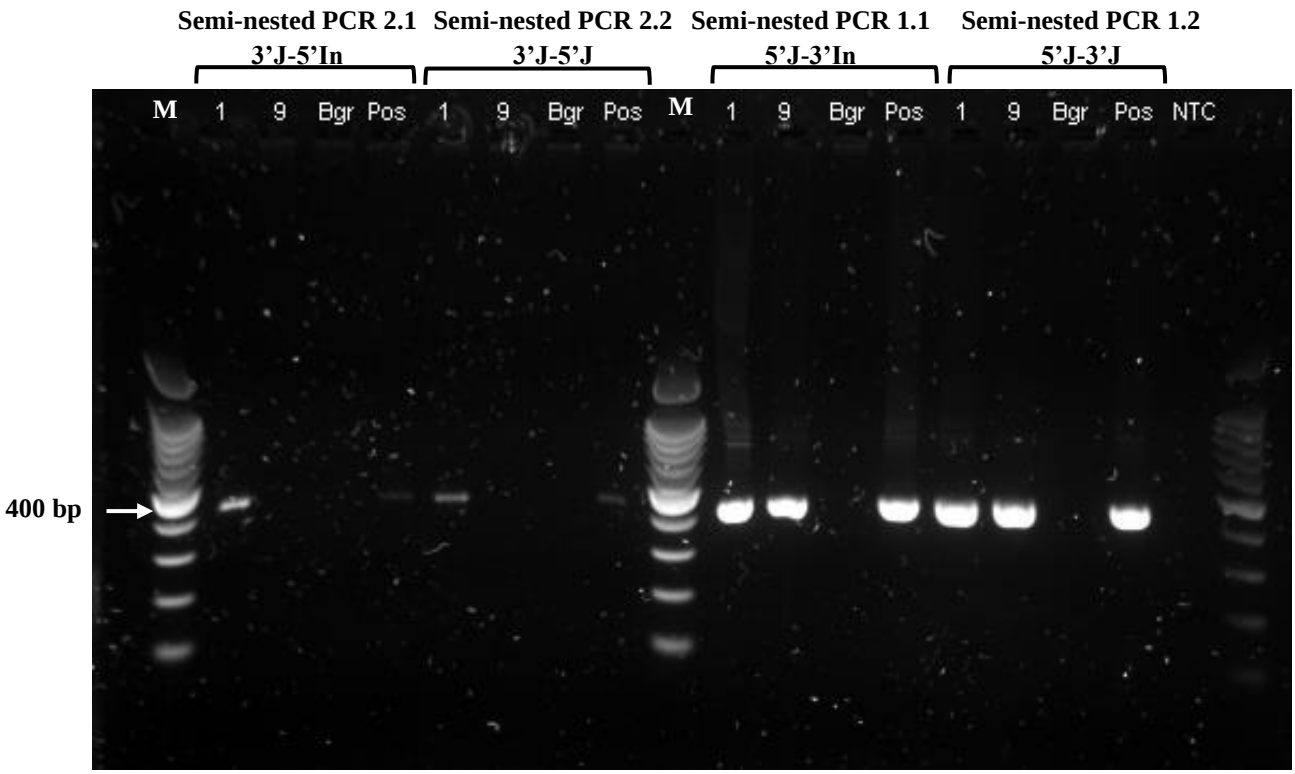


Figure 2.7: Gel electrophoresis of integrase PCR performed on wells 1 and 9 of ISSPA 1.1, 1.2, 2.1, 2.2 Bgr and Pos products. The 100 bp DNA ladder is indicated by M, along with a 400 bp marker.

These results indicate that standard primers are superior to DPO primers for the purpose of ISSPA. Standard primers can be used to successfully amplify ~10 kb HIV amplicons from low copy number ACH2 DNA.

2.3.3 Proviral amplification and nucleotide sequencing of selected proviral clones

Positive amplification was observed by fluorescence plate view (Figure 2.8A) after ISSPA, and PCR products were selected for nucleotide sequencing following gel electrophoresis (Figure 2.8B). HIV proviral sequences (red arrows) were identified from semi-nested ISSPA reactions along with human genome background amplification (white arrows) (Figure 2.8).

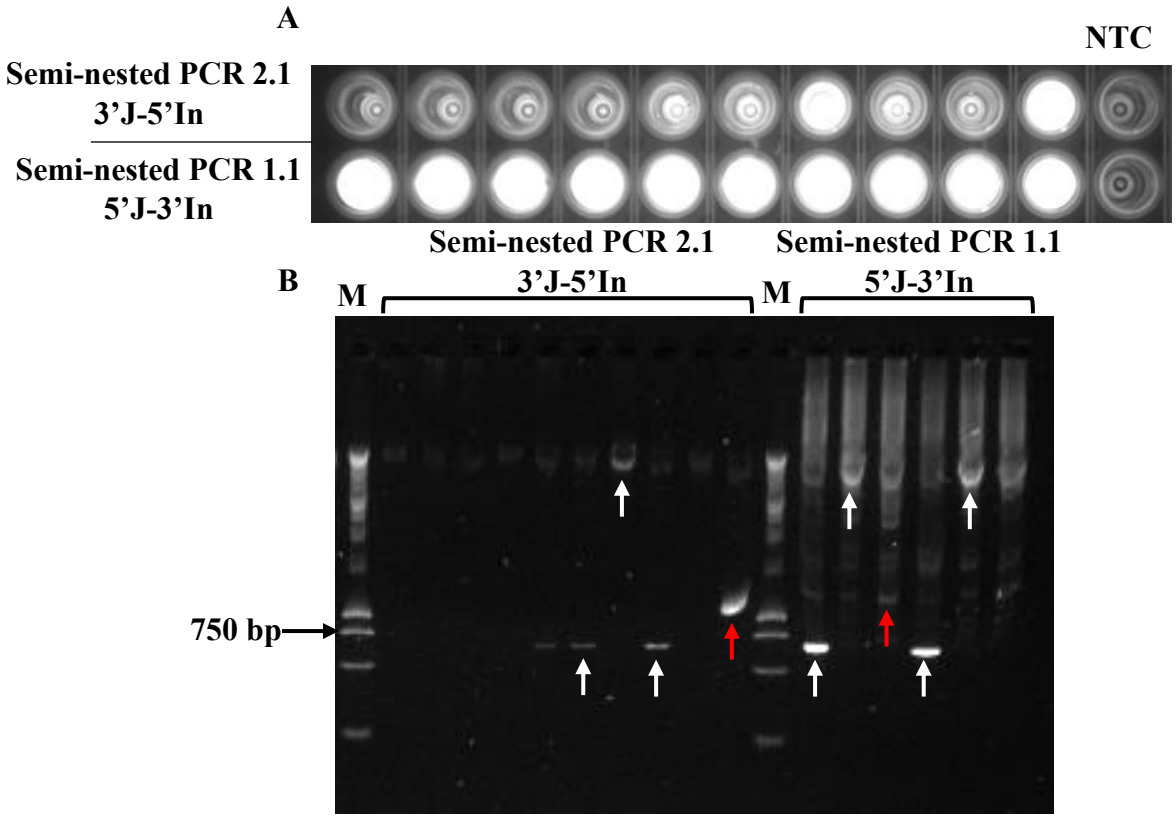


Figure 2.8: A) Amplification fluorescence plate of replicates of ISSPA semi-nested reactions 1.1 and 2.1 indicating wells of positive and negative amplification along with no template controls. B) Gel electrophoresis of ISSPA products viewed in A, with arrows indicating bands selected for nucleotide sequencing. The only HIV positive sequence products are indicated by red arrows.

Amplicons, identified as HIV by nucleotide sequencing, were obtained from 9/30 integration sites after ISSPA (Table 2.4). HIV proviruses were successfully amplified for 6/30 integration sites from the 5'- (semi-nested 1.1 and 1.2) and 3'-proviral junctions (semi-nested 2.1 and 2.2). The provirus of integration site (332406: chromosome (Chr)14_68759177) could only be amplified from the 3'-proviral junction (2.1 and 2.2), whereas provirus (335106: Chr6_45580207) could only be amplified from the 5'-proviral junction (1.1 and 1.2) (Table 2.5). Amplification of provirus (332406: Chr6_32786955) was only successful after semi nested reaction 1.2, yielding only proviral nucleotide sequence without any human genome-provirus junction signal (Table 2.5). No direct link between clone size and the ability to successfully amplify the provirus was observed.

Table 2.5: Summary of amplification results for each integration site and ISSPA semi-nested reaction as indicated by Figure 2.2. The patient ID, integration site, chromosome (Chr) and location is indicated. Each proviral clone size is indicated as a percentage of the total integration sites detected for the patient, along with the time difference between samples used for ISA and ISSPA.

Patient ID	Integration site	Clone size	Time between ISA and ISSPA	Semi-nested PCR reaction	Reactions performed per PCR set	Successful amplification
338206	Chr1_ 24294547	5.8%	5.5 Months	1.1/1.2	11	✓
				2.1/2.2	22	✓
	Chr19_ 1423908	2.3%		1.1/1.2	55	×
				2.1/2.2	55	×
	Chr11_ 73469151	1.9%		1.1/1.2	22	✓
				2.1/2.2	22	✓
	Chr1_ 231081054	1.3%		1.1/1.2	22	✓
				2.1/2.2	22	✓
	Chr1_ 153861271	0.6%		1.1/1.2	22	×
				2.1/2.2	22	×
340116	Chr17_ 58684892	2%	5.5 Months	1.1/1.2	44	×
				2.1/2.2	44	×
	Chr6_ 13641182	1.3%		1.1/1.2	33	✓
				2.1/2.2	33	✓
	Chr18_ 20529591	1.3%		1.1/1.2	33	×
				2.1/2.2	33	×
	Chr2_ 73620016	1%		1.1/1.2	22	✓
				2.1/2.2	22	✓
	Chr18_ 6467385	0.7%		1.1/1.2	22	×
				2.1/2.2	22	×
	Chr16_ 17134060	0.3%		1.1/1.2	44	×
				2.1/2.2	44	×
332406	Chr18_ 18609870	5.1%	5.5 Months	1.1/1.2	33	×
				2.1/2.2	33	×
	Chr14_ 68759177	4.4%		1.1/1.2	55	×
				2.1/2.2	11	✓
	Chr6_ 32786955	2.2%		1.2	22	✓
				2.1/2.2	22	×
339266	Chr2_ 241353495	10.9	Same visit	1.1/1.2	33	×
				2.1/2.2	33	×
	Chr9_ 137214293	7.3%		1.1/1.2	33	×
				2.1/2.2	33	×

Table 2.5 Continued

Patient ID	Integration site	Clone size	Time between ISA and ISSPA	Semi-nested PCR reaction	Reactions performed per PCR set	Successful amplification
334696	Chr6_ 135525218	7.4%	6 Months	1.1/1.2	22	×
				2.1/2.2	22	×
	Chr6_ 26227488	6.4%		1.1/1.2	22	✓
				2.1/2.2	22	✓
335106	Chr6_ 45580207	12.5%	5.5 Months	1.1/1.2	22	✓
				2.1/2.2	33	×
	Chr4_ 54281522	1.6%		1.1/1.2	22	×
				2.1/2.2	22	×
337286	Chr17_ 40421775	7.9%	6 Months	1.1/1.2	22	×
				2.1/2.2	22	×
	Chr19_ 53455758	2%		1.1/1.2	22	×
				2.1/2.2	22	×
337916	Chr16_ 47550997	2.9%	Same visit	1.1/1.2	22	×
				2.1/2.2	22	×
	Chr3_ 141257185	2.2%		1.1/1.2	22	×
				2.1/2.2	22	×
	Chr8_ 43028875	2.2%		1.1/1.2	22	×
				2.1/2.2	22	×
335836	Chr10_ 103786725	1.5%	Same visit	1.1/1.2	22	×
				2.1/2.2	22	×
	Chr3_ 49116408	1.5%		1.1/1.2	22	×
				2.1/2.2	22	×
341862	Chr17_ 42550940	3.4%	5.5 Months	1.1/1.2	22	×
				2.1/2.2	22	×
	Chr11_ 116877612	2.6%		1.1/1.2	22	×
				2.1/2.2	22	×
	Chr19_ 10287814	2.6%		1.1/1.2	22	×
				2.1/2.2	22	×

2.3.4 Solo LTR verification

Attempts to amplify any region beyond a single LTR, by priming on the integration site junction and in the HIV packaging signal and *gag* yielded no positive amplification (Figure 2.9), confirming solo LTR proviral clones detected by ISSPA.

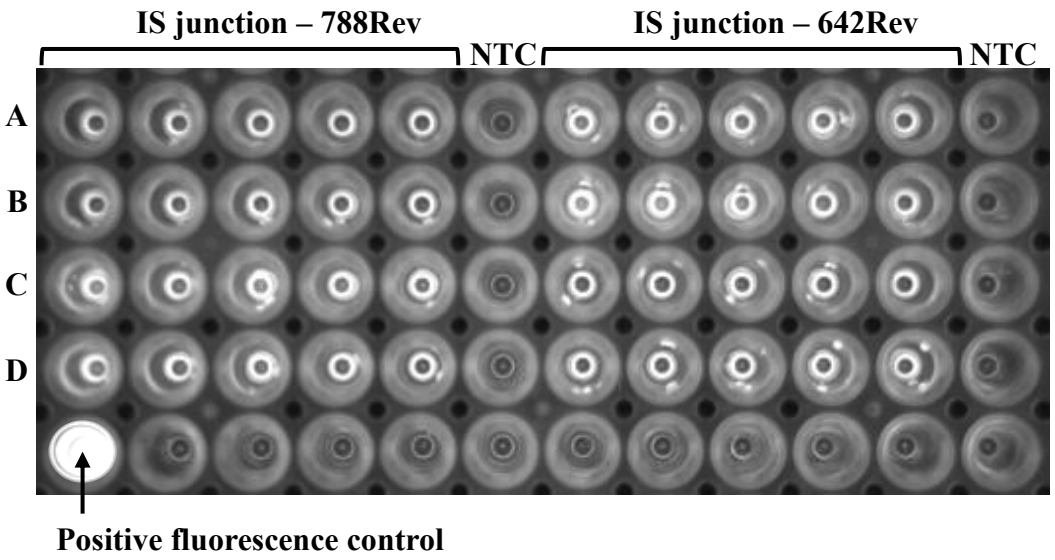


Figure 2.9: Example of an amplification fluorescence plate of solo-LTR verification reactions. Five PCR replicates performed for each primer combination integration site junction - 788Rev/642Rev) along with no template controls are indicated. Columns A, B and C indicate reactions targeting clones Chr1_24294547, Chr11_73469151 and Chr1_231081054 of patient 338206 respectively. Column D indicate reactions for clone Chr2_73620016 of patient 340116. A positive flourescence control well is indicated.

2.3.5 Half-length integration site specific proviral amplification

No amplification was observed after near half-length amplification was attempted on ISSPA negative products (Figure 2.10). This confirms that these expanded clones are not intact, or at least partially intact proviruses.

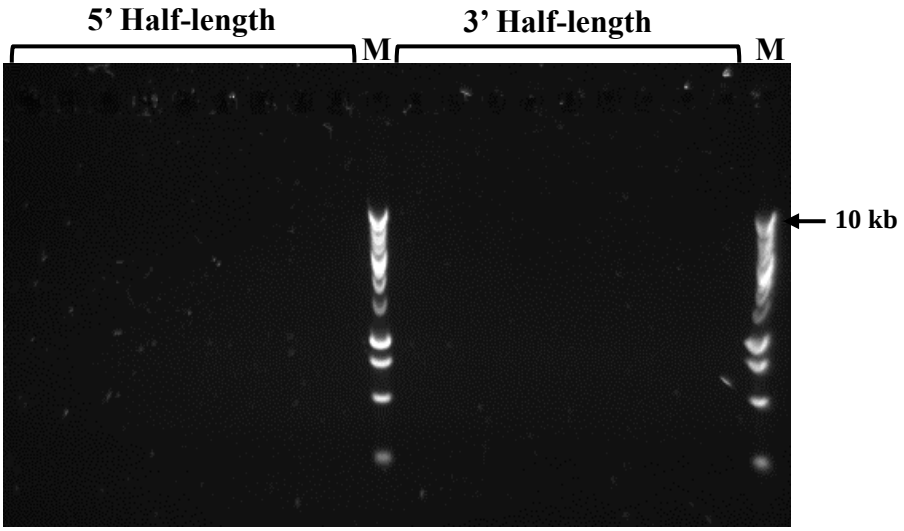


Figure 2.10: Gel electrophoresis of half-length PCR products.

2.3.6 Nucleotide analysis and proviral reconstruction

Nine proviral sequences, generated from ISSPA, were successfully reconstructed as solo-LTRs with flanking integration site junction and human genome sequences mapped to the exact genomic locations, identified by preliminary ISA data (Figure 2.11). Human genome data could only be generated for one end (either 3' or 5') of two proviral clones (332406: Chr14_68759177 and 335106: Chr6_45580207) (Figure 2.11), due to amplification success reported in section 2.3.3. Although no human genome sequence data could be obtained for proviral clone 332406: Chr6_32786955, HIV proviral sequence was obtained and reconstructed from ISSPA targeting the specific integration site (Figure 2.11). Proviral reconstruction of clone 334696: Chr6_26227488 yielded an incomplete solo-LTR, with a 240 bp deletion spanning the R and U3 regions of LTR (Figure 2.12).

Figure 2.11: Representation of reconstructed proviral nucleotide sequences showing the 3' (A) and 5' (B) junctions and human genome data. Clone integration sites are indicated along with the respective human chromosome sequences. Red bars indicate areas where ISSPA and nucleotide sequencing were unsuccessful, as reported in section 2.3.3. The closest related sequence from GenBank (AF321523) was used as reference sequence.

Figure 2.12: Representation of proviral reconstruction indicating a 240 bp deletion spanning R and U3 of the LTR of clone Chr6_26227488 (PID: 334696).

2.3.7 Phylogenetic analysis

Phylogenetic analyses indicate solo-LTR HIV proviral sequences of each patient cluster together with inter patient genetic variation (Figure 2.13). Similar solo-LTR proviral sequences were generated with ISSPA, targeting different locations in the human genome corresponding to proviral integration sites. This confirms amplicons were generated from the correct patients and further suggest that these sequences are not from artefactual PCR products.

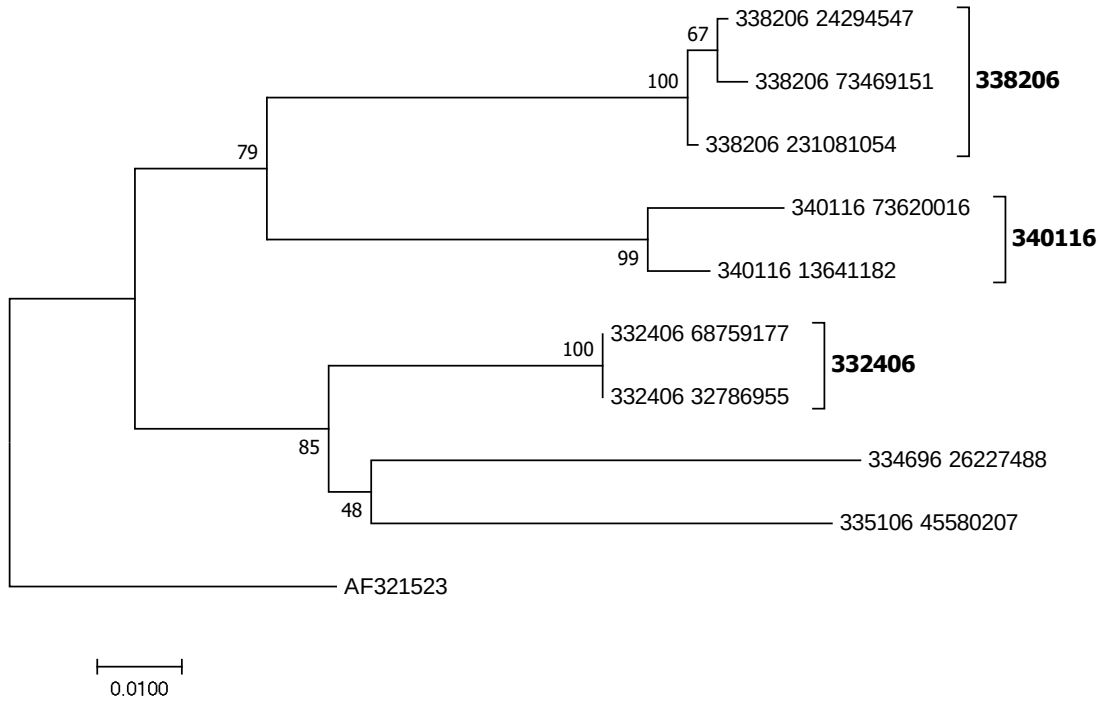


Figure 2.13: P-distance phylogenetic tree of HIV solo-LTR proviral nucleotide sequences (634 bp) obtained from ISSPA. Replicates of 1000 bootstraps were used with bootsrap percentage values indicated next to the brances. The closest HIV subtype C sequence match from GenBank was used as reference and root.

2.4 Discussion

Although some recently developed assays (Einkauf et al. 2019; Patro et al. 2019) can characterise HIV proviral clones and identify their human genome integration sites, these assays are unable to detect severely deleted proviruses linked to integration sites due to HIV genome targeting. Therefore, an assay capable of amplifying HIV proviruses irrespective of intactness and linking these proviruses to the human genome integration site can significantly assist in characterising the proviral landscape. This data could shed further light on the proportion of intact proviruses in addition to comparing the regions of proviral integration in relation to intactness. Moreover, highly defective integrated HIV-1 may not be without consequence as the promoters in LTR may impact cell survival and proliferation when activating growth gene transcription.

The novel ISSPA assay was designed with the aim of amplifying expanded proviral clones by targeting the unique human genome-proviral junctions of integration sites. Proviral amplification by ISSPA uses the adjacent 3' human genome sequence with the 5' integration site junction and the 5' human genome sequence with the 3' integration site junction in parallel. Additionally, amplification targeting both the 3' and 5' integration site junctions could be used to confirm PCR products from these parallel reactions.

The ACH2 cell line was selected for the validation of the ISSPA assay due to the close resemblance to HIV positive human PBMC samples, containing multiple proviral clones of various intactness. The major intact proviral clone of the ACH2 cell line was selected as the target of the ISSPA assay. Standard and DPO human genome primers were also compared in the validation, investigating the increased specificity of DPO primers in order to minimise human genome background amplification. The strict requirements for DPO primer design allowed for human genome targeted primers, but specific integration site junction sequences could not accommodate DPO primers. All integration site junction primers were therefore designed as standard primers initiating in the human genome, covering the five bases of the integration site and spanning six to eight bases into the HIV LTR. Intact proviral amplification performed with all standard primers yielded ~9 kb products, whereas amplification with human DPO and standard junction primers did not yield any products after multiple attempts. Standard primers were therefore considered superior in long-range proviral amplification and selected for ISSPA. Assay sensitivity, assessed by diluted input of ACH2 genetic material (~547, ~54 and ~5 major clone proviral copies) along with HIV negative DNA background (~17 000

copies) yielded ~9 kb amplicons. The ~9 kb amplicons were confirmed as HIV by *integrase* amplification. This confirmed that proviral amplification by targeting integration site junctions and adjacent human genome sequences is possible, even at low template input. A degree of background amplification was noted, therefore, only products that could be verified by junction-junction amplification would be considered for nucleotide sequencing.

Due to the expected rarity of specific integration sites, each ISSPA approach (3' and 5') was performed in replicates of 11 for each of the 30 selected integration sites. If no HIV was detected after nucleotide sequencing of several selected amplicons of various sizes, additional replicates of each ISSPA approach was performed, up to 22-55 replicates. Due to proviral intactness being unknown, amplicons of various sizes from ~600 bp to ~10 kb were selected for nucleotide sequencing. HIV sequences were identified from ISSPA amplicons of 9/30 integration sites tested, in each case yielding PCR products of ~750 bp. Nucleotide sequencing revealed the amplicons comprised of HIV LTR and human genome sequence. Proviral reconstruction of 5/30 clones (**338206**: Chr1_24294547, Chr11_73469151, Chr1_231081054; **340116**: Chr6_13641182, Chr2_73620016) yielded a single full-length HIV LTR flanked by human genome sequence at both 3' and 5' ends. Proviral clone **334696**: Chr6_26227488 reconstruction yielded an incomplete HIV LTR sequence, with a 240 bp mid LTR deletion spanning the R and U3 regions. This proviral sequence was also flanked by human genome sequence at both 3' and 5' ends. This suggests that highly deleted HIV genomes, consisting of only a partial LTR sequence, could integrate into the human genome. Proviral reconstruction of clones **335106**: Chr6_45580207 and **332406**: Chr14_68759177 yielded nearly complete HIV LTR sequences with some missing LTR sequence data and no human genome sequence at the 3' and 5' ends respectively, due to ISSPA only being successful from one end. Human genome sequences were obtained from the opposite ends (5' and 3' respectively) of these clones, that were successfully amplified and sequenced. Reconstruction of clone **332406**: Chr6_32786955 yielded a nearly intact HIV LTR sequence with only a few missing bases at the 5' end, unfortunately no human genome sequence could be generated in this case. All human genome sequences flanking HIV proviruses were successfully mapped to the exact targeted chromosome and location, confirming specific proviral and human genome junction amplification by ISSPA. These results indicate that 8/9 successfully amplified and sequenced proviral clones comprise of only a solo-LTR and 1/9 proviral clone consists of a partial solo-LTR sequence.

The unexpected nature of only identifying solo-LTR proviral clones prompted the need for result verification. The solo-LTR products were unlikely to be caused by PCR recombination as identical products were obtained from independent pre-nested and semi-nested PCR sets. Additionally, solo-LTR proviruses were confirmed by an inability to amplify any HIV signal beyond the LTR. Furthermore, phylogenetic analyses indicate solo-LTR proviral sequences of each patient cluster together and are distinctly different from solo-LTR sequences from other patients, eliminating the possibility of cross contamination. Taken together, these findings offer strong support for the existence and prevalence of solo-LTR proviral clones, and that these clones can be successfully characterised by ISSPA.

Proviruses that could not be amplified by ISSPA were further investigated with near half-length integration site specific proviral amplification. These clones were investigated for intact or partially intact proviruses possibly missed by ISSPA. Unfortunately, no additional proviral clones could be amplified using this method. Successful proviral amplification is therefore assumed to be dependent on sampling, prevalence of the clone and genomic location of the clone could complicate primer design. Additionally, successful amplification could be biased towards shorter proviruses.

The ISSPA assay allows for the investigation of specific proviral clones from identified integration sites. The ability of ISSPA to amplify proviruses, irrespective of intactness, offers additional data about the proviral landscape, notably it enables one to investigate the presence of solo-LTR proviruses in multiple patients. A limitation of the ISSPA assay is that due to the rarity of a particular clone it requires a large blood volume to obtain sufficient DNA; and to prevent inhibition by DNA overload of a reaction, multiple PCR replicates are performed. Due to the large sample volume required, ISSPA might need to be performed on a different time point than ISA, this could prohibit proviral detection due to potential waning of the clone of interest. Additionally, ISSPA may be less efficient for the amplification of larger intact proviruses in patient samples, even though long-range amplification was successful for the ACH2 control. The most expanded clones were selected for proviral characterisation. Cells harbouring proviruses that express mRNA and protein are more likely to be purged by the immune system (Pollack et al. 2017). This biased survival of deleted proviruses may also account for the overrepresentation of solo-LTRs amongst the most expanded CD4 cell clones in patients on long-term suppressive ART.

Since LTRs are duplicated at the ends of integrated HIV-1 proviruses, assays to amplify full-length proviruses have been designed to prevent ambiguous binding in LRT, and therefore would not amplify solo-LTRs (Bruner et al. 2016; Hiener et al. 2017). Therefore, until recently, HIV-1 solo-LTR proviruses have largely been overlooked, but a recent report characterised a predominant solo-LTR HIV cell clone (Anderson et al. 2020). This study utilised a mehod similar to ISSPA by targeting the unique integration site junction together with the specific region in human genome. This method is different from ISSPA, as ISSPA is performed as two independent first round amplification reactions followed by two sets of semi-nested reactions each, in several replicates. The solo-LTR described by Anderson et al was also verified by attempting to amplify from junction to *gag*, with no success. In contrast, the existence and prevalence of solo-LTR proviruses are well described in other retroviruses (Katzourakis et al. 2007), including human endogenous retroviruses (HERVs) (Buzdin et al. 2006; Contreras-Galindo et al. 2013b; Thomas, Perron & Feschotte 2018). In fact, solo-LTRs comprise as much as 90% of HERVs in the human genome (Subramanian et al. 2011). Taken together, these studies along with findings from this study, indicate that solo-LTR HIV proviruses could account for the majority of HIV proviruses in patients on long term cART, as some of the most expanded proviral clones were characterised as solo-LTRs. At this stage, the effect of HIV-1 LTR transcriptional promoters (Klaver & Berkhout 1994; Pereira et al. 2000) on host genes remains unclear in the case of a solo-LTR provirus.

In this study, a novel method capable of proviral and adjoining host nucleotide amplification by targeting the specific clone integration site was developed and termed ISSPA. Nine of the most expanded proviral clones in five patients were successfully characterised with ISSPA, and eight of these could be linked to the unique location in the human genome. The 8/9 proviruses characterised as solo-LTRs and 1/9 as partial-LTR indicate that a significant proportion of the proviral landscape is comprised of solo or partial-LTR proviruses. It would be advantageous for future analyses of the proviral landscape to include solo and partial-LTR proviruses, as this could significantly impact the proposed proportions of other proviruses. Intact proviruses may therefore represent less than 2% (Figure 1.9) of the total number of proviruses in patients on long term cART. Further studies on the quantification of specific proviral clones, including solo-LTR clones, could offer valuable data concerning clonal expansion and contraction. The ISSPA assay could be used for the investigation of important clones in adult and children, by determining the complete sequence of any selected proviral clone, as long as the integration site is known.

Chapter 3

Integration site specific proviral quantification

3.1 Introduction

Currently there are numerous assays targeting various regions of the provirus used to quantify the proviral reservoir, to determine the number of intact proviruses and so determine the barrier to cure (Yu et al. 2008; Brady et al. 2013; Cillo et al. 2014; Sherman et al. 2017; Anderson & Maldarelli 2018b; Bruner et al. 2019). These assays are capable of accurately quantifying different subsets of the proviral landscape based on the targeted proviral region, however, they are unable to quantify severely deleted proviruses such as solo-LTR and partial-LTR proviruses, potentially overlooking a significant part of the proviral landscape. Additionally, these solo-LTR proviruses may act as promoters and may impact transcription of grow genes and clonal proliferation. Determining the true proviral landscape proportions would shed light on the actual size of the barrier to cure, potentially a fraction of the total proviral load in some individuals. This will need to be taken into account when investigating the proviral reservoir, as the intact proviruses may be very rare compared to defective proviruses. Additionally, these assays cannot selectively quantify specific clones. We therefore developed an integration site specific quantification assay capable of accurately quantifying proviral clones, including solo-LTR clones. This approach consists of a nested PCR approach, targeting the unique integration site junction for highly specific quantification. A high quantity of cellular DNA is used as template in order to detect rare proviral clones. Quantification of specific proviral clones may elucidate the true dynamics of HIV-1 infected cell clonal proliferation and could assess the impact of future interventions which are aimed at limiting clonal proliferation.

3.2 Materials and methods

3.2.1 Specimen selection

Specimen selection for this study was based on successful proviral characterisation by ISSPA (Chapter 2). Nine HIV-1 proviral clones successfully characterised by ISSPA were selected as candidates for integration site specific proviral quantification (Table 3.1).

Table 3.1: Proviral clones successfully characterised by ISSPA, eligible for integration site specific proviral quantification.

Patient ID	Integration site location in human genome	Provirus orientation	Provirus
338206	Chr1_24294547	+	Solo-LTR
	Chr11_73469151	-	Solo-LTR
	Chr1_231081054	-	Solo-LTR
340116	Chr6_13641182	-	Solo-LTR
	Chr2_73620016	-	Solo-LTR
332406	Chr14_68759177	-	Solo-LTR
	Chr6_32786955	-	Solo-LTR
334696	Chr6_26227488	-	Partial-LTR
335106	Chr6_45580207	-	Solo-LTR

3.2.2 PBMC and nucleic acid isolation

PBMC and nucleic acid isolation were performed exactly as described in chapter 2 sections 2.2.2 and 2.2.3. Isolated nucleic acid was stored in 5 mM Tris-HCl (pH 8.0) (Sigma-Aldrich) at -80^{o}C until use.

3.2.3 Assay design and validation

3.2.3.1 Amplification and quantification strategy

A nested PCR approach was selected for sensitive and specific quantification of rare HIV proviruses. Increased sensitivity and specificity for the integration site specific proviral absolute quantification (ISSAQ) assay is achieved by including a pre-amplification step before *real-time* (*rt*) quantification (Figure 3.1), similar to assays described by (Bosman et al. 2015). With the expected integration site specific proviral load being less than one copy per reaction, a traditional per reaction *rt* quantification approach would simply not be sensitive enough. Due

to the rarity of individual HIV proviruses, requiring multiple ISSPA reactions to detect a specific provirus, a "digitalised" 92 reaction Poisson quantification approach was devised. With this approach, sample PBMC DNA is seeded in 92 reactions at a concentration of ~200 ng per reaction, equating to "endpoint dilution". Poisson distribution statistics were used to determine the expected number of specific proviral templates per positive reaction. The total amount of input DNA (for 92 reactions) were determined by CCR5 quantification in order to accurately quantify each proviral clone (section 3.2.5).

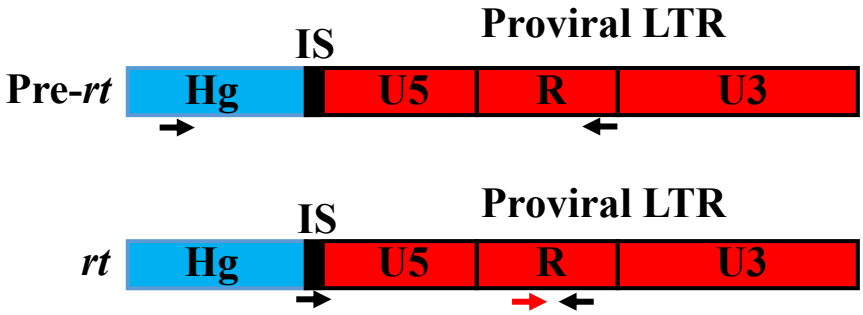


Figure 3.1: Representation of the integration site (IS) specific proviral quantification strategy, consisting of a pre-*real-time* (*rt*) amplification and *rt* quantification step. The section of human genome (Hg) is indicated along with the IS and proviral LTR. Primers are indicated by black arrows and the probe is represented by a red arrow.

3.2.3.2 Primer and probe design

Specific ISSPA inner human genome and integration site junction primers that successfully generated proviral products in chapter 2 were selected for the quantification of the respective clones. Reverse primers, targeting conserved sequences in the R region of the LTR were designed in order to bind all selected proviral clones (Figure 3.2 and Table 3.2). Primer and probe orientations are dependent on the orientation of the provirus.

Figure 3.2: Representation of specific primer and probe positioning for integration site (IS) specific proviral absolute quantification (ISSAQ). The 5' IS junction and HIV LTR U5, R and U3 regions are indicated. The IS specific 5' human genome (Hg) In and 5' IS junction primers along with ISSAQ primers (real time (rt) and pre- rt) and probe are annotated on the nucleotide sequence.

A TaqMan minor groove binding (MGB) probe design (Thermo Fisher Scientific, MA, USA) was selected for highly sensitive and specific binding. A well-conserved section in the R region of the HIV LTR was selected for the probe, which was designed to be hydrolysed by the unique integration site junction primer in order to further increase specificity (Figure 3.2 and Table 3.2). Proviral orientation was considered for the LTR primers and probe design. With these ISSAQ primers and probe, the 3' junction of a positive sense provirus and 5' junction of a negative sense provirus will be targeted.

Table 3.2: Primers and probe targeting the proviral LTR for integration site specific proviral absolute quantification (ISSAQ). Human genome (Hg) and integration site (IS) specific primers are dependant on the targeted clone, these primers are described in Addendum B.

Primer/Probe name	**Orientation**	**Sequence (5'-3')**	**Location in HXB2 genome**
Hg specific (pre- *rt*)	Forward	Dependant on IS	Upstream of IS
IS specific (*rt*)	Forward	Unique to targeted IS	IS Junction
ISSAQ pre- *rt*	Reverse	GCT GCA TAT AAG CAG CTG C	421 – 439
ISSAQ *rt*	Reverse	**W**GG **Y**AG **R**CC AGA TCT GAG CC	464 – 483
ISSAQ Probe	Forward	**FAM**-ACT CAA GGC AAG CTT	531 – 545

3.2.3.3 Assay validation

Several procedures were used to optimise and validate the ISSAQ assay. The major proviral clone of the ACH2 cell line was targeted as an initial proof of concept for the RT detection of a specific proviral clone without interference of other proviruses. Along with integration site

specific primers used for ISSPA validation in chapter 2, ACH2 specific LTR primers were designed for ISSAQ based *rt* detection (Table 3.3). The ISSAQ probe (Table 3.2) was also used for ACH2 *rt* detection.

Table 3.3: Primers targeting the major proviral clone of ACH2 for integration site specific real-time (rt) detection.

Primer name	**Orientation**	**Sequence (5'-3')**	**HXB2 location**
ACH2 In	Forward	GTG ATT TAG TAT ACC TGG AG	61 – 81 bp upstream of IS
ACH2 IS Junction	Forward	AAT TAT TAT CTT TAT TAC TGC TAG	ACH2 IS Junction
ACH2 ISSAQ pre- *rt*	Reverse	CGA GCC CTC AGA TGC TGC	408 – 425
ACH2 ISSAQ *rt*	Reverse	GTT AGA CCA GAT CTG AGC CTG G	465 – 486

Targeted ISSAQ reactions were also tested on the corresponding ISSPA products generated from patient DNA in chapter 2. The ISSPA products were serially diluted until the cycle threshold (Ct) value of the integration site specific quantification curve was ~30 and used as positive controls for ISSAQ. The ACH2 cell line, with multiple integration sites at various frequencies (Symons et al. 2017), was selected as specificity control. Assay specificity was assessed by using ~17 000 ACH2 DNA copies per reaction as template in patient integration site specific reactions. Finally, accurate quantification was assessed by performing ISSAQ on linearly diluted input DNA of patient 338206, targeting proviral clone Chr1_24294547. For the validation of each ISSAQ run, pre-amplification and quantification plates included an integration site specific positive control (diluted ISSPA amplicon), a specificity control (ACH2 DNA), and two NTCs.

3.2.4 Integration site specific proviral absolute quantification (ISSAQ)

The nested PCR approach mentioned above (section 3.2.3.1) was used for proviral quantification. The pre-amplification step was performed with Ranger mix (Bioline meridian bioscience) with each reaction (20 µl) consisting of: 1x Ranger mix, 0,4 µM forward and reverse primer respectively and nuclease free water (Promega Corp.). Pre-amplification conditions consisted of an initial denaturing step (95^{o}C, 1 min) followed by 15 cycles of

denaturation (98ºC, 15 sec), annealing (55ºC for 15 sec) and extension (68ºC, 30 sec). A final extension step (68ºC, 5 min) was included to ensure full-length amplicon. The quantification step was performed with LightCycler® 480 Probes Master (Roche, Basel, Switzerland) and each reaction (25 µl) consisted of: 1x LightCycler® mix, 0,4 µM forward and reverse primer respectively, 0,3 µM probe, nuclease free water (Promega Corp.) and 5 µl pre-amplicon. Quantification conditions consisted of an initial denature step (95ºC, 5 min) followed by 50 cycles of denaturation (95ºC, 15 sec), annealing (55ºC for 15 sec), extension (68ºC, 30 sec) and fluorescence measurement for FAM dye. Amplification curves were visually assessed, with a sigmoidal shape indicative of exponential amplification as criterion for positivity, to determine the number of positive reactions per 92 ISSAQ reactions.

3.2.5 CCR5 quantification as reference

The number of human genome cell equivalent DNA copies of each patient sample used for ISSAQ was determined in order to normalise the proviral quantification results. Cell equivalent DNA copies were determined by CCR5 quantification according to the single-copy sensitive assay described by (Malnati et al. 2008). Primers and probe as designed by Malnati et al were used for CCR5 quantification (Table 3.4).

Table 3.4: Primers and probe used for CCR5 quantification as described by (Malnati et al. 2008).

Primer name	Sequence (5'-3')
CCR5 Forward	ATG ATT CCT GGG AGA GAC GC
CCR5 Reverse	AGC CAG GAC GGT CAC CTT
CCR5 Probe	**VIC**-AAC ACA GCC ACC ACC CAA GTG ATC A

The CCR5 quantification assay was performed with LightCycler® 480 Probes Master (Roche, Basel, Switzerland) and each reaction (25 µl) consisted of: 1x LightCycler® mix, 0,9 µM forward and reverse primer respectively, 0,3 µM probe, nuclease free water (Promega Corp.) and 10 µl patient PBMC DNA. Quantification conditions consisted of an initial denature step (95ºC, 5 min) followed by 50 cycles of denaturation (95ºC, 15 sec), annealing and extension (60ºC, 1 min) and fluorescence measurement for VIC dye.

3.2.6 Proviral quantification analysis

Using Poisson probability distribution, the number of positive reactions per 92 ISSAQ reactions was analysed to calculate the probability that more than one template could be present per reaction. Nonparametric bootstrapping was used to determine the 95% confidence interval (CI), implemented in R. The calculated range of positive templates per 92 reactions, together with the total amount of cell equivalent input DNA as calculated by CCR5 quantification, was used to normalize the integration site specific proviral load to "integration sites per million cells".

3.3 Results

3.3.1 Selected specimens

Three sample time points, based on sample availability, for the nine proviral clones of five patients were selected for ISSAQ. The description of each proviral clone along with the time on treatment (in years) for each patient sample time point is provided in Table 3.5.

Table 3.5: Samples selected for proviral quantification by ISSAQ, with three time points selected for each patient. The number of years on treatment at sample time point is indicated for each patient.

Patient ID	IS location in human genome	Provirus orientation	Provirus	Years on treatment at sample time point		
				Time point 1	Time point 2	Time point 3
338206	Chr1_24294547	+	Solo-LTR	8.1	8.6	9.6
	Chr11_73469151	-	Solo-LTR			
	Chr1_231081054	-	Solo-LTR			
340116	Chr6_13641182	-	Solo-LTR	7.8	8.3	9.3
	Chr2_73620016	-	Solo-LTR			
332406	Chr14_68759177	-	Solo-LTR	9.2	9.7	10.7
	Chr6_32786955	-	Solo-LTR			
334696	Chr6_26227488	-	Partial-LTR	8.5	9.7	10.2
335106	Chr6_45580207	-	Solo-LTR	8.8	9.3	10.5

3.3.2 Assay validation outcomes

Successful amplification was observed after integration site specific *rt* detection targeting the major clone of ACH2 and patient clone ISSPA products (Figure 3.3). Amplification was observed for all patient clones, except clone Chr6_26227488, a partial-LTR proviral clone lacking the LTR R region targeted by ISSAQ. Assay specificity was confirmed by no non-specific detection of ACH2 in patient integration site targeted reactions.

A dilution effect was observed after "digitalised" quantification by ISSAQ, targeting clone Chr1_24294547, was performed on serially diluted DNA of patient 338206 (Figure 3.4). The error plot indicates the Poisson calculated values and high and low CI values at each DNA concentration.

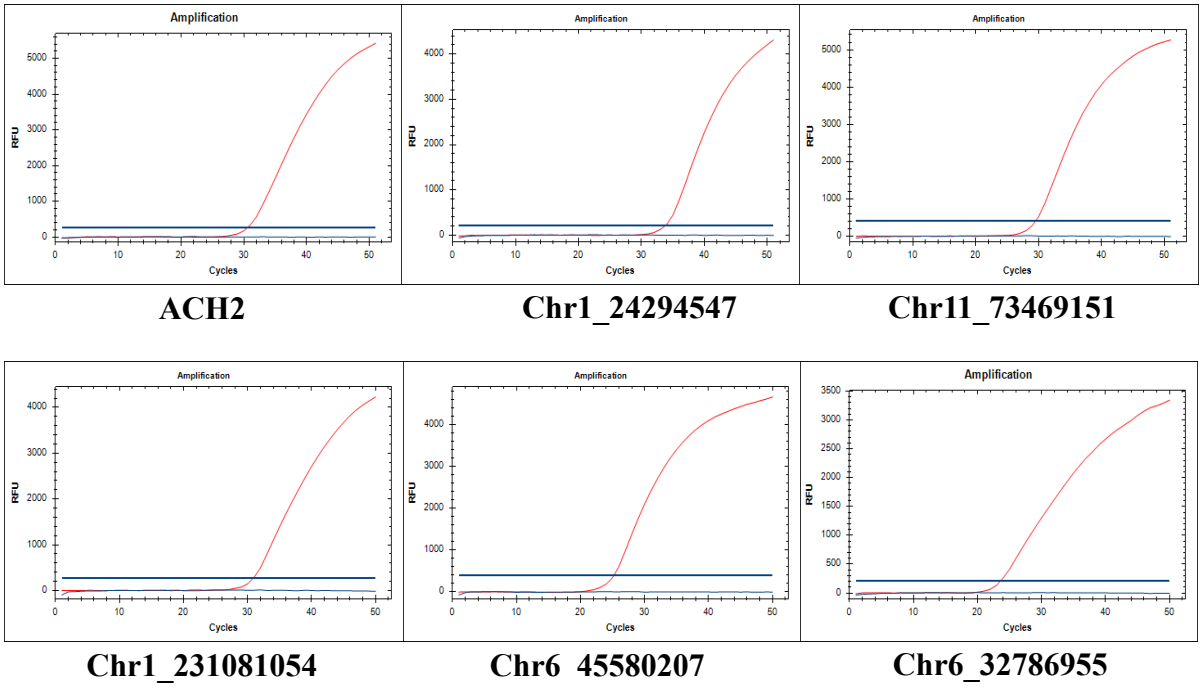


Figure 3.3: Integration site specific amplification curves of ACH2 major clone and selected diluted ISSPA products, to be used as positive controls, from selected patient proviral clones. Fluorescence is indicated on the y-axis and cycles on the x-axis. Target curves are indicated in red, NTC curves are in light blue and threshold lines are indicated in dark blue.

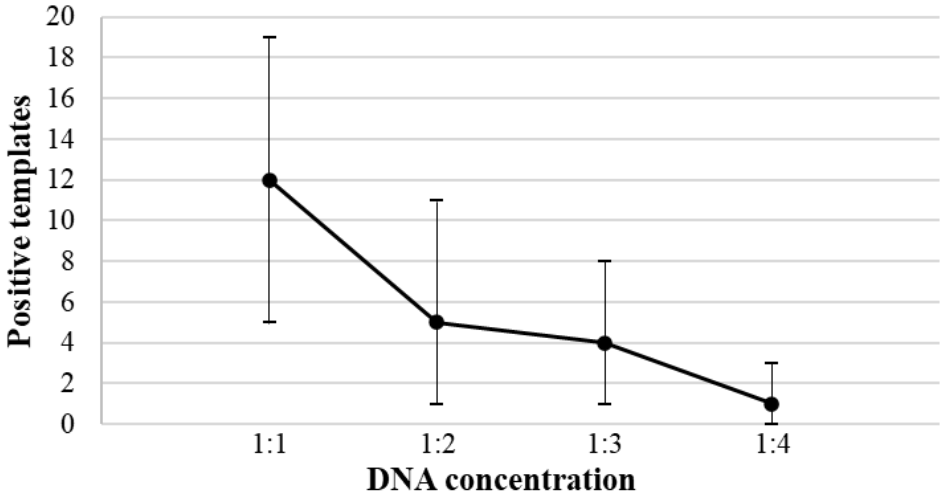


Figure 3.4: Error plot of positive templates at each DNA concentration, indicating the linear dilution effect observed after ISSAQ was performed on patient 338206 clone Chr1_24294547. The expected number of positive templates are indicated by black dots, with the 95% confidence interval values indicated by horizontal lines.

3.3.3 CCR5 quantification

Total cell equivalent DNA copies tested by ISSAQ for each patient clone was calculated from CCR5 quantification (Table 3.6). Patient 334696 clone Chr6_26227488 was excluded from CCR5 quantification and ISSAQ due to the deletion spanning the R region of LTR (discussed in chapter 2), required as ISSAQ target.

Table 3.6: Total cell equivalent DNA copies used for ISSAQ per patient clone time point, as calculated from CCR5 quantification.

Patient ID	IS location in human genome	Cell equivalent DNA copies tested per ISSAQ run at each sample time point		
		Time point 1	Time point 2	Time point 3
338206	Chr1_24294547	742817	504107	315276
	Chr11_73469151			
	Chr1_231081054			
340116	Chr6_13641182	559653	651674	389308
	Chr2_73620016			
332406	Chr14_68759177	237875	363107	516136
	Chr6_32786955			
334696	Chr6_26227488	Excluded due to LTR deletion		
335106	Chr6_45580207	542094	421816	497119

3.3.4 Proviral quantification

Integration site specific proviral quantification data for at least one sample time point could be generated for 6/9 selected proviral clones. In addition to the exclusion of clone Chr6_26227488 due to partial LTR deletion, no quantification data could be generated for patient 332406 clones Chr6_32786955 and Chr14_68759177 at any sample time point. The ISSAQ *rt* amplification curves were carefully studied to determine the number of positive reactions per 92 reaction run (Table 3.7). Amplification curves are displayed in Addendum C.

Patient ID	IS location in human genome	ISSAQ positive reactions			Range of templates calculated by Poisson and bootstrapping									Proviral copies per 1 million cell equivalent DNA copies								
		Time point 1	Time point 2	Time point 3	Time point 1			Time point 2			Time point 3			Time point 1			Time point 2			Time point 3		
					Low	Exp	High	Low	Exp	High	Low	Exp	High	Low	Exp	High	Low	Exp	High	Low	Exp	High
338206	Chr1_24294547	15	15	35	8	16	25	8	16	25	31	44	60	11	**22**	34	16	**32**	50	98	**140**	190
	Chr11_73469151	8	2	5	3	8	15	0	2	5	1	5	11	4	**11**	20	0	**4**	10	3	**16**	35
	Chr1_231081054	6	1	1	2	6	12	0	1	3	0	1	3	3	**8**	16	0	**2**	6	0	**3**	10
340116	Chr6_13641182	6	2	1	2	6	12	0	2	5	0	1	3	4	**11**	21	0	**3**	8	0	**3**	8
	Chr2_73620016	0	3	0	0	0	0	0	3	7	0	0	0	-	**-**	-	0	**5**	11	-	**-**	-
332406	Chr14_68759177	No quantification data obtained																				
	Chr6_32786955	No quantification data obtained																				
335106	Chr6_45580207	8	7	1	3	8	15	2	7	14	0	1	3	6	**15**	28	5	**17**	33	0	**2**	6

Table 3.7: Number of positive ISSAQ reactions detected per patient proviral clone at each sample time point. The range (low, expected (exp) and high) of positive reactions as calculated by Poisson, bootstrapping and proviral quantification results normalised to proviral copies per one million cell equivalent DNA copies.

The range of expected positive reactions for each ISSAQ run, calculated by Poisson probability and 95% CI determined by non-parametric bootstrapping (Table 3.7), was normalised using CCR5 quantification data (Table 3.6). The normalised specific provirus quantification data (Table 3.7) was used to construct line graphs with error bars showing 95% CI at each point (Figure 3.5) to observe proviral expansion and contraction over time.

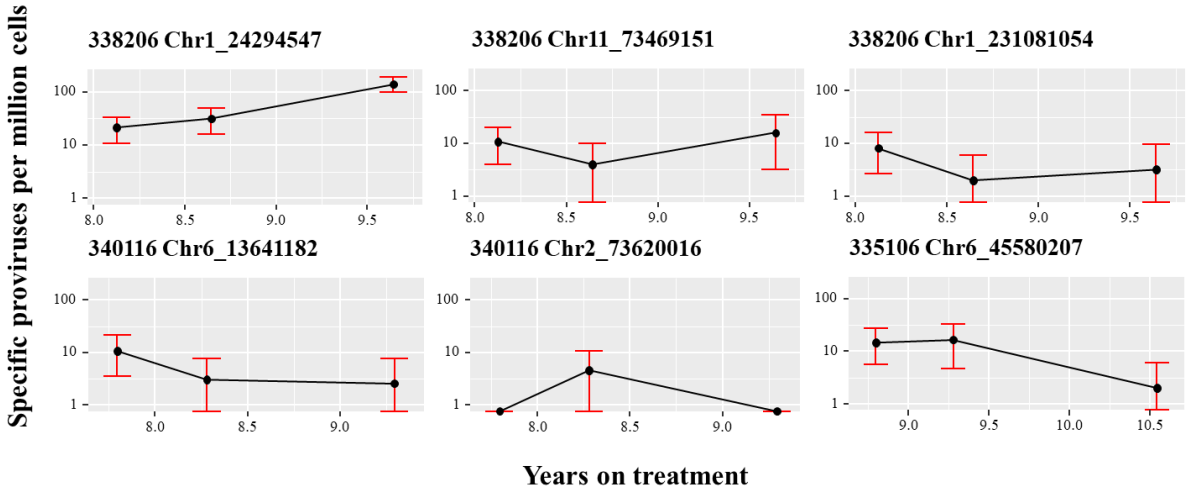


Figure 3.5: Log scale error plots representing ISSAQ results of the selected patient proviral clones over three sample timepoints (indicated by years on treatment). Quantification results are normalised to specific proviruses per one million cell equivalent DNA copies, with error bars showing 95% confidence intervals (CI). The first clone (Chr1_24294547) showed a significant expansion in the relative (normalized) clone size between the second and third time point, as evident from non-overlapping CIs and the last clone (Chr6_45580207) showed a significant decline between the first and last time points (cell yields were similar; Fisher exact p = 0.03). No other differences were statisitcally significant.

3.4 Discussion

Although proviral quantification using various approaches (Yu et al. 2008; Brady et al. 2013; Cillo et al. 2014; Sherman et al. 2017; Anderson & Maldarelli 2018; Bruner et al. 2019) has enabled improved characterisation of the proviral landscape, these assays are unable to fully characterise the proviral landscape due to various levels of proviral intactness. As more data is being generated concerning the prevalence of solo-LTR proviruses (chapter 2 and (Anderson et al. 2020), the need for proviral quantification, irrespective of proviral intactness, becomes greater for the accurate characterisation of the proviral landscape. Additionally, longitudinal quantification of a specific proviral clone could offer valuable evidence for clonal expansion and contraction. This could shed light on the driving forces of clonal waxing and waning, and proviral dynamics.

The ISSAQ assay was therefore designed to enable specific proviral quantification of clones, irrespective of proviral intactness. This is made possible by targeting HIV-1 proviral clone specific human genome sequence in a pre-amplification step, followed by amplification targeting the unique human genome-proviral integration site junction, as performed with ISSPA (chapter 2). Using these integration site specific primers with conservative HIV LTR specific primers and probe, allows for the highly specific amplification and detection of selected proviral clones. Because the prevalence of most clones is expected to be well below the detection limit of a single PCR reaction, a "digitalised" quantification approach with 92 integration site specific reactions was used. With this quantification approach, the Poisson probability distribution and non-parametric bootstrapping are used to calculate the expected number of specific proviral copies with confidence intervals from the number of positive reactions obtained after ISSAQ. The calculated values were normalised to one million cell equivalent DNA copies, obtained from CCR5 quantification of each patient time point sample.

Analyses of ISSAQ were performed to validate the assay before sample testing. To validate the assay, the nested amplification approach was first tested on the major ACH2 clone and thereafter on the nine ISSPA products obtained in chapter 2. Testing the amplification approach on these targets allowed for PCR optimisation, as ISSPA integration site specific primer binding was already confirmed. The ACH2 major clone and 8/9 ISSPA products were successfully amplified and detected by the ISSAQ amplification approach. Only the ISSPA product of patient 334696 clone Chr6_26227488 could not be amplified or detected by ISSAQ due to the R region LTR deletion, which is targeted by the ISSAQ HIV specific LTR primers and probe. Clone Chr6_26227488 was therefore excluded from further testing. Assay specificity was confirmed by performing patient integration site specific reactions on ACH2 DNA, yielding no amplification. Following preliminary ISSAQ results of patient 338206 clone Chr1_24294547, this clone was selected to assess the DNA linear dilution effect on ISSAQ as part of assay validation. A linear dilution effect was observed after performing ISSAQ on four concentrations (1:1, 1:2, 1:3 and 1:4) of patient 338206 time point 3 DNA. Taken together, results from these analyses support the use of ISSAQ for the accurate quantification of specific proviral clones.

In order to accurately quantify specific proviral clones, the number of cell equivalent DNA copies used for each ISSAQ run was determined. The highly sensitive CCR5 quantification assay described in Malnati et al. 2008 was selected to determine the patient DNA input copy number for ISSAQ. Cell equivalent DNA copies used for ISSAQ was determined to range from

237 875 to 742 817 based on patient and sample time point. This allows for the normalisation of ISSAQ data to compare intra and inter patient specific proviral clone quantification data longitudinally.

Specific proviral quantification data for at least one time point of 6/9 patient clones was successfully generated. In addition to the exclusion of patient 334696 clone Chr6_26227488, patient 332406 clones Chr6_32786955 and Chr14_68759177 did not yield any positive ISSAQ reactions after multiple attempts at all time points. Specific proviral quantification data for clone Chr2_73620016 could only be generated for time point 2. The lack of ISSAQ positive reactions at time points 1 and 3 of this clone could be the result of the specific proviral concentration being below the ISSAQ limit of detection. Five proviral clones (Chr1_24294547, Chr11_73469151, Chr1_231081054, Chr6_13641182 and Chr6_45580207) could be quantified at all time points, with an expected value (95% CI) of **2**(0-6) to **140**(98-190) specific proviral copies per one million cell equivalent DNA copies. Error plots of the proviral clone quantification data revealed longitudinal evidence of clonal expansion (338206 Chr1_24294547) and contraction (335106 Chr6_45580207). Although some level of waxing and waning was observed for the other clones, these variations were not statistically significant. No clear pattern of proviral waxing and waning could be observed between the clones. The observed variability in clone size, with two cases having definitive waxing and waning, may be in keeping with recently reported antigen-driven and cytokine-induced proliferation contributing to clonal expansion and the stability of the latent reservoir, rather than continuous cell-autonomous proliferation (Wang et al. 2018; Simonetti et al. 2020). This study yields the first report of severely deleted (solo-LTR) proviral clones persisting in paediatric patients on long term cART (~8 to ~10 years). Furthermore, these novel findings indicate that solo-LTR proviral clones may wax and wane through similar mechanism as described in literature.

The ISSAQ assay offers a novel approach to the quantification of proviral clones and offers integration site specificity capable of individual clonal quantification. Some limitations of the ISSAQ assay are that only a single proviral clone can be targeted at a time and large sample DNA volume is required for each ISSAQ run. As we had limited available sample volume from paediatric patients, the ISSAQ error margins were relatively large. Moreover, as several clones of the same patient needed investigation, available PBMC samples had to be split between assays for various clones at each time point. The precision of this assay and its ability to detect differences in clone size is dependent on the number of cells assayed; and would give more precise results in cases of larger volume blood or leukapheresis samples from adults.

Processing a large sample volume would also increase the sensitivity for the detection of very rare integration events. Another possibility to overcoming these limitations is to perform multiple displacement amplification on sample DNA distributed across 92 wells, yielding ample sample volume for repetitive ISSAQ assays targeting different proviral clones of each patient. The nested PCR approach along with targeting the unique human genome-provirus integration site offers sufficient specificity to the ISSAQ assay.

Taken together, the findings from chapters 2 and 3 demonstrate that some of the most expanded proviral clones are solo-LTRs and that these proviruses persist longitudinally by waxing and waning, suggesting that solo-LTR proviruses could make up a substantial proportion of the proviral landscape. With at least 3/5 and 2/6 of the most expanded clones from patients 338206 and 340116 respectively characterised as solo-LTR proviruses. The novel findings from this study and Anderson et al. 2020 suggest that solo-LTR proviral clones are under-represented with traditional approaches studying the proviral landscape, indicating intact proviruses could make up an even smaller proportion of all integrated HIV than previously thought. Although the prominence of solo-LTR HIV clones does not affect the barrier to cure, it does offer new data on HIV integration and proviral persistence. The prevalence and persistence of solo-LTR HIV proviruses suggest that, similar to HERVs, highly defective genome integrations may be the most frequent result of retroviral cellular infection. As these small inserts may have little impact on cellular function and could impact surrounding gene expression (Buzdin, Prassolov & Garazha 2017), it is in keeping with the incorporation and evolutionary retention of such retroviral elements in the human genome as evidence of archaic retroviral infections in the human evolutionary lineage. There is some evidence that HIV can infect spermatozoa and considering the current number of HIV infected individuals (Baccetti et al. 1998), HIV germline integration is not unfathomable, but would most likely involve highly defective proviruses. This requires further investigation.

Chapter 4

Characterisation of non-suppressible viraemia cases in USA patients (Pittsburgh, PA)

4.1 Introduction

Clinically detectable HIV viraemia in patients on long term cART has traditionally been attributed to non-adherence and/or drug resistance. However, recently a case of infectious viraemia associated with a large cell clone containing an intergenic intact HIV provirus in an individual with metastatic cancer has been reported (Simonetti et al. 2016), confirming that clonally expanded $CD4^+$ T-cell clones can contain replication competent provirus. This suggests an alternative cause of sustained detectable viraemia, rather than virologic failure which implies ongoing new rounds of replication. Therefore, a cohort of patients with sustained clinically detectable viraemia and no regimen relevant drug resistance were investigated for possible clonal viraemia (Halvas et al. 2019). Patients with ongoing viremia despite no drug resistance and objective evidence of adherence may require futher investigation, and in cases where monotypic low level viraemia is found, this may be suggestive of clonal viremia. However, to prove this requires linking plasma viraemia to an HIV-1 infected cell clone. This process is costly and cumbersome but through investigating these cases more efficient methods of characterising clonal viraemia can be developed, in future.

The data presented in this chapter formed part of the Halvas et al. 2019 study, as part of my academic visit to the HIV dynamics and replication program at the University of Pittsburgh under the supervision of Prof. John W. Mellors. The aim of this academic visit was to obtain techniques of value to the van Zyl research group at Stellenbosch University for the characterisation of potential clonal viraemia and other cases of unexplained sustained viraemia.

4.2 Methods

4.2.1 Specimen selection

Twelve patients with prolonged elevated HIV viral loads, while adherent to cART with no relevant drug resistance, and after being offered alternative regimens, who were identified by HIV care providers, were referred and enrolled into the non-suppressible HIV viraemia cohort. Three of these patients are included in this chapter. The selection criteria for this study was as follows: detectable HIV-1 RNA by Roche or Abbott platforms >40 copies/mL, detectable viraemia on multiple sample dates, no treatment-relevant drug resistance (determined by routine drug resistance testing and which would have explained current regimen failure) and no clinical suspicion of non-adherence (as determined by the primary HIV care provider). Leukopak specimens (specimens obtained by leukapheresis) collected from these patients were followed by plasma and PBMC isolation to be used for this study.

4.2.2 Quantitative viral outgrowth assay (QVOA)

Viral outgrowth was performed on PBMCs from patients, having monotypic plasma viral populations with at least one identical sequence from PBMC DNA, suggestive of clonality. Viral outgrowth was performed to identify p6ProRT sequences released from replication competent proviruses, identical to the identified sequences in plasma, In addition, identifying identical sequences in multiple QVOA replicates would suggest clonality, as this would originate from different cells. The QVOA assay was performed over 21 days as follows:

Day -1: Two days prior to initiation of the QVOA assay, PBMCs from 120 mL EDTA whole blood of an HIV negative donor were isolated and incubated in 30 mL super T cell media (STCM) [made up out of 500 mL phenol red-free RPMI (Lonza), 5 mL Glutamax (Invitrogen), 10% heat-inactivated FBS (Fisher Scientific), 3 mL Penicillin/Streptomycin (Invitrogen), 100 U/mL rh-IL-2 (Fisher Scientific) and 2% T-cell growth factor] with 0.5 µg/ml (PHA) (Sigma-Aldrich, MO, USA) at 37^{o}C until CD8 depletion (Day 2).

Day 1: On the day of QVOA initiation 120 mL EDTA whole blood from an HIV negative donor was inactivated by γ-irradiation with 5000R in a Caesium source irradiator. PBMCs were isolated from these cells and resuspended in STCM at a concentration of 2.5x10^{6} cells/mL. Cells were chilled on ice after the addition of 1 µg/ml PHA.

PBMCs previously isolated from the selected HIV positive patient were thawed at 37^{o}C and CD4^{+} T-cells were isolated using the EasySep^{TM} human CD4^{+} T-cell isolation kit (STEMCELL

Technologies, Vancouver, Canada) according to manufacturer's instructions. Isolated resting $CD4^+$ T-cells were plated as follows (Figure 4.1): 4 wells of $1x10^6$ cells/mL (1M) were plated in 6-well plates with 10 million irradiated PBMCs + PHA and the volume made up to 8 mL per well with STMC. Five wells each of HIV positive $CD4^+$ T-cells were also plated at concentrations of $3x10^5$ cells/mL (300k), $1x10^5$ cells/mL (100k) and $3x10^4$ cells/mL (30k) in 24-well plates with 2.5 million irradiated PBMCs + PHA and the volume made up to 2 mL per well with STMC. A negative control was included on the lowest concentration plate by replacing the HIV positive $CD4^+$ T-cells with STCM. Plates were incubated at 37°C until day 2.

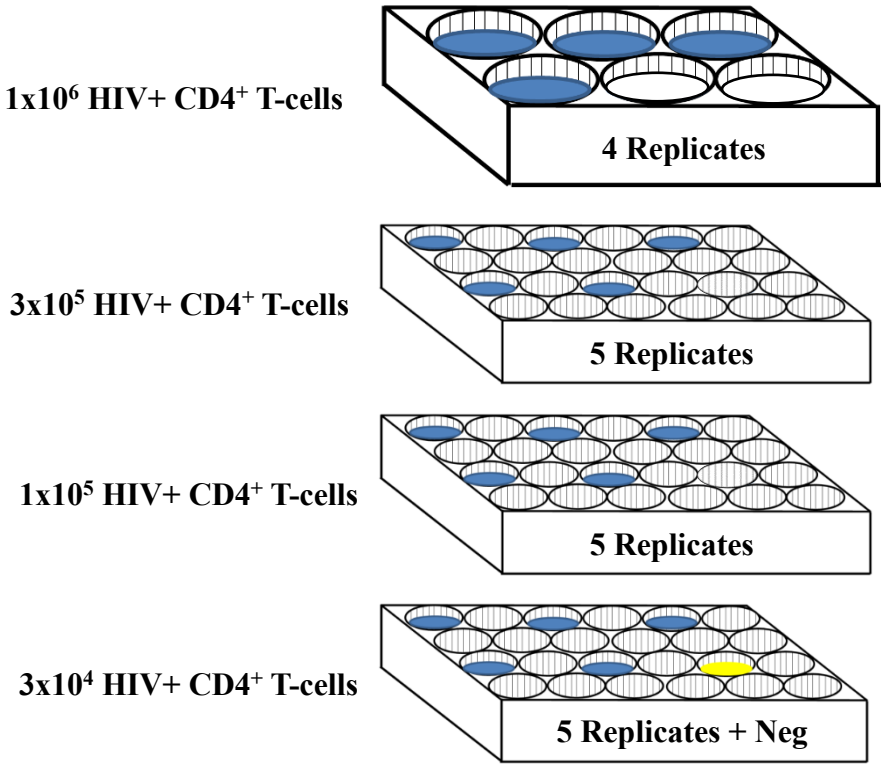


Figure 4.1: Representation of the plate layout for quantitative viral outgrowth. The number of HIV positive $CD4^+$ T-cells used are indicated along with the number of replicates performed at each cell dilution. The HIV negative control (Neg) was included on the $3x10^4$ HIV positive $CD4^+$ T-cells dilution plate, indicated in yellow.

Day 2: Media was removed from wells (6 ml from the 6-well plate and 1.5 ml from 24-well plates) and replaced with STCM, plates were incubated at 37°C for two hours before adding CD8-depleted T-lymphoblasts. PBMCs isolated from the HIV negative donor on day -1 were CD8-depleted using Anti-human CD8 magnetic particles (BD Biosciences, Franklin Lakes, NJ, USA) according to manufacturer's instructions. The CD8-depleted lymphoblasts were resuspended with STCM at a concentration of $0.67x10^6$ cells/mL and seeded into QVOA wells

($4x10^6$ cells per well for the 6-well plate and $1x10^6$ cells per plate for 24-well plates), plates were incubated at 37°C.

Day 5: Media was changed by replacing 3 mL per well for the 6-well plate and 750 µl per well for 24-well plates with STCM. PBMCs from an HIV negative donor were isolated and prepared for CD8-depletion as done on day -1.

Day 7: Supernatant (200 µl) from each QVOA well was used to perform p24 antigen detection with the Alliance HIV-1 p24 antigen ELISA kit (PerkinElmer, Waltham, MA, USA) according to manufacturer's instructions. HIV negative donor PBMCs, prepared on day 5, were CD8-depleted and resuspended in STCM at a concentration of $1x10^6$ cells/mL. Incubated QVOA cells were split by gently mixing cells in each well and discarding half of the volume. Lymphoblasts were added to replace the discarded cells, $4x10^6$ cells in 4 ml STCM was added to each well for the 6-well plate and $1x10^6$ cells in 1 ml STCM was added to each well for the 24-well plates. Plates were incubated at 37°C until day 12.

Day 12: Media change and HIV negative donor PBMC isolation proceeded as performed on day 5.

Day 14: Supernatant (200 µl) was used for HIV-1 p24 detection. The QVOA cells were split and Lymphoblasts added as performed on day 7.

Day 19: Media was changed as on days 5 and 12.

Day 21: Supernatant (200 µl) was tested for HIV-1 p24 antigen as on days 7 and 14.

Additional supernatants (700 µl) from all QVOA wells on days 7, 14 and 21 were stored at -80°C until RNA isolation.

4.2.3 PBMC and DNA isolation

PBMC and DNA isolation were performed exactly as described in chapter 2, sections 2.2.2 and 2.2.3. Isolated DNA was stored in 200 µl 5 mM Tris-HCl (pH 8.0) (Sigma-Aldrich) at -20°C until use

4.2.4 RNA isolation

Selected plasma samples and QVOA supernatants were centrifuged in an Eppendorf 5430 R centrifuge (Eppendorf, Hamburg, Germany) at 21100 rcf for 60 min at 4^{o}C prior to lysis. RNA isolation proceeded identical to DNA isolation (2.2.3). Purified plasma RNA was resuspended in 20 µl cold 5 mM Tris-HCl (pH 8.0) (Sigma-Aldrich) and QVOA RNA was resuspended in 50 µl cold 5 mM Tris-HCl (pH 8.0).

4.2.5 cDNA synthesis

Complimentary DNA (cDNA) synthesis was performed with the Superscript III first strand cDNA synthesis kit (Invitrogen Corp., Carlsbad, CA, USA) according to manufacturer's instructions. A 50 µl cDNA synthesis reaction was performed using the entire 20 µl plasma RNA volume, and 100 µl cDNA synthesis reactions were performed on dilutions (1:10, 1:100 and 1:100) of QVOA RNA. Reactions consisted of 1x Reverse transcription buffer, 5 mM $MgCl_2$, 0.01 M DTT, 2 units/µl RNaseOUT, 10 units/µl SuperScript III Reverse transcriptase, 0.5 mM dNTPs, 30 µM HIV specific reverse primer 3500 (5-CTA TYA AGT CTT TTG ATG GGT CAT AA-3'), and nuclease free water (Promega Corp.). Extracted RNA, dNTPs and the primer were denatured at 65^{o}C for 5 min and immediately chilled on ice for 2 min. After the addition of the reaction mix reverse transcription was performed at 50^{o}C for 50 min followed by enzyme inactivation at 85^{o}C for 5 min. The cDNA was used immediately after synthesis or stored at -20^{o}C until use.

4.2.6 p6PrRT single genome amplification

Single genome sequencing (SGS) covering the p6, *Protease* and *Reverse transcriptase* (p6PrRT) region of HIV-1 was performed on cDNA and PBMC DNA as described previously (Kearney et al. 2008). The p6PrRT SGS PCR protocol consists of a nested amplification approach. According to Poisson probability, single copy template amplification is most likely to be achieved when less than 30% of the total reactions are successfully amplified. Therefore, p6PrRT SGS was first performed on a "dilution plate" consisting of 1:3 serially diluted template nucleic acid to find the target single copy dilution. This was followed by p6PrRT SGS on an "expansion plate" consisting of 92 reactions at the target dilution and four NTC wells as controls (Figure 4.2).

The p6PrRT SGS amplification was performed with the Platinum™ Taq DNA Polymerase High Fidelity system (Invitrogen Corp.), with each reaction (10 µl) consisting of: 1x High Fidelity PCR Buffer, 2 mM $MgSO_4$, 0.4 units Platinum Taq High Fidelity enzyme, 0.2 mM dNTPs, 0.16 µM HIV-1 subtype B optimised forward and reverse primers (Table 4.1) and nuclease free water. Pre-nested amplification conditions consisted of an initial denaturing step

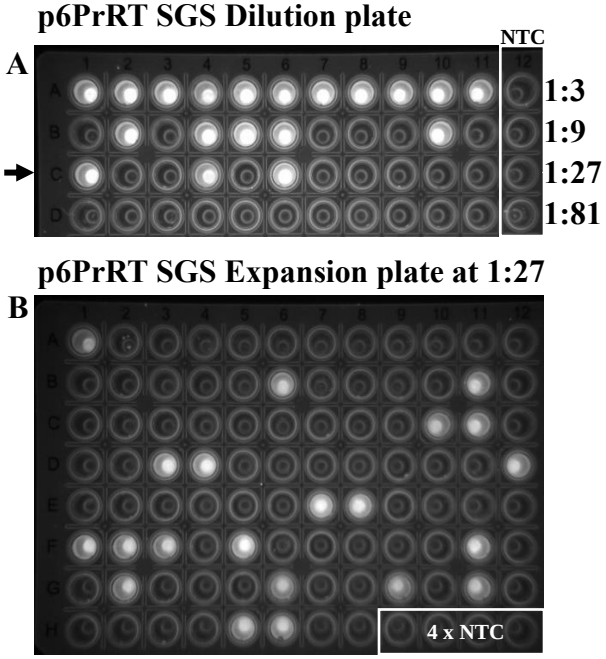


Figure 4.2: Representation of the p6-Protease-Reverse transcriptase (p6PrRT) single genome sequencing (SGS) plate layout for A) Dilution plate, indicating 11 reactions and one NTC for each 1:3 serial dilution and the arrow indicates the target dilution as 1:27. B) Expansion plate with 92 reactions of the target dilution (1:27) and four NTC reactions.

(94°C, 2 min) followed by 45 cycles of denaturation (94°C, 30 sec), annealing (50°C, 30 sec) and extension (72°C, 1 min 30 sec), with a final extension step (72°C, 3 min). The pre-nested PCR products were diluted with 83 µl 5 mM Tris-HCl (pH 8.0) (Sigma-Aldrich) per reaction and 2 µl diluted amplicon was plate stamped into the nested PCR plate. Nested amplification conditions consisted of an initial denaturing step (94°C, 2 min) followed by 41 cycles of denaturation (94°C, 30 sec), annealing (55°C, 30 sec) and extension (72°C, 1 min), with a final extension step (72°C, 3 min).

Table 4.1: Primers used for HIV-1 subtype B p6-Protease-Reverse transcriptase (p6PrRT) single genome sequencing (SGS) amplification.

Primer name	Orientation	Sequence (5'-3')	Location in HXB2 genome
Pre-nest F	Forward	GAT GAC AGC ATG TCA GGG AG	1827 – 1846
Pre-nest R	Reverse	CTA TTA AGT ATT TTG ATG GGT CAT AA	3503 – 3528
Nest-F	Forward	GAG TTT TGG CTG AGG CAA TGA G	1871 – 1892
Nest-R	Reverse	CAG TTA GTG GTA TTA CTT CTG TTA GTG CTT	3409 – 3438

4.2.7 PCR product analysis and purification

Nested p6PrRT PCR reactions were diluted with 45 µl 5 mM Tris-HCl (pH 8.0) (Sigma-Aldrich) and assessed for amplicon as described in chapter 2 section 2.2.8. PCR wells containing amplicon as observed by fluorescence were purified by adding 2 µl Exonuclease 1 (Exo) (New England Biolabs Inc. (NEB), Ipswich, MA, USA) and 4 µl Shrimp Alkaline Phosphatase (rSAP) (NEB) to each selected well. Reactions were incubated at 37^{o}C for 20 min followed by enzyme inactivation at 80^{o}C for 15 min and diluted with 40 µl 5 mM Tris-HCl (pH 8.0) (Sigma-Aldrich). Purified PCR products were stored at -20^{o}C until use.

4.2.8 Nucleotide sequencing and analysis

Nucleotide sequences of purified p6ProRT amplicons were determined as discussed in chapter 2 section 2.2.9, using 3 µl Exo-rSAP purified amplicon per reaction and four overlapping primers (Table 4.2) in independent reactions. Nucleotide Sanger sequence chromatograms were analysed, and sequences verified with Sequencher^{TM} 4.10.1 (Gene Codes Corp.) and BioEdit Sequence Alignment Editor (V7.2.5). Consensus sequences were generated for each positive p6PrRT SGS amplicon.

Table 4.2: Primers used for overlapping nucleotide Sanger sequencing of p6-Protease-Reverse transcriptase (p6PrRT) single genome sequencing (SGS) amplicons

Primer name	Orientation	Sequence (5'-3')	Location in HXB2 genome
p6PrRT-A	Forward	TGT TGG AAA TGT GGA AAG GAA GGA C	2026 – 2050
p6PrRT-B	Forward	ATG GCC CAA AAG TTA AAC AAT GGC	2599 – 2622
p6PrRT-C	Reverse	TTC TTC TGT CAA TGG CCA TTG TTT AAC	2610 – 2636
p6PrRT-D	Reverse	TTG CCC AAT TCA ATT TTC CCA CTA A	3327 – 3351

4.2.9 Phylogenetic analyses

Nucleotide sequences were aligned with the consensus B sequence (MN919177.1) using MAFFT Version 7 (https://mafft.cbrc.jp/alignment/server/). Neighbour joining p-distance analysis with 1000 bootstraps were performed to construct phylogenetic trees in MEGA7 (Kumar, Stecher & Tamura 2016). Phylogenetic trees were analysed for identical sequences spanning the p6PrRT region. Aligned consensus sequences were analysed for polymorphisms responsible for branching observed on phylogenetic trees.

4.2.10 Multiple displacement amplification and product purification

Multiple displacement amplification (MDA) was performed on PBMC DNA from patients where identical p6PrRT sequences, as determined by phylogenetic analysis, were obtained from plasma, PBMCs and QVOA. The MDA reactions were performed to yield greatly amplified genomic material where possible HIV-1 proviral clones are present to be used for downstream molecular analyses such as integration site determination and near full length proviral amplification. Isolated PBMC DNA was diluted in nuclease free water (Promega Corp.) to the determined p6PrRT endpoint and seeded at 2 µl per well in a 96 well PCR plate. This was followed by the addition of the denature solution comprising of 200 mM KOH (Sigma-Aldrich) and 5 mM EDTA (Invitrogen Corp.). Each reaction was pipette mixed and incubated at 25^{o}C for 3 min, the plate was then chilled at 4^{o}C for 2 min prior to the addition of the neutralisation solution comprising of 100 mM Tris-HCl pH 7.5 (Sigma-Aldrich) and 332 mM Trehalose (Sigma-Aldrich). The MDA master mix consisting of: 1x Phi29 Polymerase Buffer (NEB), 2 mM DTT (Invitrogen Corp.), 100 mM bovine serum albumin (BSA) (NEB),

3 mM dNTPs (Invitrogen Corp.), 600 mM Trehalose, 0.02 mM random heptamer primer, 0.04 mM Hg19 anti-HIV primers and 0.75 mM Phi29 Polymerase enzyme (NEB) was added to the cold neutralised reaction to a total of 25 µl. The reactions were incubated for 20 hours at 40^{o}C followed by an enzyme deactivation step for 10 min at 65^{o}C.

Solid phase reversible immobilisation (SPRI) DNA fragment purification beads from Beckman Coulter Life Sciences (Brea, CA, USA) were used for MDA product purifications. SPRI beads (20 µl) were added to MDA reactions and mixed well for 5 min. Reactions were incubated on a magnet for 2 min and the supernatants were carefully removed and discarded without disturbing the bead pellet. Each reaction was washed twice with 170 µl 80% Ethanol for 30 sec followed by air drying of the bead pellets for 3 min while on the magnet. Bead pellets were resuspended in 35 µl 5 mM Tris-HCl (pH 8.0) (Sigma-Aldrich) by incubation at 37^{o}C for 5 min with intermittent pulse vortexing. Each reaction plate was then incubated on the magnet again for 2 min followed by the removal of the supernatant containing purified MDA products and stored in 96 well PCR plates at -20^{o}C until use.

4.2.11 Post MDA p6PrRT SGS

Purified MDA products were screened by p6PrRT SGS for identical p6PrRT sequences previously detected from plasma, PBMC and QVOA sources, in order to match MDA product wells with possible proviral clones for further analyses. Purified MDA reactions were plate stamped for p6PrRT SGS, 2 µl of each purified MDA reaction was used for p6PrRT amplification in final reaction volumes of 10 µl, as described in section 4.2.6. Post p6PrRT amplification proceeded as previously mentioned (section 4.2.7 and 7.2.8) and nucleotide sequence data was included in neighbour joining phylogenetic trees. Purified MDA products that yielded p6PrRT sequences identical to monotypic phylogenetic clusters (possible clones) were sent to the HIV Dynamics and Replication Program at the NCI for integration site (IS) determination by integration site analysis (ISA).

4.2.12 Host to full-length provirus to host amplification (HFH)

After obtaining integration site data from p6PrRT clones of interest, HFH was attempted to investigate and characterise the proviral clone. Using HFH, the provirus can be amplified in two overlapping half genome segments with ample host genome on both 5' and 3' ends too definitively link the provirus to the IS. For this assay, p6PrRT primers targeting HIV are used together with human genome targeted primers up-stream and down-stream of the proviral integration site (Figure 4.3). The Hg primers were designed with similar melting temperatures as the corresponding p6PrRT primers, and special consideration was taken to avoid primer dimers. Proviral orientation was taken into consideration to pair up the correct HIV and human genome primers (Table 4.3).

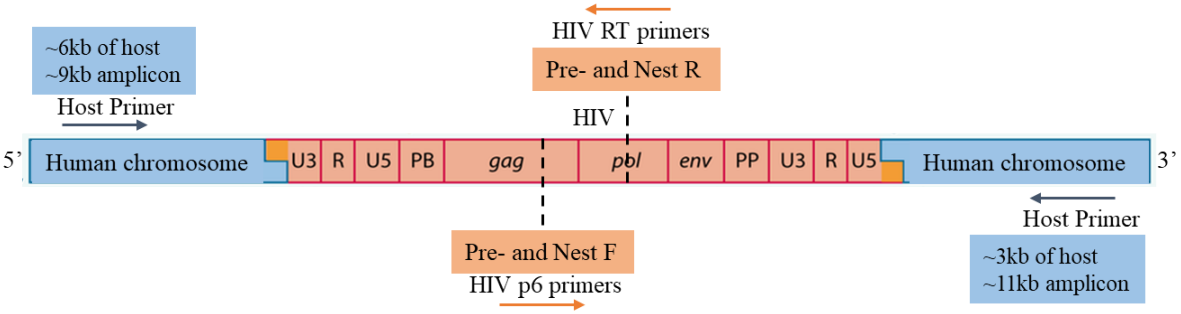


Figure 4.3: Respresentation of the host to full length to host proviral amplification strategy, indicating HIV and human genome specific primers.

Table 4.3: Human genome and HIV specific primers used for host to full length provirus to host proviral amplification of a selected integration site (IS).

Primer name	Orientation	Sequence (5'-3')	Location in HXB2 genome
Host-9kb F	Forward	GGA ATT CTA TTC CCA GCA TTG TGG	9250 bases upstream of IS
Host-6kb F	Forward	GAA TCA CTA TCC TGC TAT ATG AAA TCT AC	5830 bases upstream of IS
Pre-nest R	Reverse	CTA TTA AGT ATT TTG ATG GGT CAT AA	3503 – 3528
Nest-R	Reverse	CAG TTA GTG GTA TTA CTT CTG TTA GTG CTT	3409 – 3438
Pre-nest F	Forward	GAT GAC AGC ATG TCA GGG AG	1827 – 1846
Nest-F	Forward	GAG TTT TGG CTG AGG CAA TGA G	1871 – 1892
Host-R	Reverse	TAA GAA GTC AAG AGA TTG TTT CAT TC	3410 bases down-stream of IS

A semi-nested PCR approach was selected for HFH, pairing a single host primer with both of the respective p6PrRT primers (Figure 4.3). Amplification was performed with Ranger mix (Bioline meridian bioscience) and each reaction (10 µl) consisted of: 1x Ranger mix, 0.4 µM forward and reverse primer respectively and nuclease free water (Promega Corp.). Two sets of amplification conditions were attempted, first: an initial denaturing step (95^{o}C, 2 min) followed by 30 cycles of denaturation (98^{o}C, 10 sec), annealing and extension in one step (58^{o}C, 12 min). A final extension (58^{o}C, 15 min) was included to ensure full-length products. Alternatively, the second set of amplification conditions consisted of an initial denaturing step (95^{o}C, 2 min), followed by 30 cycles of denaturation (95^{o}C, 10 sec), annealing (58^{o}C for 30 sec) and extension (68^{o}C, 12 min). A final extension step (68^{o}C, 15 min) was included to ensure full-length amplicons. Amplicons were analysed as discussed in chapter 2 section 2.2.8.

4.3 Results

4.3.1 Selected specimens

Three patients presenting with clinical data indicative of clonal viraemia (ie. prolonged elevated viral load with no regimen relevant drug resistance) that required further investigation, were selected for this project. Viral load data and a brief clinical overview for each patient are provided, and selected specimen time points are indicated. Patient T13 clinical overview and viral load data indicate sustained detectable viraemia for 17.3 months. Samples from two time points (TP), ~3 months apart, were collected (Figure 4.4). The collected sample viral loads are indicated as 78 copies/mL at TP1 and 100 copies/mL at TP2.

Age	48
Sex	Male
Race	Caucasian
Diagnosis	06/2010
Pre-ART VL (c/mL)	181,000
Roche Taqman (c/mL)	TP1: 78 TP2: 100
Duration of Suppression Below LOD	unknown
Duration of Detectable Viremia	17,3 months
Current ART Regimen	ECF-TAF and Darunavir

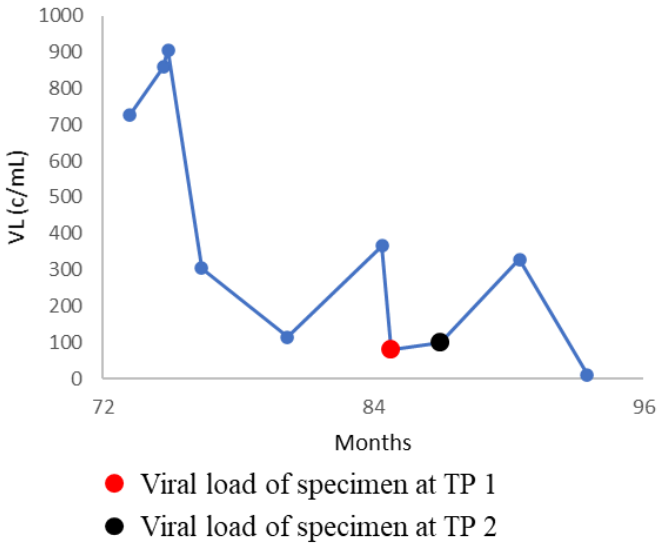


Figure 4.4: Clinical overview and viral load data for patient T13. Time point 1 (TP1) sample is indicated in red and TP2 in black. (Elvitegravir/ Cobicistat/ Emtricitabine/ Tenofovir Alafenamide (ECF-TAF)).

The clinical overview and viral load data for patient L14 indicate detectable viraemia for 1.8 months, and one sample collected (TP1) with a viral load of 583 copies/mL (Figure 4.5).

Age	50
Sex	Male
Race	Caucasian
Diagnosis	12/2002
Pre-ART VL (c/mL)	2,346,667
Roche Taqman (c/mL)	583
Duration of Suppression Below LOD	06/2004-02/2017
Duration of Detectable Viremia	1.8 months
Current ART Regimen	ECF-TAF and Darunavir

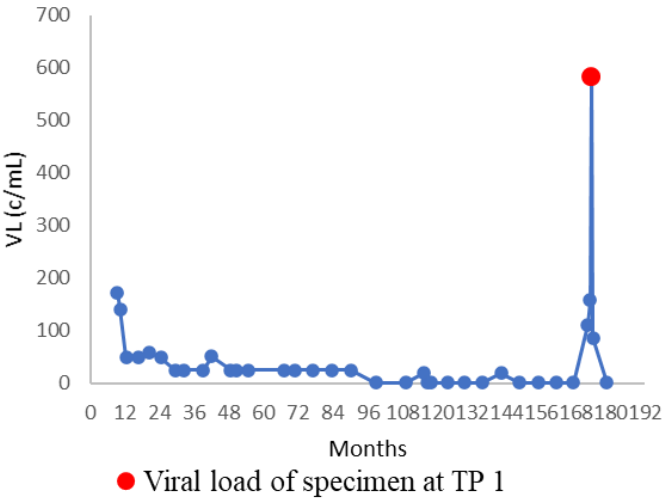


Figure 4.5: Clinical overview and viral load data for patient L14. Time point 1 (TP1) sample is indicated in red. (Elvitegravir/ Cobicistat/ Emtricitabine/ Tenofovir Alafenamide (ECF-TAF)).

Patient D15 clinical overview and viral load data indicate a longitudinal unstable viral load. After the last period of viral suppression, HIV was detectable for over 12 months at sampling TP1, with a viral load of 116 copies/mL (Figure 4.6).

Age	50
Sex	Male
Race	Caucasian
Diagnosis	02/2004
Pre-ART VL (c/mL)	>500,000
Roche Taqman (c/mL)	116
Duration of Suppression Below LOD	03/2012-08/2013 03/2015-09/2016
Duration of Detectable Viremia	Fluctuating, >12 months at TP1
Current ART Regimen	Darunavir and Cobicistat

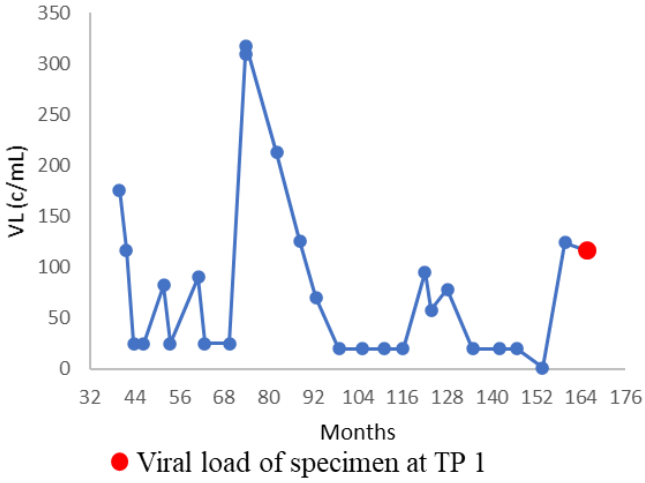


Figure 4.6: Clinical overview and viral load data for patient D15. Time point 1 (TP1) sample is indicated in red.

4.3.2 Preliminary p6PrRT phylogenetic data

Preliminary p6PrRT SGS data generated from samples of patients T13, L14 and D15 indicate various levels of identical sequences as indicated by p-distance neighbour joining phylogenetic trees (Figure 4.7).

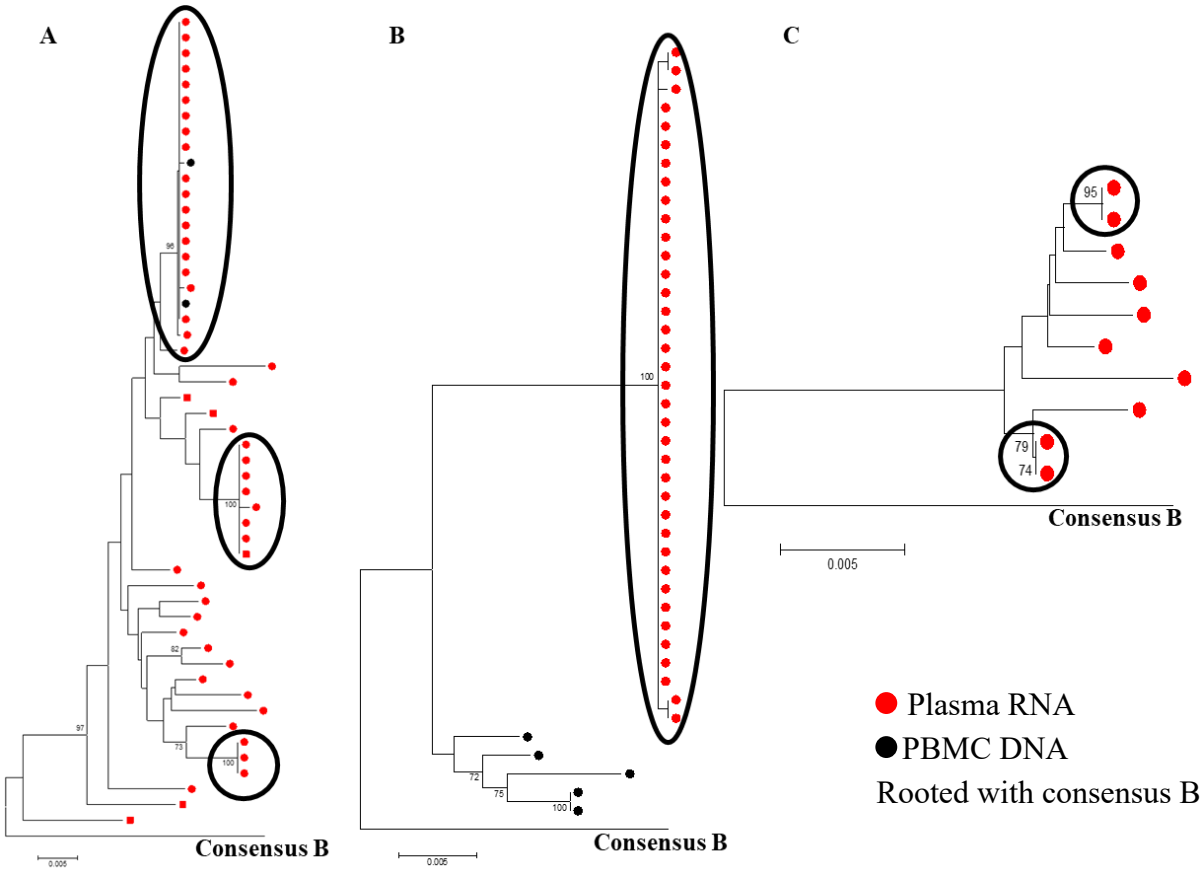


Figure 4.7: P-distance neighbour joining phylogenetic trees of p6PrRT SGS sequence data from patients A) T13, B) L14 and C) D15. Sequence data obtained from plasma RNA are indicated by red dots and data from PBMC DNA by black dots. All phylogenetic trrees were rooted with HIV consensus B sequence (MN919177.1) and identical sequences are circled.

4.3.3 QVOA p24 detection

Based on preliminary phylogenetic analyses (Figure 4.7) indicating identical p6PrRT sequences obtained from plasma RNA and PBMC DNA, patient T13 TP1, with ample PBMC sample volume, was selected for QVOA. HIV p24 detection from days 7, 14 and 21 of QVOA indicated significant replication with nearly all replicates being detectible at days 14 and 21 (Table 4.4).

Table 4.4: p24 detection results of day7, 14 and 21 quantitative viral outgrowth assay supernatants for patient T13.

Day	Dilution	Rep	p24
7	$1x10^6$ (1M)	1	Pos
		2	Low Pos*
		3	Pos
		4	Pos
	$3x10^5$ (300k)	1	Pos
		2	Pos
		3	Pos
		4	Pos
		5	Pos
	$1x10^5$ (100k)	1	Low Pos*
		2	Pos
		3	Low Pos*
		4	Low Pos*
		5	Pos
	$3x10^4$ (30k)	1	Neg
		2	Low Pos*
		3	No data#
		4	No data#
		5	No data#
	Neg		Neg
Day	**Dilution**	**Rep**	**p24**
21	$1x10^6$ (1M)	1	Pos
		2	Pos
		3	Pos
		4	Pos
	$3x10^5$ (300k)	1	Pos
		2	Pos
		3	Pos
		4	Pos
		5	Pos
	$1x10^5$ (100k)	1	Pos
		2	Pos
		3	Pos
		4	Pos
		5	Pos
	$3x10^4$ (30k)	1	Pos
		2	Pos
		3	Neg
		4	Pos
		5	Neg
	Neg		Neg

Day	Dilution	Rep	p24
14	$1x10^6$ (1M)	1	Pos
		2	Pos
		3	Pos
		4	Pos
	$3x10^5$ (300k)	1	Pos
		2	Pos
		3	Pos
		4	Pos
		5	Pos
	$1x10^5$ (100k)	1	Pos
		2	Pos
		3	Pos
		4	Pos
		5	Pos
	$3x10^4$ (30k)	1	Pos
		2	Pos
		3	Neg
		4	Pos
		5	Neg
	Neg		Neg

*****A slight colour change in the p24 detection well was deemed as "Low Pos"
#p24 detection failed

The infectious units per million (IUPM) for patient T13 was calculated as 36.1 using the online IUPMStats v1.0 calculator (https://silicianolab.johnshopkins.edu/) based on p24 results from days 14 and 21. Day 14 QVOA supernatants with detectable HIV p24 antigen were selected for p6PrRT SGS along with day 7 300k replicate 3 supernatant. Due to identical p24 results for day 14 and 21, day 14 replicates were selected for sequence analyses for a greater chance to detect both fast and slower growing HIV variants.

4.3.4 p6PrRT SGS

The p6PrRT SGS assay was performed on genomic cell associated DNA from PBMCs and plasma RNA of patients T13, L14 and D15. For patient T13 p6PrRT SGS was performed on TP1 and 2 cell associated DNA and plasma sequences, along with TP1 QVOA supernatants (day 7 300k-3 and day 14) and TP1 MDA products. A total of 309 p6PrRT sequences were generated for patient T13 from these sources (Table 4.5).

Table 4.5: Patient T13 p6PrRT SGS nucleotide sequences generated from various sources as indicated, including plasma, cell associated DNA (CAD), quantitative viral outgrowth assay (QVOA) supernatant and multiple displacement amplification (MDA).

Source			Sequences generated	
TP1 Plasma			51	
TP1 CAD			65	
TP1 Day 7 QVOA 300k-3			5	
TP1 Day 14 QVOA	1M	1	4	Day 14 QVOA total: 74
		2	3	
		3	4	
		4	5	
	300k	1	4	
		2	4	
		3	3	
		4	3	
		5	4	
	100k	1	6	
		2	5	
		3	4	
		4	4	
		5	6	
	30k	1	5	
		2	5	
		4	5	
TP1 MDA			55	
TP2 Plasma			24	
TP2 CAD			35	

Time point 1 p6PrRT SGS amplification of patients L14 and D15 produced a total of 82 and 36 nucleotide sequences respectively (Table 4.6).

Table 4.6: Patient L14 and D15 p6PrRT sequences generated from cell associated DNA (CAD) and plasma

Source		Sequences generated
L14	TP1 Plasma	27
	TP1 CAD	55
D15	TP1 Plasma	10
	TP1 CAD	26

4.3.5 Phylogenetic and nucleotide sequence analyses

Neighbour-joining phylogenetic trees constructed of p6PrRT sequence data from patients T13, L14 and D15 with evolutionary distance calculated using the p-distance model were analysed for the presence of monotypic clusters suggestive of clonality. Two distinct clusters of identical sequences (clones 1 and 2) consisting of sequences obtained from TP1 plasma, TP2 plasma, TP1 cell associated DNA, TP2 cell associated DNA, various QVOA supernatants and MDA products were identified with phylogenetic analysis for patient T13 (Figure 4.8). A potential sub-clone 1 consisting of identical sequences with a single base mutation compared to clone 1 sequences were identified (Figure 4.8). A third cluster (clone 3) consisting of identical sequences obtained from TP1 plasma, TP1 cell associated DNA, TP2 cell associated DNA and MDA products were also identified., The inducibility and replication competence of clone 3 are unknown due to the lack of identical sequences from QVOA supernatants (Figure 4.8).

Nucleotide sequences spanning p6PrRT of clone 1 and sub-clone 1 were further investigated to identify the polymorphisms responsible for the separate phylogenetic clustering. A single nucleotide polymorphism (SNP) (G → A) was identified to be responsible for the formation of sub-clone 1 (Figure 4.9).

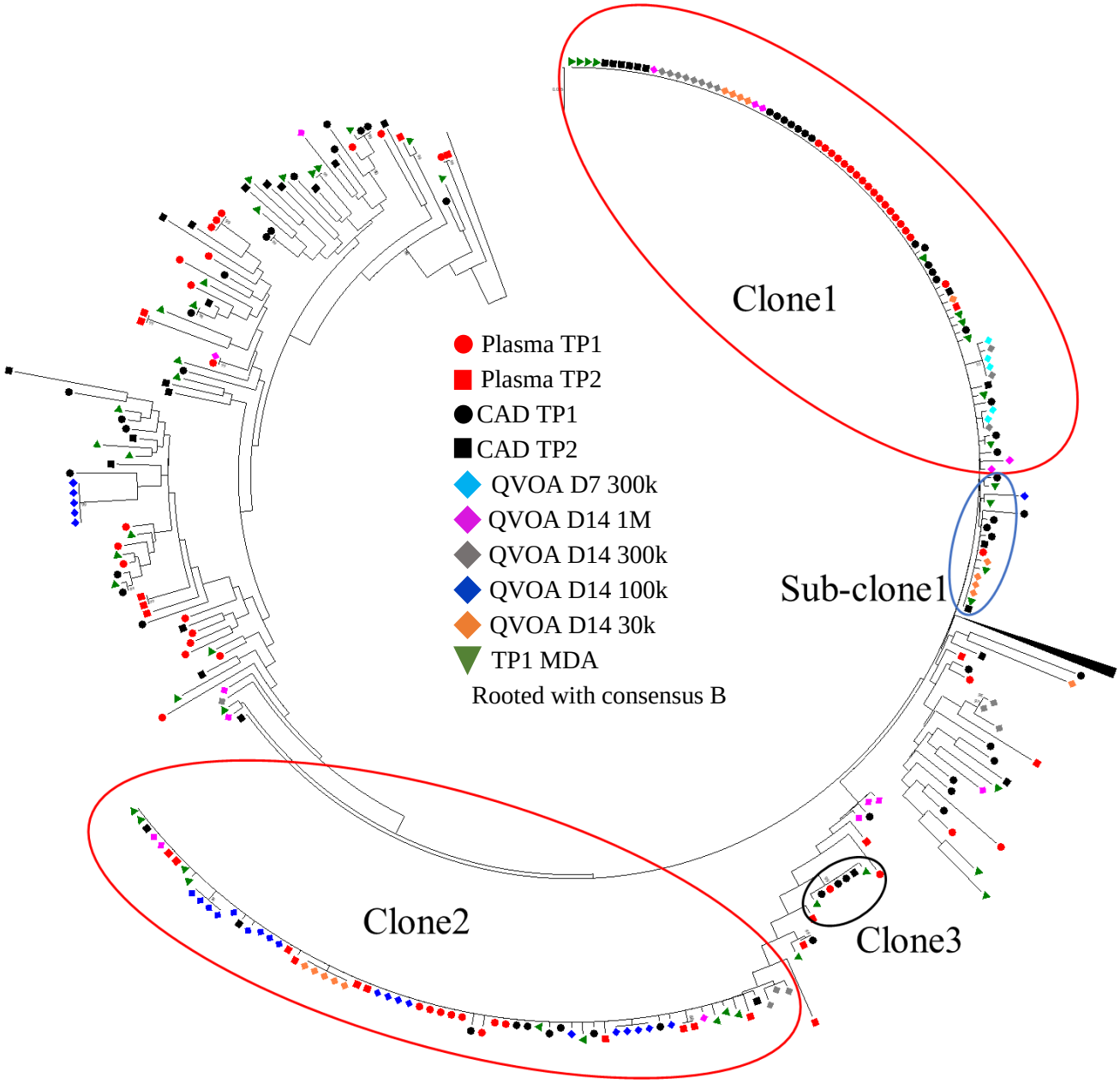


Figure 4.8: Neighbour joining p-distance phylogenetic tree of p6PrRT nucleotide sequences generated from TP1 and 2 plasma, TP1 and 2 cell associated DNA (CAD), TP1 QVOA supernatants and TP1 MDA products of patient T13. Identical sequence clusters "Clones" 1 and 2 circled in red, sub-clone 1 circled in blue and clone 3 circled in black are indicated. Sequence origins are represented with various symbols as indicated by the legend. A cluster of hypermutated sequences adjacent but not related to sub-clone 1 was collapsed into a wedge. The tree was constructed with 1000 bootstraps and rooted to Consensus B (MN919177.1).

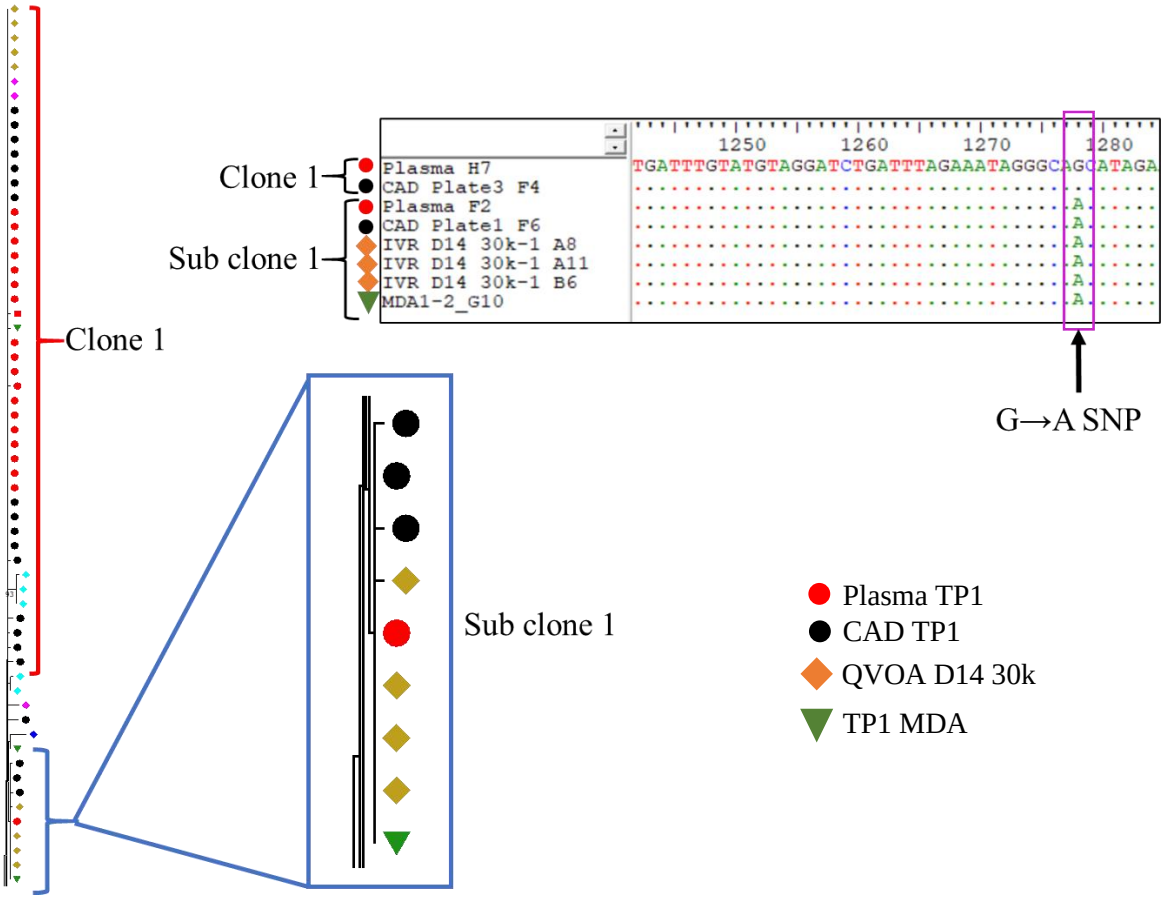


Figure 4.9: Excerpt of clone 1 and sub-clone 1 from patient T13 phylogenetic tree along with nucleotide sequences of clone 1 and sub-clone 1 indicating the single nucleotide polymorphism (SNP) responsible for the phylogenetic branching.

All (27) plasma p6PrRT SGS nucleotide sequences obtained from patient L14 are identical or nearly identical, as indicated by the phylogenetic analysis (Figure 4.10). No identical clusters were observed with patient L14 cell associated DNA sequences, and none of the 55 cell associated DNA sequences were similar to any plasma sequences.

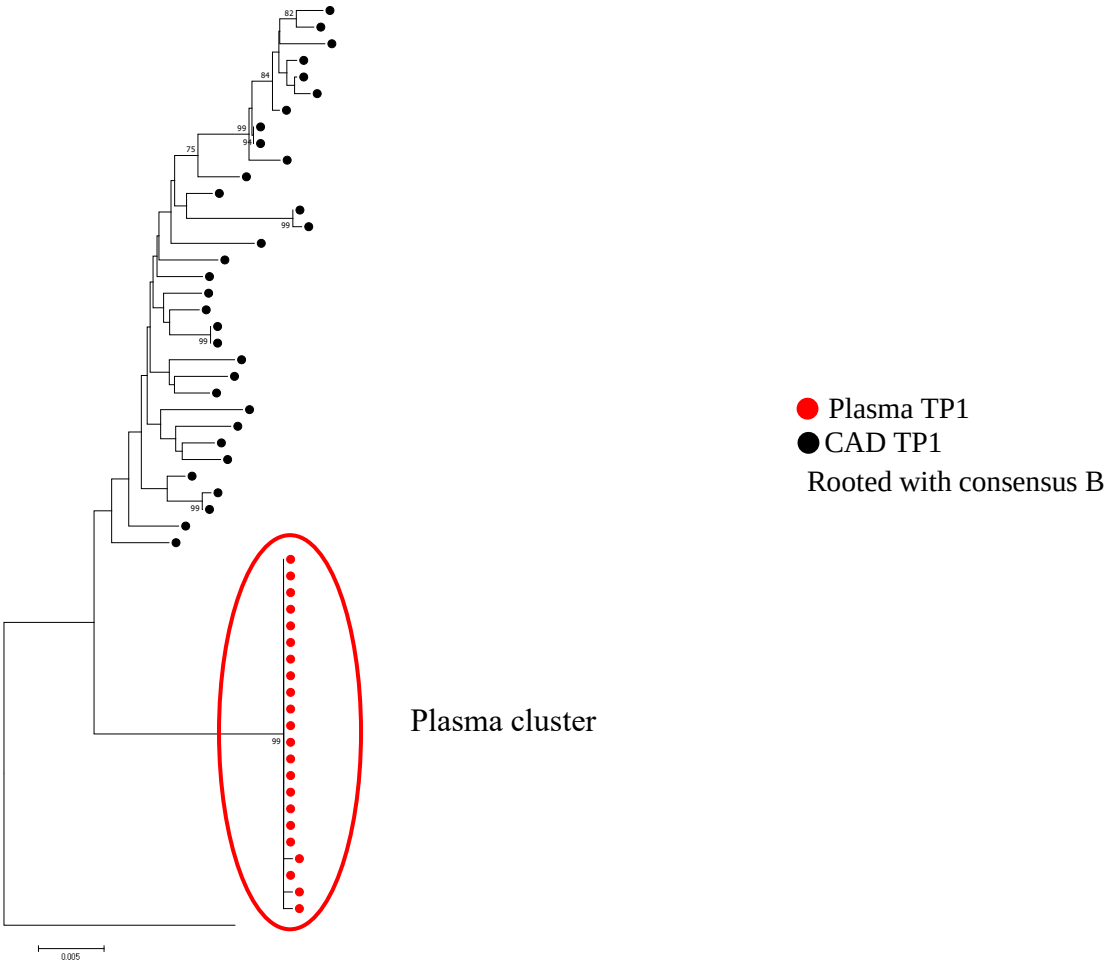


Figure 4.10: Neighbour joining p-distance phylogenetic tree of p6PrRT nucleotide sequences generated from TP1 plasma and cell associated DNA (CAD) samples of patient L14. The identical plasma sequence cluster is circled in red. Sequence origins are represented with symbols as indicated by the legend. The tree was constructed with 1000 bootstraps and rooted to Consensus B (MN919177.1).

Phylogenetic analysis of TP1 plasma and cell associated DNA sequences from patient D15 did not indicate any major clustering of identical sequences (Figure 4.11).

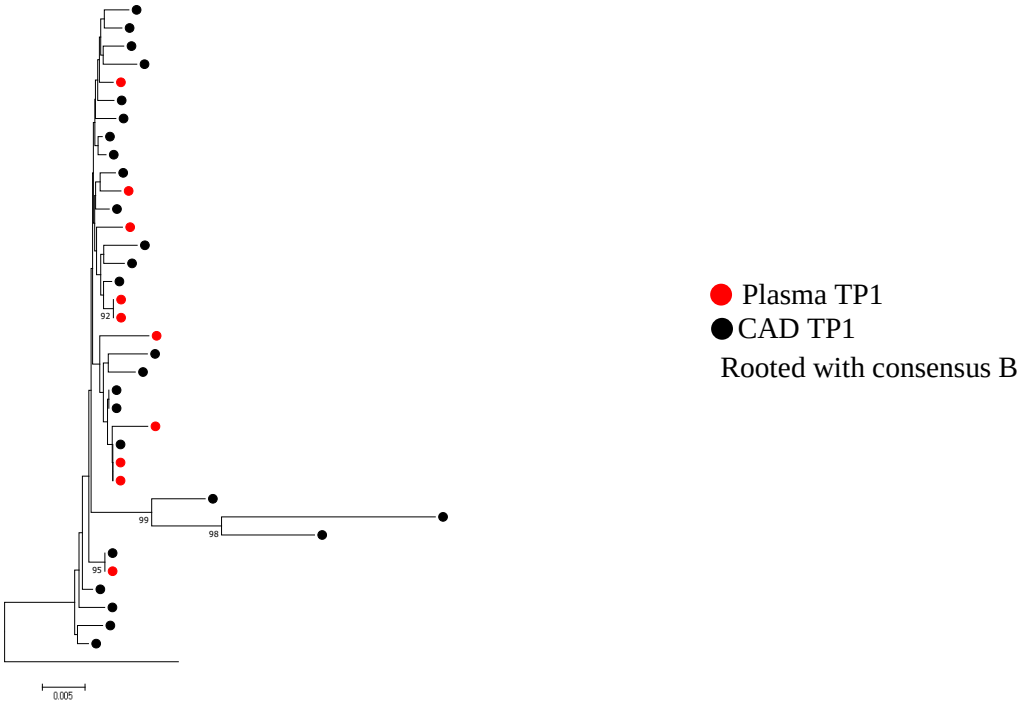


Figure 4.11: Neighbour joining p-distance phylogenetic tree of p6PrRT nucleotide sequences generated from TP1 plasma and cell associated DNA (CAD) samples of patient D15. Sequence origins are represented with symbols as indicated by the legend. The tree was constructed with 1000 bootstraps and rooted to Consensus B (MN919177.1).

4.3.6 Preliminary HFH data of T13 sub-clone 1

Preliminary ISA data of MDA products from patient T13 sub-clone 1 indicate the presence of the provirus in an intergenic region of Chr1 (position 146,354,082) in the positive orientation. No clear amplicon, without background, was obtained from the combined annealing and extension PCR approach. Separating the annealing and extension steps and increasing the extension temperature to 68^{o}C resulted in a reduction in background amplification (Figure 4.12). A single amplicon of the expected size (9.1 kb) from the 5' ~6kb-p6PrRT approach was obtained and could be visualised on gel electrophoreses (Figure 4.12). However, no nucleotide sequence could be obtained for this HFH amplicon.

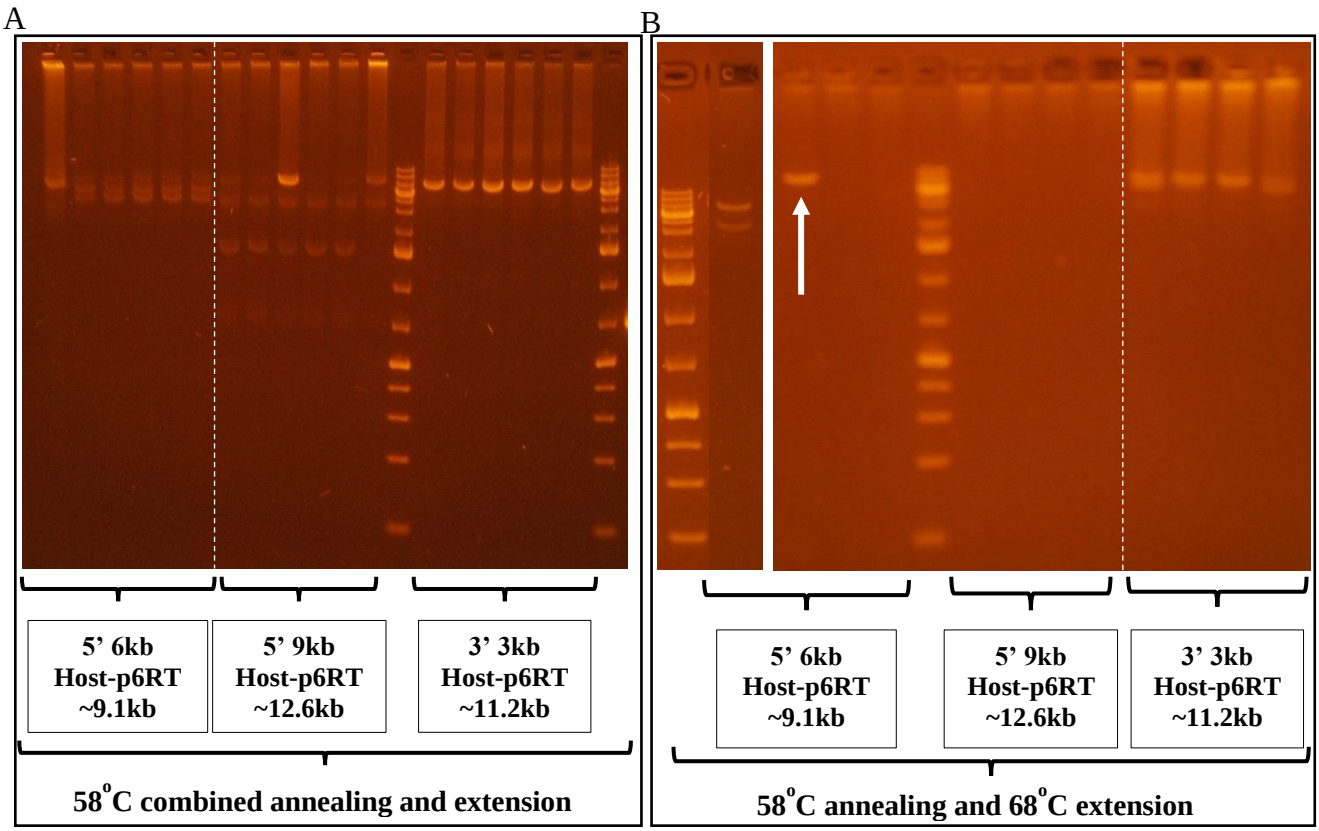


Figure 4.12: Gel electrophoresis images of sub-clone 1 provirus Chr1_146354082 specific HFH. A) Combined annealing and extension step on 5' and 3' HFH halves. B) Separate annealing and extension steps on 5' and 3' HFH halves. Amplicon of expected size (9.1kb) is indicated by a white arrow. Lanes contain replicates of the indicated targets.

4.4 Discussion

It is widely accepted that the major proviral reservoir of infectious HIV-1 is harboured in resting memory T-cells (Chun 2013; Siliciano & Siliciano 2014), presenting a significant barrier to cure. Clonally expanded $CD4^+$ T-cells have been implicated in harbouring infectious virus and shedding virus into the periphery for several years in a patient on cART (Simonetti et al. 2016). The frequency of clonally expanded cells harbouring infectious HIV is unknown, it is however an additional barrier to achieving cure. Moreover, when clonally expaned cells produce viraemia detectable by commercial assays it may complicate clinical care as this may be attributed to treatment adherence or drug resisance and result in unnecessary and ineffective treatment substitutions.

Therefore, in this study a cohort of patients with prolonged unexplained HIV viraemia was enrolled to investigate the clonally expanded proviral reservoir. This part of the study consisted

of three patients presenting with clinical data consistent with expected clonal viraemia: persistent non-suppressible viraemia and no treatment relevant drug resistance, as indicated by p6PrRT single genome sequences and routine drug resistance testing, which included integrase, while maintaining good adherence on cART. Genetic material from plasma and PBMCs of these patients were investigated with molecular assays targeting the semi-conserved p6PrRT region. Virus recovery was attempted from $CD4^+$ T-cells of patients where multiple identical p6PrRT nucleotide sequences were obtained from plasma and PBMCs. HIV p24 antigen positive QVOA supernatants were investigated for infectious virus with identical p6PrRT nucleotide sequences to plasma virus and PBMC provirus. The integration site data of infectious suspected clones were sought by screening amplified genomic MDA reactions for identical p6PrRT sequences. Host genome amplified reactions containing the suspected proviral clones were used for ISA to identify the integration sites and thereby confirm clonality. With confirmed integration site data and p6PrRT nucleotide sequences, the sequence of the entire provirus and flanking Hg regions can be determined using HFH.

Although all nucleotide sequences from plasma of patient L14 were identical or nearly identical, a strong indication of clonal viraemia, no identical or closely related sequences could be obtained from PBMCs. Clonal viraemia maintained by expanded cell clones could therefore not be further investigate for this patient. The spike in p6PrRT identical HIV viraemia lasting 1.8 months, as indicated by the clinical data, could be the result of the activation and HIV expression of a minor cell clone however, this remains unconfirmed. This period of detectable viraemia spontaneously resolved with HIV re-suppression.

The unstable viral load of patient D15 with the latest viral load spike lasting for over 12 months was investigated for clonal viraemia. Initial p6PrRT plasma sequencing indicated a low level of identical sequences. Further p6PrRT analyses indicated significant variation of the plasma and cell associated DNA population, with no substantiated clonal viraemia. Due to these initial results further investigation of clonally expanded infectious proviruses were abandoned for this patient.

Initial p6PrRT SGS on plasma and cell associated DNA of patient T13 TP1 indicated multiple clusters of identical plasma sequences and one plasma cluster with identical sequences from cell associated DNA. This patient was therefore selected for QVOA to investigate the inducible infectious proviral reservoir. Almost all viral outgrowth replicates, except for two, tested HIV p24 antigen positive after 14 and 21 days, with a calculated IUPM of 36.1. A total of 74 p6PrRT

SGS sequences were obtained from QVOA supernatants of a single positive replicate of day 7 and all positive day 14 replicates. Nucleotide sequences of the p6PrRT region were also obtained from TP2 plasma and cell associated DNA. Phylogenetic analysis of these sequences identified two major clusters of identical sequences obtained from all sources, termed clone 1 and 2. A cluster of identical plasma, cell associated DNA and QVOA sequences, removed from clone 1 by a SNP confirmed in all sequences, was identified and termed sub-clone 1. Sub-clone 1 formation could be the result of either of two events. Firstly, an almost identical HIV strain to the clone 1 strain could be responsible for a separate integration event, followed by some level of clonal expansion. Alternatively, the polymorphism could have occurred during clonal proliferation of cell clone 1, resulting in the formation of a seemingly separate clone. This phenomenon could be clarified by comparing the full-length HIV sequences of both proviral clusters, however, integration site data is required to confirm the true origin of sub-clone 1. In addition, a third separate phylogenetic cluster, consisting of plasma and cell associated DNA sequences, were identified and termed clone 3. Given the detection of multiple identical cell associated DNA sequences but no related sequences from QVOA supernatants, the replication competence of clone 3 is brought into question, but not ruled out. Identical p6PrRT sequences from MDA products were obtained for clones 1, 2, 3 and sub-clone 1. The integration site of sub-clone 1 was determined to be in an intergenic region of Chr1 at position 146354082, with the provirus in the positive orientation. Observations of the Hg flanking this position revealed a high degree of repetitive sequences. Reliable primers for HFH targeting this position were therefore designed over 6 kb up-stream and over 3 kb down-stream of the IS, requiring amplicons of at least 9 kb and ~12 kb respectively for successful HFH. No amplicons of interest could be generated with the standard HFH PCR cycle conditions. The HFH protocol was therefore optimised by separation of the annealing and extension steps, resulting in a reduction of background amplification and a single 5’ amplicon of interest. Unfortunately, no nucleotide sequence data could be generated for the HFH amplicon of interest. Although the clonality of the T13 identical sequence clusters remain to be confirmed by ISA and HFH, the high number of identical p6PrRT sequences from numerous sources is highly suggestive of detectable viraemia sustained by clonally expanded cells harbouring infectious provirus.

Results from this project were shared as part of the non-suppressible viraemia cohort study at CROI 2019 (Halvas et al. 2019). This study indicated that intact proviruses are integrated into introns in intergenic regions in either orientation. It offers further evidence that non-adherence and drug resistance are not the only causes of clinically detectable and sustained viraemia.

Infectious proviral clones present new implications to achieve cure as they may cause rapid rebound of viraemia after cART interruption. In order to achieve HIV cure, expanded infectious proviral clones along with the rest of the reservoir will need to be eliminated or suppressed successfully, this is further complicated by the potential regrowth of proviral clones. Further investigations regarding the mechanism of clonal escape are required.

Chapter 5

Non-suppressible HIV viraemia in South African adult patients on long term cART

5.1 Introduction

In keeping with observations made in chapter 4, South African adult patients presenting with non-suppressible HIV subtype C viraemia, as evident from sustained detectable viral loads, which could not be explained by detected drug resistance or insufficient treatment adherence, were referred to our lab for further investigation. Molecular assays used for clonal viraemia characterisation in chapter 4 were applied to these patient samples to investigate the nature of the sustained viraemia.

Findings from patient 1 discussed in this chapter have been shared at the 28th International Workshop on HIV Drug Resistance and Treatment Strategies (Botha et al. 2019) and at CROI 2020 (Botha et al. 2020).

5.2 Materials and methods

5.2.1 Specimen selection

Two HIV-1 positive South African adult patients on long term cART presenting with non-suppressible viraemia were referred to our lab by HIV care providers. The selection criteria for this study was a detectable HIV viral load sustained for over 6 months with no expected non-adherence. Plasma and EDTA whole blood specimens over multiple time points were collected for these patients. Ethics approval for this study was obtained from the University of the Witwatersrand from where the patients were enrolled (Clearance certificate: M190294).

5.2.2 PBMC and nucleic acid isolation

PBMC, DNA and plasma RNA isolation proceeded as described previously in chapter 2 sections 2.2.2 and 2.2.3 and chapter 4 section 4.2.4.

Additionally, cell associated RNA was isolated from PBMCs as follows: selected PBMC specimens were split to consist of ~1.25 million cells per isolation reaction. PBMCs were centrifuged in a Prism R refrigerated microcentrifuge (Labnet) at 500 rcf for 15 min at 4°C. After removal of the supernatant, 600 µl custom lysis-buffer consisting of 3.6 M GuSCN, 0.13 M DTT, 0.67 mg/mL Glycogen, 10 mg/mL N-Lauroylsarcosine, 10 mg/mL Sodium Citrate and 132 µl nuclease free water were added. Cell pellets were resuspended by pulse vortex mixing for 4 seconds and incubated at 25°C for 10 min. Reactions were inverted at least 30 times after the addition of 600 µl 100% Isopropanol and centrifuged at 17 200 rcf for 20 min at 10°C. The supernatant was carefully removed without disturbing the nucleic acid pellet followed by the addition of 1mL ice cold 70% ethanol. Reactions were inverted and centrifuged at 17 200 rcf for 20 min at 10°C followed by the complete removal of all ethanol by centrifuging reactions twice at 17 200 rcf for 1 min and removing any supernatant. The nucleic acid pellets were dried for about 5 – 10 min at 25°C followed by resuspension in 8ul Purelink DNase buffer (Invitrogen), 62ul nuclease free water and 10ul Purelink DNase (Invitrogen). The nucleic acid was incubated in DNase mix for 15 min at 25°C followed by the addition of 200 µl resuspension buffer (5.7 M GuSCN, 50 mM Tris-HCl pH 7.5 and 1 mM EDTA) and 250 µl 100% Isopropanol. Mixtures were inverted and centrifuged at 17 200 rcf for 20 min at 10°C. The supernatant was carefully removed without disturbing the RNA pellet followed by the addition of 1mL 70% ethanol and centrifugation at 17 200 rcf for 20 min at 10°C. Supernatant was removed and RNA pellets were stored in 1mL 70% ethanol at -80°C until use.

Prior to the use of cell associated RNA, pellets stored at -80°C in 70% ethanol were thawed and centrifuged 17 200 rcf for 20 min at 10°C followed by the removal of supernatant. Centrifugation was repeated twice for 1 min for the complete removal of all supernatant. The RNA pellets were dried for 5 – 10 min at 25°C followed by resuspension in 24 µl cold 5 mM Tris-HCl (pH 8.0) (Sigma-Aldrich). The resuspended RNA was used immediately.

5.2.3 Molecular workup (cDNA synthesis, p6PrRT SGS, phylogenetic analysis)

cDNA synthesis was performed as described in chapter 4 section 4.2.5 on plasma RNA and cell associated RNA in 50 µl reactions. SGS targeting the p6PrRT region was performed as described in chapter 4 section 4.2.6 on PBMC DNA, DNA of cell debris (obtained from low-speed centrifugation of plasma), stored buffy coat DNA and cDNA (plasma virus pellet RNA and PBMC RNA) based on availability from each patient. HIV-1 subtype C specific p6PrRT

primers (Table 5.1) were designed and used for South African patients. For cell associated RNA-cDNA p6PrRT, successful DNase was confirmed by including two no-cDNA synthesis control reactions.

Table 5.1: Primers used for HIV-1 subtype C p6-Protease-Reverse transcriptase single genome sequencing amplification.

Primer name	Orientation	Sequence (5'-3')	Location in HXB2 genome
Pre-nest F	Forward	GAT GAC AGC ATG TCA GGG AG	1827 – 1846
Pre-nest R	Reverse	CTA T**Y**A AGT CTT TTG ATG GGT CAT AA	3503 – 3528
Nest-F	Forward	GAG TGT TGG CTG AGG CAA TGA G	1871 – 1892
Nest-R	Reverse	CAG TTA GTC GTA CTA TGT CTG TTA GTG CTT	3409 – 3438

PCR products analysis, nucleotide sequencing with subtype C specific primers (Table 5.2) and phylogenetic analysis proceeded as described in chapter 4. Nucleotide sequences obtained in this study were aligned with consensus C sequence (KR820366.1) for accurate phylogenetic analysis.

Table 5.2: Primers used for overlapping nucleotide Sanger sequencing of HIV-1 subtype C p6-Protease-Reverse transcriptase single genome sequencing amplicons

Primer name	Orientation	Sequence (5'-3')	Location in HXB2 genome
p6PrRT-A	Forward	TGT TGG AAA TGT GGA AAG GAA GGA C	2026 – 2050
p6PrRT-B	Forward	ATG GCC CAA AAG TTA AAC AAT GGC	2599 – 2622
p6PrRT-C	Reverse	**Y**TC TTC TGT CAA TGG CCA TTG TTT AAC	2610 – 2636
p6PrRT-D	Reverse	TTG CCC AGT TTA ATT TTC CCA CTA A	3327 – 3351

5.2.4 Near full-length amplification

Near full-length (NFL) amplification of HIV-1 was performed on PBMC DNA of selected patients. Similar to p6PrRT SGS (section 4.2.6), the NFL protocol consists of a nested amplification approach and makes use of Poisson probability to determine the single template endpoint dilution, followed by amplification at the determined endpoint.

Amplification was performed with Ranger mix (Bioline meridian bioscience) and each reaction (10 µl) consisted of: 1x Ranger mix, 0.4 µM forward and reverse primers (Table 5.3) and nuclease free water (Promega Corp.). Pre-nested amplification conditions consisted of an initial denaturing step (95ºC, 2 min) followed by 30 cycles of denaturation (98ºC, 10 sec), annealing and extension in one step (61ºC, 10 min) followed by a final extension step (68ºC, 15 min). The pre-nested PCR products were diluted with 83 µl 5 mM Tris-HCl (pH 8.0) (Sigma-Aldrich) per reaction and 2 µl diluted amplicon was plate stamped into the nested PCR plate. Nested amplification conditions consisted of an initial denaturing step (95ºC, 2 min) followed by 30 cycles of denaturation (98ºC, 10 sec), annealing and extension in one step (61.5ºC, 10 min) and a final extension step (68ºC, 15 min).

Table 5.3: Primers used for HIV-1 subtype C near full-length single genome amplification. *(Li et al. 2010)

Primer name	**Orientation**	**Sequence (5'-3')**	**Location in HXB2 genome**
*Li_Out_F	Forward	AAA TCT CTA GCA GTG GCG CCC GAA CAG	623 – 649
*Li_Out_R	Reverse	TGA GGG ATC TCT AGT TAC CAG AGT C	9662 – 9686
NFL_In_F	Forward	CCG AAC AGG GAC BHG AAA GCG AA	642 – 664
*Li_In_R	Reverse	GCA CTC AAG GCA AGC TTT ATT GAG GCT TA	9604 – 9632

Successful amplification reactions were determined followed by gel-electrophoresis to determine amplicon size, as discussed in chapter 2 section 2.2.8. Amplicons of ~9 kb was selected for p6PrRT sequencing (as described above). The packaging signal, p17 and partial p24 nucleotide sequences of NFL amplicons were determined by Sanger sequencing (as described) with the NFL_In_F nested amplification primer (Table 5.3). Theses nucleotide sequences were aligned with and analysed against consensus C sequence (KR820366.1).

5.3 Results

5.3.1 Specimen selection

Two patients presenting with unexplained non-suppressible HIV viraemia were referred to us for further investigation. The known viral load history is indicated on log scale data plots with samples collected for analyses. A brief clinical history is also provided for each patient. Clinical data from patient 1 indicates an unstable viral load that was finally suppressed after adapting the treatment to AZT/FTC/TDF/LPV/R (Figure 5.1). However, after a brief period of viral suppression, HIV-1 was once again detectable at 4230 copies/mL and sustained for over 15 months. Drug resistance population sequencing performed on samples 1 and 2 yielded hypermutated sequences only. Plasma virus component and plasma cell debris from samples 1 and 2, along with plasma virus component and PBMCs from sample 4 were investigated to characterise the non-suppressible viraemia in patient 1.

Age	46
Sex	Female
Diagnosis	2013
Viral Load (VL) copies/mL(c/mL)	S1:(Plasma), 4230 c/mL S2:(Plasma), 3880 c/mL S4:(EDTA), 2388 c/mL
Duration of Detectable Viremia	>15,3 months
Current ART Regimen	AZT/FTC/TDF/LPV/R

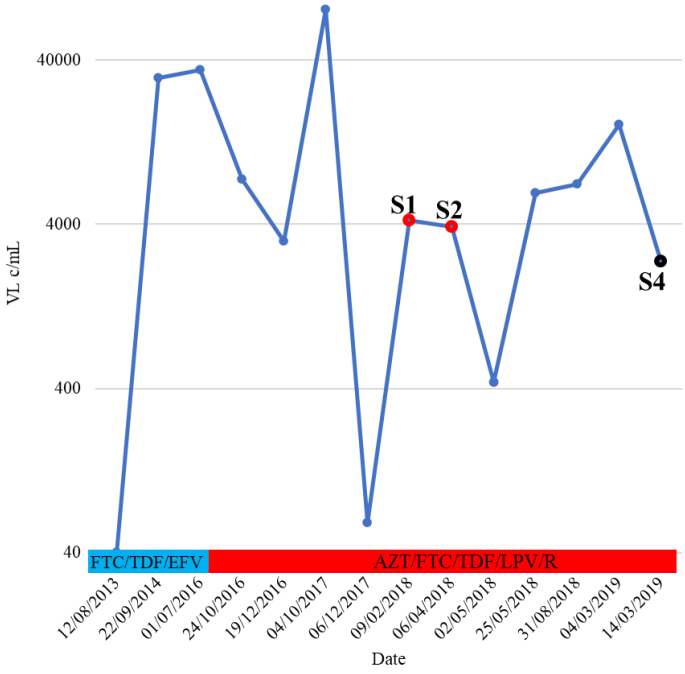


Figure 5.1: Clinical overview and viral load (VL) data plot for patient 1. Collected samples (S) 1, 2 and 4 are indicated with the respective VL.

Patient 2 clinical data with known viral loads of the study period indicates a very slow but steady decline of detectable HIV-1 (Figure 5.2). No viral load data prior to sample 1 is available for this patient. Initial poor adherence to treatment was reported for 2012 – 2015

(FTC/TDF/EFV), followed by treatment adjustment to 3TC/AZT/LPV/R with good adherence from 2015 – early 2018. However, side effects led to treatment adjustment to 3TC/AZT/DTG, the treatment combination for the duration of study. After initial viral decline to 6500 copies/mL the viral load returned to over 10 000 copies/mL and non-adherence was suspected. Although freely admitting poor adherence during 2012-2015 the patient claimed perfect adherence since early 2018. The patient missed no clinic appointments and pill refill records indicated towards adherence. The patient further provided video evidence of taking his tablets every evening for over two months, to the satisfaction of the HIV care provider. Finally, adherence and drug absorption were confirmed with 3TC and DTG drug level testing on four occasions. Although unconventional, these findings prove adherence to treatment. Drug resistance testing indicated high-level resistance to 3TC, resulting in increased susceptibility to AZT over the study period (Liu & Shafer 2006). *Integrase* sequencing indicated susceptibility to DTG from January 2018 until September 2019 (21 months), after which resistance to DTG (N155H and R263K) was detected (January 2020). Even though DTG resistance was detected, the viral load continued to decline (Figure 5.2). Samples 4 (plasma and buffy coat), 5 (plasma) and 6 (plasma and PBMCs) were investigated to characterise the source of sustained viraemia in patient 2. Samples 1, 2 and 3 were excluded due to elevated viral loads.

Age	43
Sex	Male
Diagnosis	2012
Viral Load (VL) copies/mL(c/mL)	S1:(Plasma), 149 000 c/mL S2:(Plasma), 6500 c/mL S3:(Plasma), 3780 c/mL S4:(Plasma and Buffy Coat), 1420 c/mL S5:(Plasma), 1290 c/mL S6:(EDTA), 1128 c/mL
Duration of Detectable Viremia	>24 months
Current ART Regimen	3TC/AZT/DTG

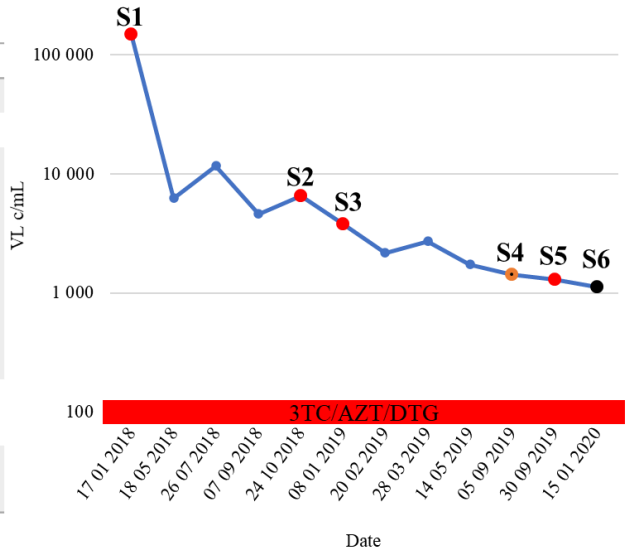


Figure 5.2: Clinical overview and viral load (VL) data plot for patient 2. Collected samples (S) 1, 2, 3, 4, 5 and 6 are indicated with the respective VL.

5.3.2. p6PrRT SGS

For patient 1, several p6PrRT nucleotide sequences were generated from sample 1 (plasma virus component and plasma cell debris), sample 2 (plasma cell debris) and sample 4 (plasma virus component, plasma cell debris, cell associated DNA and RNA) (Table 5.4). Notably sample 4 p6PrRT SGS endpoint for cell associated DNA was 1:27 compared to cell associated RNA endpoint 1:2187, indicating a much higher HIV-1 RNA than HIV-1 DNA concentration.

Table 5.4: Patient 1 p6PrRT nucleotide sequences obtained from various sources for each sample. *(Liu & Shafer 2006)

Sample	Sequence source	Number of sequences obtained	*Mutations detected
Sample 1 Plasma (09/02/2018)	Population drug resistance	1 consensus	Hypermutated
	Plasma cell debris	4	Hypermutated
	Plasma virus component	3	Hypermutated
Sample 2 Plasma (06/04/2018)	Population drug resistance	1 consensus	Hypermutated
	Plasma cell debris	2	Hypermutated
	Plasma virus component	PCR negative	PCR negative
Sample 4 Plasma (14/03/2019)	Population drug resistance	1 consensus	D67N
	Plasma cell debris	1	Hypermutated
	Plasma virus component	4	D67N
Sample 4 PBMC (14/03/2019)	DNA	10	1/10 – D67N
			9/10 – Hypermutated
	RNA	11	Hypermutated

Patient 2 p6PrRT SGS yielded nucleotide sequences from samples 4 (plasma RNA and buffy coat DNA), 5 (plasma RNA) and 6 (plasma RNA, cell associated DNA and RNA) (Table 5.5).

Table 5.5: Patient 2 p6PrRT nucleotide sequences obtained from various sources for each sample. *(Liu & Shafer 2006)

Sample	Sequence source	Number of sequences obtained	*Drug resistance mutations		
			NRTI	NNRTI	Integrase
Sample 4 (05/09/2019)	Plasma RNA	8	A62V, K65R, M184V	L100I, K103N	None
	Buffy coat DNA	24	None		
Sample 5 (30/09/2019)	Plasma RNA	12	A62V, K65R, M184V	L100I, K103N	None
Sample 6 (15/01/2020)	Plasma RNA	22	A62V, K65R, M184V	L100I, K103N	N155H, R263K
	PBMC DNA	24	None		
	PBMC RNA	21	None	K103N	None

5.3.3 NFL amplification and nucleotide sequencing

In total, 10 NFL amplicons were generated from sample 4 cell associated DNA of patient 1 (Figure 5.3). Products from all amplicon positive wells were analysed on gel electrophoresis (Figure 5.3). Nucleotide sequences covering the p6PrRT region and the packaging signal, p17 and partial p24 were successfully generated for all 10 NFL amplicons.

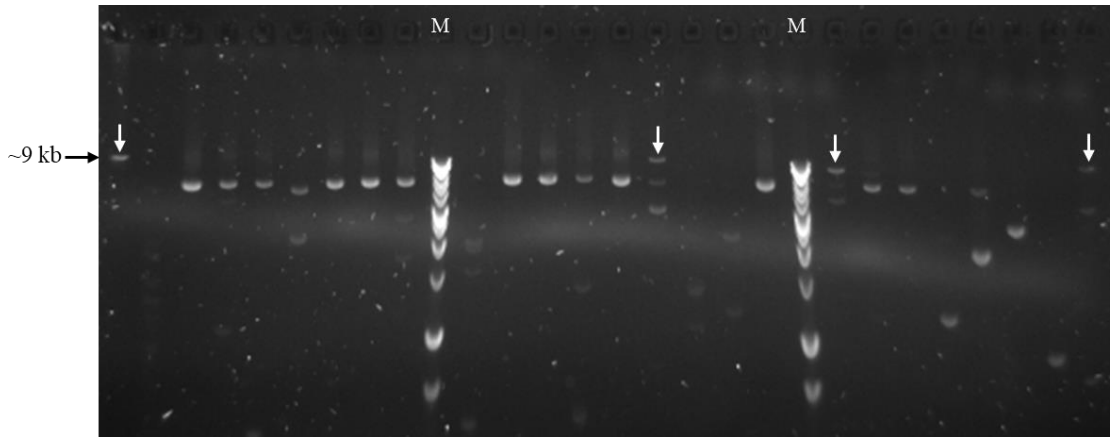


Figure 5.3: Gel electrophoresis of amplicons detected after near full-length (NFL) amplification of patient 1 sample 4 PBMC DNA. The expected NFL amplicon size (~9 kb) is indicated (black arrow) along with positive NFL amplicons (white arrows).

Five possible NFL amplicons were obtained from patient 2 sample 6 cell associated DNA (Figure 5.4). Products from all amplicon positive wells were analysed on gel electrophoresis (Figure 5.4). The p6PrRT region of only 1/5 amplicons could be obtained.

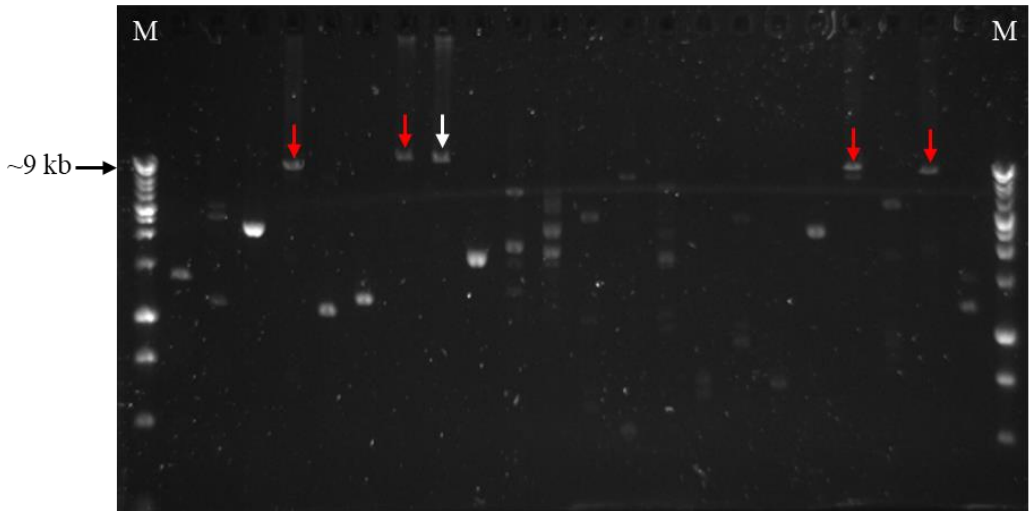


Figure 5.4: Gel electrophoresis of amplicons detected after near full-length (NFL) amplification of patient 2 sample 6 PBMC DNA. The expected NFL amplicon size (~9 kb) is indicated (black arrow), along with possible NFL amplicons (red arrows) and one confirmed NFL product (white arrow).

5.3.4 Phylogenetic and nucleotide sequence analyses

Neighbour-joining phylogenetic trees constructed of p6PrRT nucleotide sequences and p6PrRT sequence data from NFL products from patients 1 and 2 with evolutionary distance calculated using the p-distance model were analysed for monotypic clusters that, indicative of proviral clones. A single identical nucleotide sequence cluster consisting of all p6PrRT sequences obtained from samples 1 and 2, along with 9/10 cell associated DNA, the plasma cell debris, all cell associated RNA and NFL sequences obtained from sample 4 was phylogenetically identified (Figure 5.5). All these sequences had APOBEC3G G to A signature mutations, as identified by the Stanford HIVDB (Liu & Shafer 2006) (Table 5.4). All sample 4 plasma virus component sequences along with 1/10 cell associated DNA and the consensus drug resistance sequences, formed a separate variable phylogenetic cluster. A D67N mutation was identified in all these sequences (Table 5.4). Analysis of the identical packaging signal, p17 and partial p24 nucleotide sequences obtained from 9/10 NFL amplicons, against consensus C sequence (KR820366.1) revealed a G → A SNP (HXB2 nucleotide position 1425), resulting in a stop codon within the p24 gene (HXB2 amino acid position 212) (Figure 5.6).

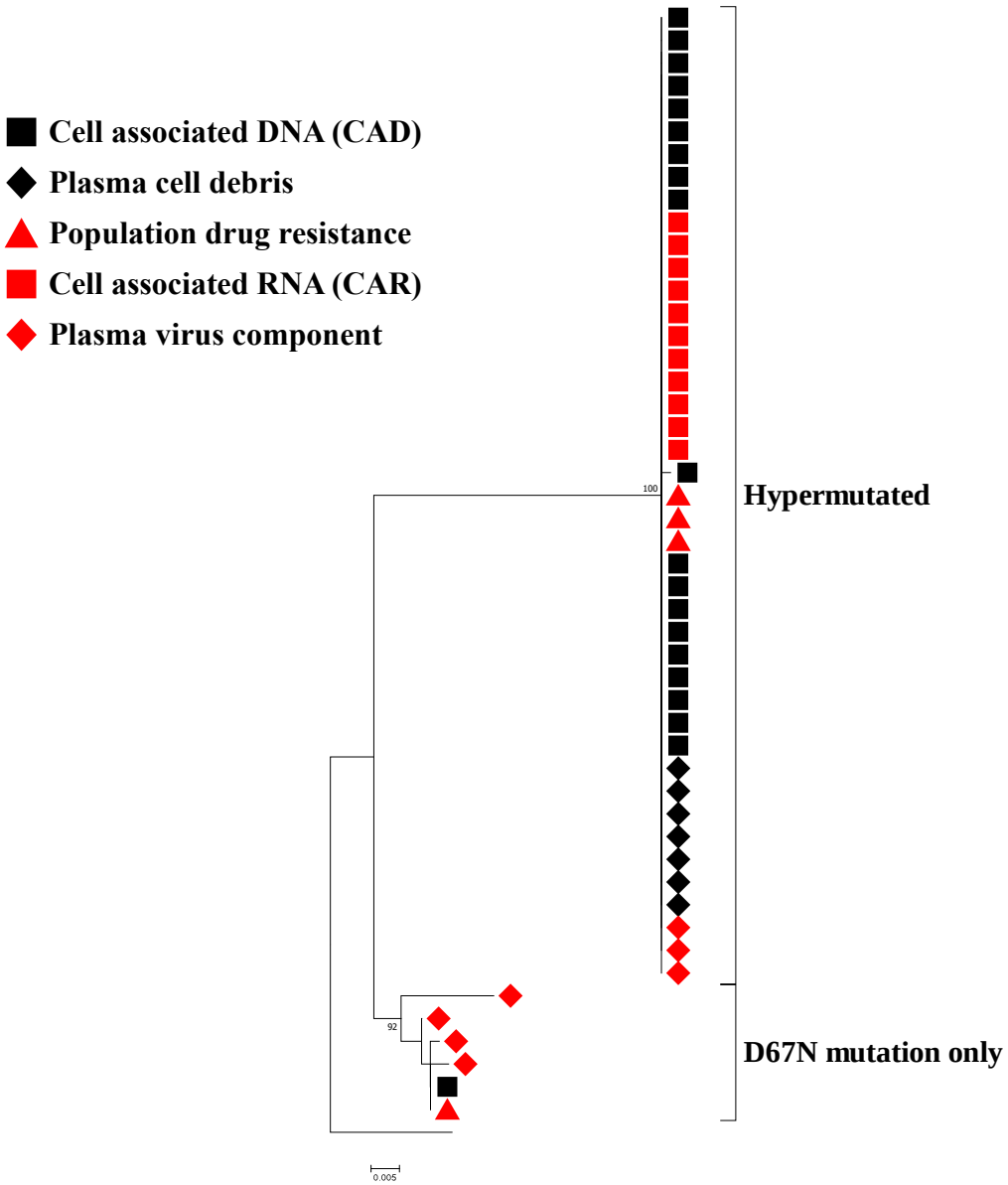


Figure 5.5: Neighbour joining p-distance phylogenetic tree of patient 1 p6PrRT nucleotide sequences generated from plasma virus component and plasma cell debris from sample 1 and plasma cell debris from sample 2, along with sequences from plasma virus component, plasma cell debris, CAD, CAR and NFL from sample 4. The identical APOBEC hypermutated nucleotide sequence cluster is indicated along with the sample 4 sequence cluster consisting of D67N mutation sequences. Sequence origins are represented with various symbols as indicated by the legend. The tree was constructed with 1000 bootstraps and rooted to Consensus C (KR820366.1).

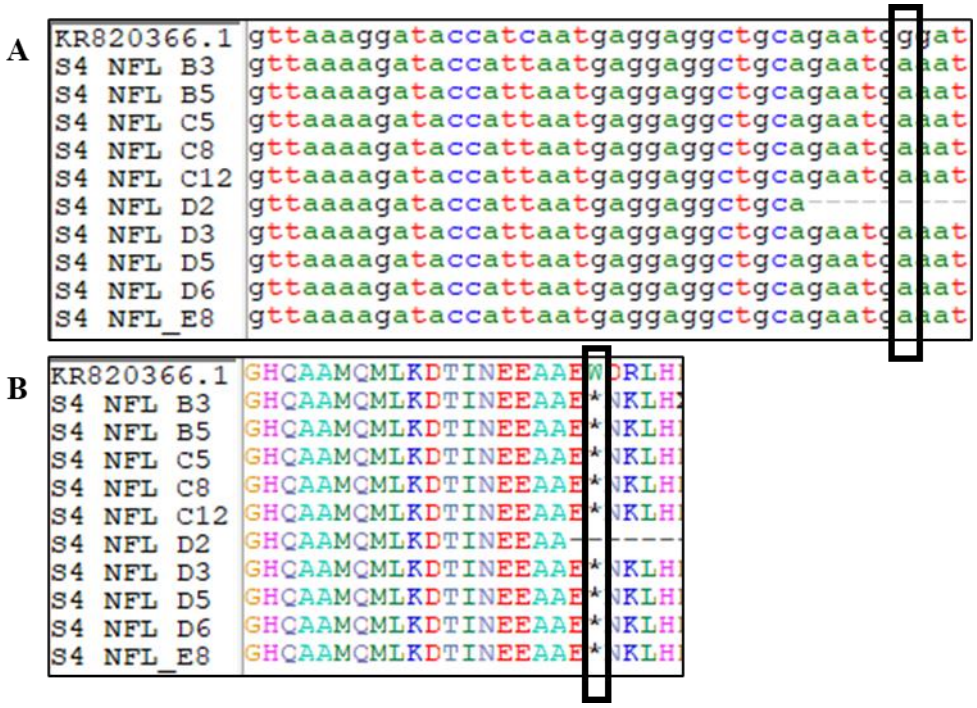


Figure 5.6: Patient 1 sample 4 PBMC DNA partial p24 sequence data obtained from NFL amplicons. A) Nucleotide sequences indicating the G → A SNP (HXB2 nucleotide position 1425). B) Transcribed amino acid sequences indicating the stop codon within the p24 gene (HXB2 amino acid position 212). The NFL sequences were analysed against consensus C sequence (KR820366.1).

Several identical p6PrRT nucleotide sequence clusters were observed after phylogenetic analysis for patient 2 (Figure 5.7). A monotypic sequence cluster consisting of sample 4 buffy coat and sample 6 cell associated DNA sequences, not related to any cell associated RNA or plasma sequences, was identified. Most sequences obtained from sample 6 cell associated RNA clustered together in two identical sequence clusters however, no cell associated RNA sequences identical to any of the identical plasma or cell associated DNA monotypic clusters were identified. Plasma virus RNA sequences yielded two identical clusters termed plasma cluster 1 and 2. Plasma cluster 1 consisted of sample 4, 5 and 6 sequences, whereas plasma cluster 2 consisted of only sample 4 and 6 sequences. No PBMC DNA or buffy coat DNA sequences were identical to any of the plasma clusters.

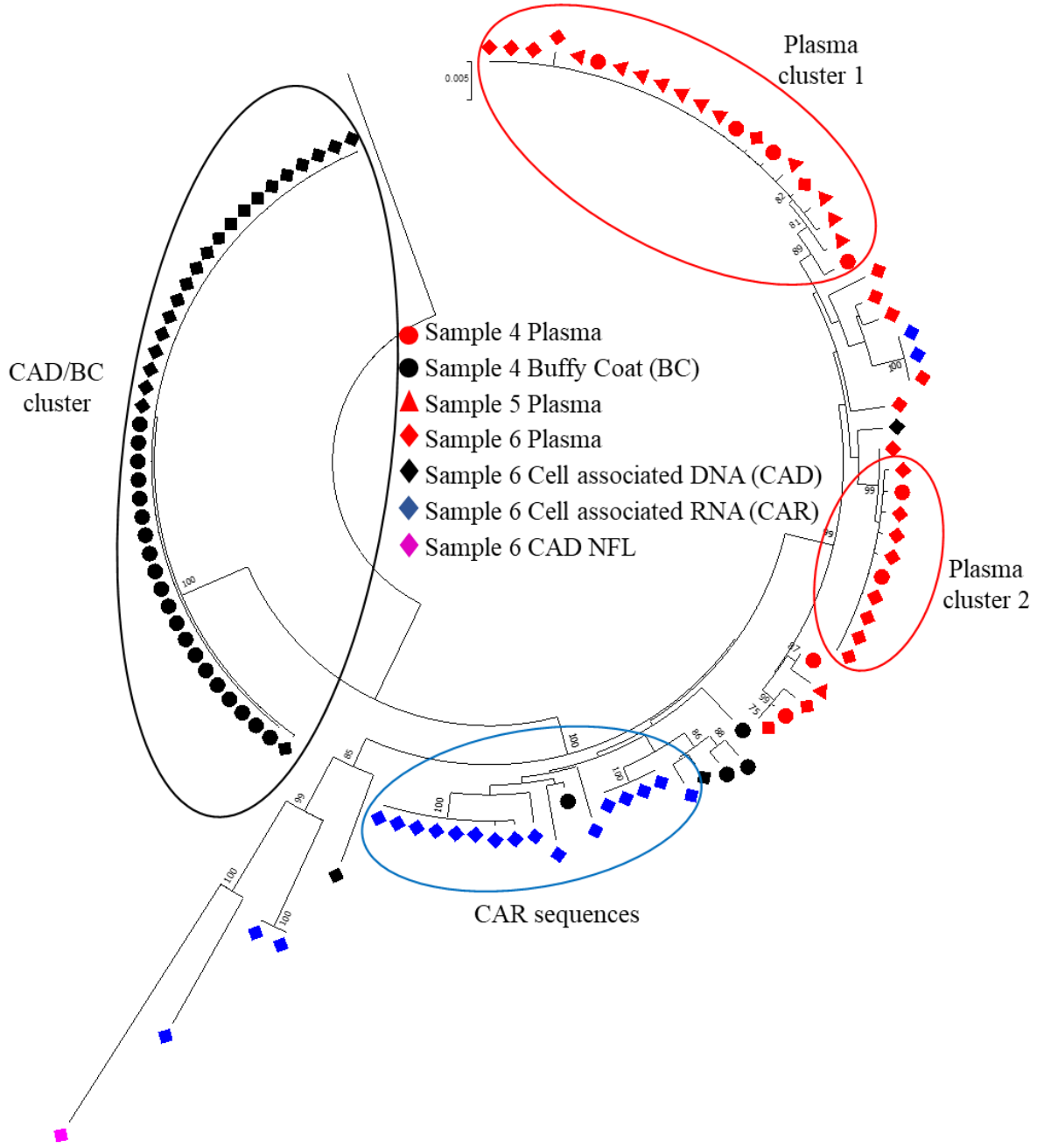


Figure 5.7: Neighbour joining p-distance phylogenetic tree of patient 2 p6PrRT nucleotide sequences generated from samples 4, 5 and 6 (plasma virus), sample 4 (buffy coat (BC)) and sample 6 (cell associated DNA (CAD), cell associated RNA (CAR) and near full-length amplification on CAD. Identical sequence clusters, CAD/BC monotypic cluster (sample 4 BC and sample 6 CAD), CAR sequences (two separate clusters of identical sequences), plasma cluster 1 (samples 4, 5 and 6 plasma virus sequences) and plasma cluster 2 (samples 4 and 6 plasma virus sequences) are circled in corresponding colours. Sequence origins are represented with various symbols as indicated by the legend. The tree was constructed with 1000 bootstraps and rooted to Consensus C (KR820366.1).

5.4 Discussion

Findings from Simonetti et al. 2016 and Halvas et al. 2019 discussed in chapter 4 indicated that patients may present with persistent vireamia, as detected by commercial viral load assays, over several months, due to the release of viruses from expanded cell clones in the absence of ongoing viral replication. Based on these studies the incidence of clonal viraemia could be more common than previously thought. Patients presenting with non-suppressible viraemia therefore require further investigation to characterise the source of the sustained viraemia, to avoid being mischaracterised as having virologic failure.

In this study two South African adult patients presenting with non-suppressible viraemia for over 15 months and 24 months respectively were referred to our lab for further investigation. No treatment relevant drug resistance was observed, and the patients were adherent to cART. Molecular methods targeting the p6PrRT region of the HIV genome, as discussed in chapter 4, were used to characterise the possible clonal viraemia on various longitudinal samples from these patients. As these patients were referred from clinical care, samples were a combination of residual samples from routine tests (viral load and HIV drug resistance) and additional testing after informed consent for this case series. Sample volumes were limited for these longitudinal plasma samples. In one case (patient 2 sample 4) the only available cells were from a buffy coat sample. Although this is a suboptimal sample due to poor preservation of cells, several p6PrRT nucleotide sequences could be obtained from buffy coat DNA. For the proper investigation of HIV-1 cell associated DNA and RNA, a whole blood EDTA sample for PBMC isolation was requested at enrolment into this study. In addition to the molecular methods of chapter 4, cell associated RNA was extracted, amplified and p6PrRT sequenced to investigate whether the cell associated DNA was transcribed and NFL amplicons to investigate the intactness of the possible clones.

Phylogenetic analysis of patient 1 p6PrRT SGS revealed a major identical cluster consisting of products from all sources expect sample 2 and 4 plasma virus pellets. These sequences were present in the virus pellet of sample 1, PBMCs (cell associated DNA and RNA) and could be separated from plasma by low speed pre-centrifugation (cell debris) of samples 1, 2 and 4. The hypermutated sequences had a critical mutation in *gag* p24 which would prevent packaging. The sample 4 PBMC p6PrRT SGS endpoints of DNA (1:27) and RNA (1:2187) suggest that this APOBEC hypermutated strain is highly transcribed. The partial *gag* sequences from NFL amplicons revealed a SNP resulting in a stop codon in the p24 gene. If translated this would

result in truncated p24 capsid proteins, which in turn would preclude virion formation. Taken together these findings suggest that the source of the non-suppressible viraemia appear to be a large population of cells with identical hypermutated provirus (a possible $CD4^+$ T-cell clone) that, through cytolysis, are releasing cellular nucleic acid as cellular fragments of various sizes into the plasma component. This is the first reported case of false virologic failure of a patient on long-term cART, as a result of defective proviral DNA and RNA released into plasma. This case was detected due to the hypermutated nature of the HIV viraemia, however the release of cellular material into plasma may be an underrepresented cause of false treatment failure and could impact patient care by unnecessarily changing drug regimens. Moreover, it has recently been reported that delayed processing of plasma before viral load testing could result in false high HIV viral loads due to the release of cellular material into the plasma. For that reason, pre-centrifugation of samples is recommended before viral load testing (Hardie et al. 2019).

Patient 2 phylogenetic analysis revealed several monotypic sequence clusters of PBMC and buffy coat DNA, cell associated RNA and plasma respectively. Similar to patient L14 of chapter 4 no identical cell associated DNA and plasma sequences could be obtained. However, unlike patient L14 nearly all cell associated DNA sequences were identical to one another. These observations suggest that some level of proviral clonality might be present in patient 2. The proviral strain responsible for the apparent clonal viraemia could not be identified from 48 p6PrRT SGS cell associated DNA/buffy coat sequences. The viraemia could therefore be the result of cellular activation of a rare provirus, which may be underrepresented in peripheral blood and possibly more prevalent in particular tissues (Wong & Yukl 2016). This does not exclude clonality of this strain. The large identical cell associated DNA cluster hints at the existence of a predominant cell clone, with no evidence of trasncription. The identical cell associated RNA sequences, that are not matching the DNA cluster, suggest that additional proviral transcription is occurring in this patient, however no evidence was observed to suggest that this is responsible for the sustained viraemia. Although the identical plasma nucleotide sequences suggest that clonal viraemia could be the cause of the uncommon very slow viral decay, no matching cellular DNA sequences could be found. However, after a period of almost 2 years without treatment relevant HIV drug resistance and the presence of monotypic sequences (at time points S4 and S5), significant treatment relevant high level DTG drug resistance was detected. As this patient did not have prior integrase strand transfer inhibitor treatment (InSTI), such as raltegravir before being initiated on DTG, this remains unusual. It raises the question whether other mutations outside the integrase gene, such as the polypruine

tract, which confer DTG drug resistance might have emerged first (Malet et al. 2017). Recently, envelope mutations were shown to allow in vitro HIV replication by cell to cell spread which also confer DTG resistance. However, this has not yet been described in a patient (van Duyne et al. 2019). Although the initial lack of mutations conferring DTG resistance, for 21 months, suggest that clonal viraemia could be the primary cause of the very slow viral decay, the identification of DTG resistance at 24 months suggests that other mechanisms may have contributed to treatment failure such as ongoing compartmentalised replication or mutations outside of *integrase* contributing to initial reduced susceptibility of DTG. Further investigation of this case is required to identify the underlying cause of initial non-suppressible viremia and origin of the sustained identical viraemia.

Perelson et al. previously described a theory of HIV-1 plasma virus decay based on the differential half-lives of various cellular sources of HIV viraemia (Perelson et al. 1997). This model however does not explain viral "blips" (Lee et al. 2006; Sörstedt et al. 2016) observed in some patients on treatment and does not account for the recent findings of Simonetti et al. and Halvas et al. or the two cases described here. Although the cases described here may be extreme outliers, as most cases of monotypic viremia have been described in participans with residual viremia (< 50 copies/mL) recent evidence suggests the HIV infected cell population is very dynamic and that HIV-1 infected cell clones may wax and wane (Wang et al. 2018). Moreover, when these clones harbour HIV-1 DNA transcribed to RNA or released as virions in the plasma, it may be a cause of non-suppressed viraemia and it remains unknown how often clonal viremia may result in persistent low viremia above 50 copies/mL. Therefore, HIV-1 viraemia in patients receiving cART may not be the result of either non-adherence or drug resistance, but a third alternative of clonal viraemia could explain non-suppressed viraemia. When these clones consist of large cell populations, not only viruses released from clones with intact p24 protein but also cellular debris may account for detectable viraemia, as suggested by the findings in this study. Antiretroviral therapy is now characterised by more tolerable and potent regimens with a high genetic barrier to drug resistance. These factors helped to address adherence barriers and have resulted in a low drug resistance risk. Therefore, as the most important causes of viraemia decrease and converge with highly sensitive assays to monitor HIV-1 viral loads, it may enrich for previously rare or overlooked causes of non-suppressed viraemia. Moreover, most HIV-1 clonal proliferation appears to be antigen driven (Wang et al. 2018; Simonetti et al. 2020) and clonal viraemia after vaccination or infection with a pathogen, such as influenza, may activate HIV infected T cells clones, that release virus on activation,

resulting in viral blips. Novel mechanisms of HIV drug resistance that occur outside target genes, or adaptation of HIV to replicate by cell to cell spread may also contribute to treatment failure (Malet et al. 2017; Bracq et al. 2018; van Duyne et al. 2019). The contribution of these different mechanisms of drug resistance and clonal viremia to unsuppressed viremia is unknown and require further investigation. Our findings corroborate that of others that suggest that a new framework for viral load kinetics in patients on cART is needed, that does not only take account of the different initial phases of decay but also includes clonal viraemia and other cases of apparent virologic failure.

Chapter 6

General discussion and conclusions

This doctoral project was aimed at investigating the HIV reservoir by characterising and quantifying specific proviral clones in paediatric patients on suppressive therapy. Proviral clones were further studied in adult patients presenting with non-suppressible HIV viraemia. The developed assays illustrated in chapters 2 and 3 may be used to further investigate proviral clones responsible for unsuppressed viraemia presented in chapters 4 and 5.

A novel method (ISSPA), capable of detecting severely deleted proviruses, was developed for the amplification and characterisation of specific proviruses and linking these proviruses to specific integration sites. The characterisation of proviral clones in paediatric patients with this novel assay revealed that severely deleted solo-LTR proviruses are probably more prevalent than previously expected as most assays that characterise proviruses exclude them, due to the duplication of LTRs at the 5' and 3' genome ends. In two patients multiple of the most expanded clones were characterised as solo-LTRs. Although this does not directly affect the size of the true replication competent proviral reservoir, it does suggest that when considering all HIV-1 integration events the proposed proportions of other proviruses, including intact proviruses, require adjustments to account for the prevalence of solo-LTR and partial-LTR proviruses. Recent evidence also suggests that intact proviruses decay faster (Peluso et al. 2020), most likely since these and other proviruses capable of antigen expression can be targeted by the immune system which shapes the proviral landscape (Pollack et al. 2017). In this investigation we selected the most expanded clones as they were the most feasible to detect. This may have biased our investigation towards highly defective proviruses, as such severely deleted proviruses may be undetectable by the immune system which may favour the expansion of these clones. All the proviral clones characterised in this study were integrated in untranslated/intronic regions, with 5/9 in the opposite orientation (related genes: *SRSF10, ALMS1, RAD51B, HIST1H3E, RUNX2*) and 3/9 in the same orientation (related genes: *RAB6A, TTC13* and *RANBP9*) than the related gene (Table 2.4). Human genome data for one integration site was unknown. The effects of integrated solo-LTRs on adjacent genes are not currently

known, but as they may act as promoters, they may impact on transcription of grow genes and clonal proliferation as has been seen described with proviral integration in general (Maldarelli et al. 2014).

Further investigation of these specific proviral clones by means of a novel assay (ISSAQ), capable of quantifying proviruses based on unique integration sites, was performed. This "digital" semi-nested quantitative PCR approach, like others (Bosman et al. 2015), offers efficient sensitivity and specificity for the accurate quantification of specific proviruses. This revealed that solo-LTR proviruses can be detected in up to 140 (98-190) copies per million PBMCs. The longitudinal waxing and waning of proviral clones was observed, supporting previous reports that antigen-driven and cytokine-induced proliferation contributes to clonal expansion and stability of the proviral reservoir rather than cell-autonomous proliferation (Wang et al. 2018; Simonetti et al. 2020). The ISSAQ assay offers a novel approach to the sensitive quantification of proviral clones. However, this assay does require a high volume of DNA to successfully detect rare integration events. The ISSPA and ISSAQ assays, like many other assays, are completely dependent on sampling and large blood volumes or leukopaks are required for in depth studies of the PBMC harboured proviral reservoirs.

The findings from chapters 2 and 3 together with recent data from Anderson et al, indicate that solo-LTR proviruses are a major component of the proviral landscape (Anderson et al. 2020). The transcription factor binding sites in the HIV LTR could regulate the expression or suppression of related genes (Buzdin et al. 2006; Burnett et al. 2009) however, further investigation is required to determine the effects, if any, that solo-LTR proviruses have on adjacent host genes. Furthermore, the high proportion of solo-LTR proviruses shed light on previous studies that observed up to 10-fold higher HIV DNA quantities when targeting the LTR (Rozera et al. 2010). Predominant solo-LTR proviruses could have further implications such as interference with near full-length amplification methods that prime in the HIV LTR (Li et al. 2007; Ho et al. 2013; Bruner et al. 2016), and result in LTR targeted integration assays (Maldarelli et al. 2014) yielding mostly solo-LTR related integration sites and potentially overlooking the true reservoir. Additional studies of solo-LTR and partial-LTR proviruses are required to shed light on the true proviral landscape proportions (Maldarelli 2016).

Proviral clonality was further studied in HIV subtype B positive adult patients on long term cART presenting with episodes of non-suppressible HIV viraemia. Monotypic plasma viraemia was detectable in some treatment adherent patients, similar to previously described cases

(Anderson et al. 2011; Wagner et al. 2013). In one patient sequences identical to the monotypic plasma viraemia were also detected in PBMCs and shown to be replication competent by QVOA. The identification of monotypic HIV sequences longitudinally is in agreement with previous studies showing that the reservoir is maintained by clonal expansion rather than low level replication, indicating that proviral clones are the major barrier to cure (Shen & Siliciano 2008; Dinoso et al. 2009; Kearney et al. 2017; van Zyl et al. 2017). Although identical p6PrRT SGS nucleotide sequences in patients with high diversity strongly suggest that these monotypic sequences are clonal, integration site analyses are required for confirmation (Maldarelli et al. 2014). Replication competent proviral clones are a major barrier to cure and viruses produced by such clones, resulting in clonal viraemia, can present in clinical care as non-suppressible viraemia (Simonetti et al. 2016; Halvas et al. 2019). Moreover, these proviral clones are harboured in long-lived cells that can self-renew through clonal proliferation (Maldarelli et al. 2014; Wagner et al. 2014) and are able to wax and wane over time as a result of antigen stimulation (Wang et al. 2018; Simonetti et al. 2020). The ability of these reservoir harbouring cell clones to self-renew has severe implications for curative strategies, as incomplete clearance of HIV infected cells can result in the regrowth of the reservoir (Deeks 2012; Sengupta & Siliciano 2018; Abner & Jordan 2019). In depth analyses are required to shed light on the mechanisms of persisting HIV viraemia during cART (Jacobs et al. 2019).

Further investigation of non-suppressible viraemia was performed on HIV subtype C positive adult patients. This revealed two uncommon cases of sustained viraemia in patients on long term cART, with no treatment relevant drug resistance. Firstly, a probable large cell clone harbouring non-infectious virus was found to release cell debris that contained proviral nucleic acid into plasma, resulting in apparent treatment failure. This monotypic hypermutated provirus was quantifiable at multiple time points for over a year at between 400 copies/mL and 4300 copies/mL. Nucleotide sequencing revealed a stop codon in *gag-p24*, which would result in truncated gag proteins, and therefore no virion formation (Göttlinger 2010). Secondly, possible clonal viraemia along with indications of compartmentalised replication were observed in a patient with very slow decaying viral load, with no relevant drug resistance for 21 months and excellent confirmed adherence to treatment during the study period. An initial decrease in viral load from 149 000 copies/mL to 6500 copies/mL was observed four months after amended regimen initiation, followed by very slow viral load decline by an average of ~270 copies/mL per month and which still remains above 1000 copies/mL, 26 months after treatment initiation. The observed viral load kinetics are vastly different from the expected 8 – 24 weeks to achieve

viral suppression (AIDSinfo 2020), with no adequate clinical explanation for these observations. The sustained viraemia in this case revealed two separate clusters of monotypic p6PrRT sequences. Although some monotypic sequences were also identified in cell associated RNA, these were not related to the monotypic viraemia. Thereafter a sudden emergence of high level DTG associated drug resistance due to N155H and R263K was detected, four months after the only mutation that may have been suggestive of DTG selection was a polymorphic mutation M50I. Furthermore, almost all cell associated DNA sequences were identical, suggesting the existence of a possible large cell clone however, also not responsible for the sustained viraemia. The very slow decaying viral load remains uncharacterised, but with indications of clonal viraemia and possible compartmentalised replication. The sustained viraemia is thought to be the result of either virion expression from a very small peripheral clone that remains uncharacterised, or a partially compartmentalised tissue reservoir that contains clonal cell populations that express virus or which has insufficient drug penetration to achieve viral suppression (Wong & Yukl 2016; Henderson et al. 2019) or could have been due to mutations outside of *integrase* that reduce susceptibility to DTG; and require further investigation to confirm the underlying cause. This case indicates that, similar to HIV subtype B cases, clonal viraemia occurrence may be possible in other HIV subtypes and is more prevalent than expected. Cases like these pose alternative explanations to ongoing viraemia apart from insufficient treatment exposure due to inadequate adherence, dosing or absorption, or HIV drug resistance. These cases of non-suppressible viraemia persisted too long to be classified as transient viraemia or "blips" and cannot be characterised as treatment failure, as the viraemia is refractory to adjusting treatment and therefore complicates patient care. With more sensitive viral load assays being available, the detection rate of "low-level" viraemia is becoming more common (Greub et al. 2002; Palmer et al. 2008). The cases described here suggest that an elevated viral load is not necessarily an indicator of drug resistance or non-adherence. We therefore suggest that alternative explanations should be considered in the management of low-level viraemia and apparent treatment failure. Based on these findings, cases of unexplained viraemia, despite objective evidence of adherence and without drug resitance that would explain failure, require further investigation and more in-depth investigation than routine drug resistance testing before drug regimens are altered.

Taken together, these findings further confirm the importance of proviral clones in maintaining the reservoir. These findings also suggest that severely deleted proviral clones follow waxing and waning trends similar to other proviral clones and represent an important marker of the

proviral landscape. This suggests that proviral clones may wane to below the detection limit of many assays while still being maintained, only to proliferate at a later stage. This poses challenges to achieve HIV cure as very rare, non-detectable latent proviral clones could harbour infectious virus and are able to self-renew and replenish the reservoir. Consequently, the dynamics of clonal waxing and waning require further investigation to clarify the requirements for HIV cure and remission. In addition, findings from this study suggest that viral load kinetics are more complex than previously thought and factors other than non-adherence and drug resistance should be considered in cases of apparent virologic failure.

The ISSPA assay expands on available assays to investigate the proviral landscape, as any selected proviral clone, such as the patient 1 probable clone, or severely defective clones can be characterised by using knowledge of the human genome integration position only. Furthermore, the ISSAQ assay allows for the accurate quantification of selected proviral clones and could in future be used to assess the longitudinal waxing and waning of possible intact proviral clones and the impact of antigenic stimuli on the proliferation of particular clones. It is as of yet not known how many cases of apparent virological failure could be due to clonal viremia. We therefore suggest that samples from patients with unexplained virological failure, which remains unexplained after obtaining objective evidence of adherence and HIV-1 drug resistance sequencing, should be screened for probable clonal viremia. Such screening could be done by first assessing whether plasma viremia is monotypic followed by recovering matching proviruses from PBMC DNA, after which it could be confirmed by integration site analysis. This approach is currently costly, but with minitiurization of single genome sequencing, or improvements in template tagging and next generation sequencing approaches to recovery true proviral and plasma virus haplotypes (Boltz et al., 2019), efficient identification of cases with likely clonal viremia may become more feasible in future.

Chapter 7

References

Abner, E. & Jordan, A., 2019, "HIV 'shock and kill' therapy: In need of revision," *Antiviral Research*, 166, 19–34.

Adams, M., Sharmeen, L., Kimpton, J., Romeo, J.M., Garcia, J. v., Peterlin, B.M., Groudine, M. & Emerman, M., 1994, "Cellular latency in human immunodeficiency virus-infected individuals with high CD4 levels can be detected by the presence of promoter-proximal transcripts.," *Proceedings of the National Academy of Sciences*, 91(9), 3862–3866.

Adams, M., Wong, C., Wang, D. & Romeo, J., 1999, "Limitation of Tat-Associated Transcriptional Processivity in HIV-Infected PBMC," *Virology*, 257(2), 397–405.

Ahlenstiel, C., Mendez, C., Lim, S.T.H., Marks, K., Turville, S., Cooper, D.A., Kelleher, A.D. & Suzuki, K., 2015, "Novel RNA Duplex Locks HIV-1 in a Latent State via Chromatin-mediated Transcriptional Silencing," *Molecular Therapy - Nucleic Acids*, 4, e261.

Ahmed, M., Merga, H. & Jarso, H., 2019, "Predictors of virological treatment failure among adult HIV patients on first-line antiretroviral therapy in Woldia and Dessie hospitals, Northeast Ethiopia: a case-control study," *BMC Infectious Diseases*, 19(1), 305.

Ahmed, R., Bevan, M.J., Reiner, S.L. & Fearon, D.T., 2009, "The precursors of memory: models and controversies.," *Nature reviews. Immunology*, 9(9), 662–668.

Aibekova, L., Foley, B., Hortelano, G., Raees, M., Abdraimov, S., Toichuev, R. & Ali, S., 2018, "Molecular epidemiology of HIV-1 subtype A in former Soviet Union countries," *PLOS ONE*, 13(2), e0191891.

AIDSinfo, 2020, "Guidelines for the Use of Antiretroviral Agents in Adults and Adolescents with HIV," *Panel on Antiretroviral Guidelines for Adults and Adolescents, Available at https://www.aidsinfo.nih.gov/contentfiles/lvguidelines/adultandadolescentgl.pdf*.

Aiken, C., Konner, J., Landau, N.R., Lenburg, M.E. & Trono, D., 1994, "Nef induces CD4 endocytosis: Requirement for a critical dileucine motif in the membrane-proximal CD4 cytoplasmic domain," *Cell*, 76(5), 853–864.

Aldous, J.L. & Haubrich, R.H., 2009, "Defining treatment failure in resource-rich settings," *Current Opinion in HIV and AIDS*, 4(6), 459–466.

Alexander, L., Weiskopf, E., Greenough, T.C., Gaddis, N.C., Auerbach, M.R., Malim, M.H., O'Brien, S.J., Walker, B.D., Sullivan, J.L. & Desrosiers, R.C., 2000, "Unusual Polymorphisms in Human Immunodeficiency Virus Type 1 Associated with Nonprogressive Infection," *Journal of Virology*, 74(9), 4361–4376.

Alexiev, I., Presti, A. lo, Dimitrova, R., Foley, B., Gancheva, A., Kostadinova, A., Nikolova, L., Angeletti, S., Cella, E., Elenkov, I., Stoycheva, M., Nikolova, D., Doychinova, T., Pekova, L. & Ciccozzi, M., 2018, "Origin and Spread of HIV-1 Subtype B Among Heterosexual Individuals in Bulgaria," *AIDS Research and Human Retroviruses*, 34(3), 244–253.

Ananworanich, J., Chomont, N., Eller, L.A., Kroon, E., Tovanabutra, S., Bose, M., Nau, M., Fletcher, J.L.K., Tipsuk, S., Vandergeeten, C., O'Connell, R.J., Pinyakorn, S., Michael, N., Phanuphak, N. & Robb, M.L., 2016, "HIV DNA Set Point is Rapidly Established in Acute HIV Infection and Dramatically Reduced by Early ART," *EBioMedicine*, 11, 68–72.

Ananworanich, J. & Fauci, A.S., 2015, "HIV cure research: a formidable challenge.," *Journal of virus eradication*, 1(1), 1–3.

Ananworanich, J., Fletcher, J.L., Pinyakorn, S., Griensven, F. van, Vandergeeten, C., Schuetz, A., Pankam, T., Trichavaroj, R., Akapirat, S., Chomchey, N., Phanuphak, P., Chomont, N., Michael, N.L., Kim, J.H. & Souza, M. de, 2013, "A novel acute HIV infection staging system based on 4thgeneration immunoassay," *Retrovirology*, 10(1), 56.

Anderson, E. M., & Maldarelli, F., 2018a. "The role of integration and clonal expansion in HIV infection: Live long and prosper," *Retrovirology*, 15(1).

Anderson, E.M. & Maldarelli, F., 2018b, "Quantification of HIV DNA Using Droplet Digital PCR Techniques," *Current Protocols in Microbiology*, 51(1).

Anderson, E.M., Simonetti, F.R., Gorelick, R.J., Hill, S., Gouzoulis, M.A., Bell, J., Rehm, C., Pérez, L., Boritz, E., Wu, X., Wells, D., Hughes, S.H., Rao, V., Coffin, J.M., Kearney, M.F. & Maldarelli, F., 2020, "Dynamic shifts in the HIV proviral landscape during long term combination antiretroviral therapy: Implications for persistence and control of HIV infections," *Viruses*, 12(2).

Anderson, J.A., Archin, N.M., Ince, W., Parker, D., Wiegand, A., Coffin, J.M., Kuruc, J., Eron, J., Swanstrom, R. & Margolis, D.M., 2011, "Clonal Sequences Recovered from Plasma from Patients with Residual HIV-1 Viremia and on Intensified Antiretroviral Therapy Are Identical to Replicating Viral RNAs Recovered from Circulating Resting CD4+ T Cells," *Journal of Virology*, 85(10), 5220–5223.

Anderson, J.L., Khoury, G., Fromentin, R., Solomon, A., Chomont, N., Sinclair, E., Milush, J.M., Hartogensis, W., Bacchetti, P., Roche, M., Tumpach, C., Gartner, M., Pitman, M.C., Epling, C.L., Hoh, R., Hecht, F.M., Somsouk, M., Cameron, P.U., Deeks, S.G. & Lewin, S.R., 2020, "Human Immunodeficiency Virus (HIV)–Infected CCR6+ Rectal CD4+ T Cells and HIV Persistence On Antiretroviral Therapy," *The Journal of Infectious Diseases*, 221(5), 744–755.

Arts, E.J. & Hazuda, D.J., 2012, "HIV-1 Antiretroviral Drug Therapy," *Cold Spring Harbor Perspectives in Medicine*, 2(4), a007161–a007161.

Ashorn, P., McQuade, T.J., Thaisrivongs, S., Tomasselli, A.G., Tarpley, W.G. & Moss, B., 1990, "An inhibitor of the protease blocks maturation of human and simian immunodeficiency viruses and spread of infection.," *Proceedings of the National Academy of Sciences*, 87(19), 7472–7476.

Assoumou, L., Weiss, L., Piketty, C., Burgard, M., Melard, A., Girard, P.-M., Rouzioux, C. & Costagliola, D., 2015, "A low HIV-DNA level in peripheral blood mononuclear cells at antiretroviral treatment interruption predicts a higher probability of maintaining viral control," *AIDS*, 29(15), 2003–2007.

Avert, 2020, Global HIV and AIDS statistics, *Retrieved July 16, 2020, from Avert, https://www.avert.org/global-hiv-and-aids-statistics*.

Baccetti, B., Benedetto, A., Burrini, A.G., Collodel, G., Elia, C., Piomboni, P., Renieri, T., Sensini, C. & Zaccarelli, M., 1991, "HIV particles detected in spermatozoa of patients with AIDS," *Journal of Submicroscopic Cytology and Pathology*, 23, 339–345.

Baccetti, B., Benedetto, A., Collodel, G., Caro, A. di, Garbuglia, A.R. & Piomboni, P., 1998, "The debate on the presence of HIV-1 in human gametes," *Journal of Reproductive Immunology*, 41, 41–67.

Bagasra, O., Farzadegan, H., Seshamma, T., Oakes, J.W., Saah, A. & Pomerantz, R.J., 1994, "Detection of HIV-1 proviral DNA in sperm from HIV-1-infected men," *AIDS*, 8(12), 1669–1674.

Bailey, J.R., Sedaghat, A.R., Kieffer, T., Brennan, T., Lee, P.K., Wind-Rotolo, M., Haggerty, C.M., Kamireddi, A.R., Liu, Y., Lee, J., Persaud, D., Gallant, J.E., Cofrancesco, J., Quinn, T.C., Wilke, C.O., Ray, S.C., Siliciano, J.D., Nettles, R.E. & Siliciano, R.F., 2006, "Residual Human Immunodeficiency Virus Type 1 Viremia in Some Patients on Antiretroviral Therapy Is Dominated by a Small Number of Invariant Clones Rarely Found in Circulating CD4+ T Cells," *Journal of Virology*, 80(13), 6441–6457.

Bale, M.J., Katusiime, M.G., Wells, D., Wu, X., Spindler, J., Halvas, E.K., Cyktor, J.C., Wiegand, A., Shao, W., Coffin, J.M., Cotton, M., Hughes, S.H., Mellors, J.W., van Zyl, G.U. & Kearney, M.F., 2019, "Long-term persistence of HIV-infected cell clones in children treated early," *CROI 2019*, Abstract 805.

Banga, R., Procopio, F.A., Noto, A., Pollakis, G., Cavassini, M., Ohmiti, K., Corpataux, J.-M., Leval, L. de, Pantaleo, G. & Perreau, M., 2016, "PD-1+ and follicular helper T cells are responsible for persistent HIV-1 transcription in treated aviremic individuals," *Nature Medicine*, 22(7), 754–761.

Bangsberg, D.R., 2006, "Less Than 95% Adherence to Nonnucleoside Reverse-Transcriptase Inhibitor Therapy Can Lead to Viral Suppression," *Clinical Infectious Diseases*, 43(7), 939–941.

Bangsberg, D.R., Perry, S., Charlebois, E.D., Clark, R.A., Roberston, M., Zolopa, A.R. & Moss, A., 2001, "Non-adherence to highly active antiretroviral therapy predicts progression to AIDS," *AIDS*, 15(9), 1181–1183.

Banhegyi, D., Katlama, C., Arns da Cunha, C., Schneider, S., Rachlis, A., Workman, C., Meyer, S. de, Vandevoorde, A., Casteele, T. van de & Tomaka, F., 2012, "Week 96 Efficacy, Virology and Safety of Darunavir/r Versus Lopinavir/r in Treatment-Experienced Patients in TITAN," *Current HIV Research*, 10(2), 171–181.

Barboric, M., Nissen, R.M., Kanazawa, S., Jabrane-Ferrat, N. & Peterlin, B.M., 2001, "NF-κB Binds P-TEFb to Stimulate Transcriptional Elongation by RNA Polymerase II," *Molecular Cell*, 8(2), 327–337.

Barton, K., Winckelmann, A. & Palmer, S., 2016, *HIV-1 Reservoirs During Suppressive Therapy*, *Trends in Microbiology*, 24(5), 345–355.

Basu, V.P., Song, M., Gao, L., Rigby, S.T., Hanson, M.N. & Bambara, R.A., 2008, "Strand transfer events during HIV-1 reverse transcription," *Virus Research*, 134(1–2), 19–38.

Bbosa, N., Kaleebu, P. & Ssemwanga, D., 2019, "HIV subtype diversity worldwide," *Current Opinion in HIV and AIDS*, 14(3), 153–160.

Benjamin, J., Ganser-Pornillos, B.K., Tivol, W.F., Sundquist, W.I. & Jensen, G.J., 2005, "Three-dimensional Structure of HIV-1 Virus-like Particles by Electron Cryotomography," *Journal of Molecular Biology*, 346(2), 577–588.

Berkowitz, R., Fisher, J. & Goff, S.P., 1996, "RNA Packaging," *Current Topics in Microbiology and Immunology*, 177–218.

Bernstein, H.B., Tucker, S.P., Kar, S.R., Mcpherson, S.A., Mcpherson, D.T., Dubay, J.W., Lebowitz, J., Compans, R.W. & Hunter, E., 1995, *Oligomerization of the Hydrophobic Heptad Repeat of gp41*, vol. 69.

Besson, G.J., Lalama, C.M., Bosch, R.J., Gandhi, R.T., Bedison, M.A., Aga, E., Riddler, S.A., McMahon, D.K., Hong, F. & Mellors, J.W., 2014, "HIV-1 DNA decay dynamics in blood during more than a decade of suppressive antiretroviral therapy," *Clinical Infectious Diseases*, 59(9), 1312–1321.

Billings, E., Sanders-Buell, E., Bose, M., Kijak, G.H., Bradfield, A., Crossler, J., Arroyo, M.A., Maboko, L., Hoffmann, O., Geis, S., Birx, D.L., Kim, J.H., Michael, N.L., Robb, M.L., Hoelscher, M. & Tovanabutra, S., 2017, "HIV-1 Genetic Diversity Among Incident Infections in Mbeya, Tanzania," *AIDS Research and Human Retroviruses*, 33(4), 373–381.

Bishop, K.N., Holmes, R.K., Sheehy, A.M., Davidson, N.O., Cho, S.-J. & Malim, M.H., 2004, "Cytidine Deamination of Retroviral DNA by Diverse APOBEC Proteins," *Current Biology*, 14(15), 1392–1396.

Blankson, J.N., Finzi, D., Pierson, T.C., Sabundayo, B.P., Chadwick, K., Margolick, J.B., Quinn, T.C. & Siliciano, R.F., 2000, "Biphasic Decay of Latently Infected CD4+ T Cells in Acute Human Immunodeficiency Virus Type 1 Infection," *The Journal of Infectious Diseases*, 182(6), 1636–1642.

Blazkova, J., Trejbalova, K., Gondois-Rey, F., Halfon, P., Philibert, P., Guiguen, A., Verdin, E., Olive, D., Lint, C. van, Hejnar, J. & Hirsch, I., 2009, "CpG Methylation Controls Reactivation of HIV from Latency," *PLoS Pathogens*, 5(8), e1000554.

Blumenthal, R., Durell, S. & Viard, M., 2012, "HIV Entry and Envelope Glycoprotein-mediated Fusion," *Journal of Biological Chemistry*, 287(49), 40841–40849.

Boadu, R., Darko, G., Nortey, P., Akweongo, P. & Sarfo, B., 2016, "Assessing the sensitivity and specificity of First Response HIV-1-2 test kit with whole blood and serum samples: a cross-sectional study," *AIDS Research and Therapy*, 13(1), 9.

Boltz, V. F., Shao, W., Bale, M. J., Halvas, E. K., Luke, B., McIntyre, J. A., Schooley, R. T., Lockman, S., Currier, J. S., Sawe, F., Hogg, E., Hughes, M. D., Kearney, M. F., Coffin, J. M., & Mellors, J. W., 2019, "Linked dual-class HIV resistance mutations are associated with treatment failure," *JCI Insight*, 4(19), e130118.

Borzy, M.S., Connell, R.S. & Kiessling, A.A., 1988, "Detection of human immunodeficiency virus in cell-free seminal fluid," *Journal of Acquired Immune Deficiency Syndromes*," 1, 419–424.

Bosman, K.J., Nijhuis, M., Ham, P.M. van, Wensing, A.M.J., Vervisch, K., Vandekerckhove, L. & Spiegelaere, W. de, 2015, "Comparison of digital PCR platforms and semi-nested qPCR as a tool to determine the size of the HIV reservoir," *Scientific Reports*, 5, 13811.

Botha, J.C., Steegan, K., Hans, L., Karstaedt, A., Carmona, S., Reddy, D., Kearney, M., Mellors, J. & Zyl, G. van, 2020, "'False ART failure' from identical hypermutated HIV nucleic acid in plasma," *CROI 2020*, Abstract 373.

Botha, J.C., Steegan, K., Hans, L., Karstaedt, A., Carmona, S., Reddy, D. & Zyl, G. van, 2019, "HIV Pseudo viremia from large defective cell population," *28th International Workshop on HIV Drug Resistance and Treatment Strategies*.

Bracq, Lucie, Maorong Xie, Serge Benichou, and J. Bouchet. 2018. "Mechanisms for Cell-to-Cell Transmission of HIV-1." *Frontiers in Immunology*, 9 (FEB), 260.

Brady, T., Agosto, L.M., Malani, N., Berry, C.C., O'Doherty, U. & Bushman, F., 2009, "HIV integration site distributions in resting and activated CD4+ T cells infected in culture," *AIDS*, 23(12), 1461–1471.

Brady, T., Kelly, B.J., Male, F., Roth, S., Bailey, A., Malani, N., Gijsbers, R., O'Doherty, U. & Bushman, F.D., 2013, "Quantitation of HIV DNA integration: Effects of differential integration site distributions on Alu-PCR assays," *Journal of Virological Methods*, 189(1), 53–57.

Brenner, B.G. & Wainberg, M.A., 2017, "Clinical benefit of dolutegravir in HIV-1 management related to the high genetic barrier to drug resistance," *Virus Research*, 239, 1–9.

Briggs, J.A.G., 2003, "Structural organization of authentic, mature HIV-1 virions and cores," *The EMBO Journal*, 22(7), 1707–1715.

Briggs, J.A.G., Grünewald, K., Glass, B., Förster, F., Kräusslich, H.-G. & Fuller, S.D., 2006, "The Mechanism of HIV-1 Core Assembly: Insights from Three-Dimensional Reconstructions of Authentic Virions," *Structure*, 14(1), 15–20.

Broder, S., 2010, "The development of antiretroviral therapy and its impact on the HIV-1/AIDS pandemic," *Antiviral Research*, 85(1), 1–18.

Bruner, K.M., Murray, A.J., Pollack, R.A., Soliman, M.G., Laskey, S.B., Capoferri, A.A., Lai, J., Strain, M.C., Lada, S.M., Hoh, R., Ho, Y.-C., Richman, D.D., Deeks, S.G., Siliciano, J.D. & Siliciano, R.F., 2016, "Defective proviruses rapidly accumulate during acute HIV-1 infection," *Nature Medicine*, 22(9), 1043–1049.

Bruner, K.M., Wang, Z., Simonetti, F.R., Bender, A.M., Kwon, K.J., Sengupta, S., Fray, E.J., Beg, S.A., Antar, A.A.R., Jenike, K.M., Bertagnolli, L.N., Capoferri, A.A., Kufera, J.T., Timmons, A., Nobles, C., Gregg, J., Wada, N., Ho, Y.-C., Zhang, H., Margolick, J.B., Blankson, J.N., Deeks, S.G., Bushman, F.D., Siliciano, J.D., Laird, G.M. & Siliciano, R.F., 2019, "A quantitative approach for measuring the reservoir of latent HIV-1 proviruses," *Nature*, 566(7742), 120–125.

Bui, J.K., Sobolewski, M.D., Keele, B.F., Spindler, J., Musick, A., Wiegand, A., Luke, B.T., Shao, W., Hughes, S.H., Coffin, J.M., Kearney, M.F. & Mellors, J.W., 2017, "Proviruses with identical sequences comprise a large fraction of the replication-competent HIV reservoir," *PLoS Pathogens*, 13(3), 1–18.

Bukrinsky, M.I., Haggerty, S., Dempsey, M.P., Sharova, N., Adzhubei, A., Spitz, L., Lewis, P., Goldfarb, D., Emerman, M. & Stevenson, M., 1993, "A nuclear localization signal within HIV-1 matrix protein that governs infection of non-dividing cells," *Nature*, 365(6447), 666–669.

Bull, M.E., Learn, G.H., McElhone, S., Hitti, J., Lockhart, D., Holte, S., Dragavon, J., Coombs, R.W., Mullins, J.I. & Frenkel, L.M., 2009, "Monotypic Human Immunodeficiency Virus Type 1 Genotypes across the Uterine Cervix and in Blood Suggest Proliferation of Cells with Provirus," *Journal of Virology*, 83(12), 6020–6028.

Burnett, J.C., Miller-Jensen, K., Shah, P.S., Arkin, A.P. & Schaffer, D. v., 2009, "Control of Stochastic Gene Expression by Host Factors at the HIV Promoter," *PLoS Pathogens*, 5(1), e1000260.

Buzdin, A.A., Kovalskaya-Alexandrova, E., Gogvadze, E. & Sverdlov, E., 2006, "At Least 50% of Human-Specific HERV-K (HML-2) Long Terminal Repeats Serve In Vivo as Active Promoters for Host Nonrepetitive DNA Transcription," *Journal of Virology*, 80(21), 10752–10762.

Buzdin, A.A., Prassolov, V. & Garazha, A.V., 2017, Friends-Enemies: Endogenous retroviruses are major transcriptional regulators of human DNA, "*Frontiers in Chemistry"*, 5(35).

Campbell, S.M., Crowe, S.M. & Mak, J., 2001, "Lipid rafts and HIV-1: from viral entry to assembly of progeny virions," *Journal of Clinical Virology*, 22(3), 217–227.

Canaud, G., Dejucq-Rainsford, N., Avettand-Fenoël, V., Viard, J.-P., Anglicheau, D., Bienaimé, F., Muorah, M., Galmiche, L., Gribouval, O., Noël, L.-H., Satie, A.-P., Martinez, F., Sberro-Soussan, R., Scemla, A., Gubler, M.-C., Friedlander, G., Antignac, C., Timsit, M.-O., Onetti Muda, A., Terzi, F., Rouzioux, C. & Legendre, C., 2014, "The Kidney as a Reservoir for HIV-1 after Renal Transplantation," *Journal of the American Society of Nephrology*, 25(2), 407–419.

Cassan, E., Arigon-Chifolleau, A.-M., Mesnard, J.-M., Gross, A. & Gascuel, O., 2016, "Concomitant emergence of the antisense protein gene of HIV-1 and of the pandemic," *Proceedings of the National Academy of Sciences*, 113(41), 11537–11542.

Castley, A., Sawleshwarkar, S., Varma, R., Herring, B., Thapa, K., Dwyer, D., Chibo, D., Nguyen, N., Hawke, K., Ratcliff, R., Garsia, R., Kelleher, A. & Nolan, D., 2017, "A national study of the molecular epidemiology of HIV-1 in Australia 2005–2012," *PLOS ONE*, 12(5), e0170601.

CDC, 2020, "Types of HIV tests," *Retrieved July 17, 2020, from CDC, https://www.cdc.gov/hiv/basics/hiv-testing/test types.html#:~:text=There%20are%20three%20types%20of,also%20be%20performed %20on%20urine.*

Cesana, D., Santoni De Sio, F. R., Rudilosso, L., Gallina, P., Calabria, A., Beretta, S., Merelli, I., Bruzzesi, E., Passerini, L., Nozza, S., Vicenzi, E., Poli, G., Gregori, S., Tambussi, G., & Montini, E., 2017, "HIV-1-mediated insertional activation of STAT5B and BACH2 trigger viral reservoir in T regulatory cells," *Nature Communications*, 8(1).

Chaillon, A., Gianella, S., Dellicour, S., Rawlings, S.A., Schlub, T.E., Oliveira, M.F. de, Ignacio, C., Porrachia, M., Vrancken, B. & Smith, D.M., 2020, "HIV persists throughout deep tissues with repopulation from multiple anatomical sources," *Journal of Clinical Investigation*, 130(4), 1699–1712.

Chakrabarti, A.K., Caruso, L., Ding, M., Shen, C., Buchanan, W., Gupta, P., Rinaldo, C.R. & Chen, Y., 2009, "Detection of HIV-1 RNA/DNA and CD4 mRNA in feces and urine from chronic HIV-1 infected subjects with and without anti-retroviral therapy," *AIDS Research and Therapy*, 6(1), 20.

Chan, D.C. & Kim, P.S., 1998, "HIV Entry and Its Inhibition," *Cell*, 93(5), 681–684.

Chen, H. & Engelman, A., 1998, "The barrier-to-autointegration protein is a host factor for HIV type 1 integration.," *Proceedings of the National Academy of Sciences of the United States of America*, 95(26), 15270–4.

Cheng-Mayer, C. & Levy, J.A., 1988, "Distinct biological and serological properties of human immunodeficiency viruses from the brain," *Annals of Neurology*, 23(S1), S58–S61.

Chiodi, F., Valentin, A., Keys, B., Schwartz, S., åsjö, B., Gartner, S., Popovic, M., Albert, J., Sundqvist, V.-A. & Fenyö, E.M., 1989, "Biological characterization of paired human immunodeficiency virus type 1 isolates from blood and cerebrospinal fluid," *Virology*, 173(1), 178–187.

Chomont, N., El-Far, M., Ancuta, P., Trautmann, L., Procopio, F.A., Yassine-Diab, B., Boucher, G., Boulassel, M.R., Ghattas, G., Brenchley, J.M., Schacker, T.W., Hill, B.J., Douek, D.C., Routy, J.P., Haddad, E.K. & Sékaly, R.P., 2009, "HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation," *Nature Medicine*, 15(8), 893–900.

Chomont, N., Hocini, H., Grésenguet, G., Brochier, C., Bouhlal, H., Andréoletti, L., Becquart, P., Charpentier, C., Dieu Longo, J. de, Si-Mohamed, A., Kazatchkine, M.D. & Bélec, L., 2007, "Early archives of genetically-restricted proviral DNA in the female genital tract after heterosexual transmission of HIV-1," *AIDS*, 21(2), 153–162.

Chun, J.Y., Kim, K.J., Hwang, I.T., Kim, Y.J., Lee, D.H., Lee, I.K. & Kim, J.K., 2007a, "Dual priming oligonucleotide system for the multiplex detection of respiratory viruses and SNP genotyping of CYP2C19 gene," *Nucleic Acids Research*, 35(6), 6–11.

Chun, T.-W., 2013, "Tracking replication-competent HIV reservoirs in infected individuals," *Current Opinion in HIV and AIDS*, 8(2), 111–116.

Chun, T.W., Nickle, D.C., Justement, J.S., Meyers, J.H., Roby, G., Hallahan, C.W., Kottilil, S., Moir, S., Mican, J.M., Mullins, J.I., Ward, D.J., Kovacs, J.A., Mannon, P.J. & Fauci, A.S., 2008, "Persistence of HIV in gut-associated lymphoid tissue despite long-term antiretroviral therapy.," *The Journal of infectious diseases*, 197(5), 714–20.

Chun, T.W., Stuyver, L., Mizell, S.B., Ehler, L.A., Mican, J.A.M., Baseler, M., Lloyd, A.L., Nowak, M.A. & Fauci, A.S., 1997, "Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy," *Proceedings of the National Academy of Sciences*, 94(24), 13193–13197.

Cillo, A.R., Krishnan, A., Mitsuyasu, R.T., McMahon, D.K., Li, S., Rossi, J.J., Zaia, J. a & Mellors, J.W., 2013, "Plasma viremia and cellular HIV-1 DNA persist despite autologous hematopoietic stem cell transplantation for HIV-related lymphoma.," *Journal of acquired immune deficiency syndromes (1999)*, 63(4), 438–41.

Cillo, A.R., Vagratian, D., Bedison, M.A., Anderson, E.M., Kearney, M.F., Fyne, E., Koontz, D., Coffin, J.M., Piatak, M. & Mellors, J.W., 2014, "Improved single-copy assays for quantification of persistent HIV-1 viremia in patients on suppressive antiretroviral therapy," *Journal of Clinical Microbiology*, 52(11), 3944–3951.

Ciuffi, A. & Barr, S.D., 2011, "Identification of HIV integration sites in infected host genomic DNA," *Methods*, 53(1), 39–46.

Clotet, B., Bellos, N., Molina, J.-M., Cooper, D., Goffard, J.-C., Lazzarin, A., Wöhrmann, A., Katlama, C., Wilkin, T., Haubrich, R., Cohen, C., Farthing, C., Jayaweera, D., Markowitz, M., Ruane, P., Spinosa-Guzman, S. & Lefebvre, E., 2007, "Efficacy and safety of darunavir-ritonavir at week 48 in treatment-experienced patients with HIV-1 infection in POWER 1 and 2: a pooled subgroup analysis of data from two randomised trials," *The Lancet*, 369(9568), 1169–1178.

Coffin, J., 1995, "HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy," *Science*, 267(5197), 483–489.

Cohen, C., 2009, "Low-Level Viremia in HIV-1 Infection: Consequences and Implications for Switching to a New Regimen," *HIV Clinical Trials*, 10(2), 116–124.

Cohen, E.A., Dehni, G., Sodroski, J.G. & Haseltine, W.A., 1990, "Human Immunodeficiency Virus vpr Product Is a Virion-Associated Regulatory Protein," *Journal of Virology*, 3097–3099.

Cohen, M.S., Gay, C.L., Busch, M.P. & Hecht, F.M., 2010, "The Detection of Acute HIV Infection," *The Journal of Infectious Diseases*, 202(S2), S270–S277.

Cohn, L.B. & Nussenzweig, M.C., 2017, "HIV: Persistence through division.," *Journal of experimental medicine*, 214(4), 875–876.

Cohn, L.B., Silva, I.T., Oliveira, T.Y., Rosales, R.A., Parrish, E.H., Learn, G.H., Hahn, B.H., Czartoski, J.L., McElrath, M.J., Lehmann, C., Klein, F., Caskey, M., Walker, B.D., Siliciano, J.D., Siliciano, R.F., Jankovic, M. & Nussenzweig, M.C., 2015, "HIV-1 integration landscape during latent and active infection," *Cell*, 160(3), 420–432.

Contreras-Galindo, R., Kaplan, M.H., He, S., Contreras-Galindo, A.C., Gonzalez-Hernandez, M.J., Kappes, F., Dube, D., Chan, S.M., Robinson, D., Meng, F., Dai, M., Gitlin, S.D., Chinnaiyan, A.M., Omenn, G.S. & Markovitz, D.M., 2013, "HIV infection reveals widespread expansion of novel centromeric human endogenous retroviruses," *Genome Research*, 23(9), 1505–1513.

Coombs, R.W., Lockhart, D., Ross, S.O., Deutsch, L., Dragavon, J., Diem, K., Hooton, T.M., Collier, A.C., Corey, L. & Krieger, J.N., 2006, "Lower Genitourinary Tract Sources of Seminal HIV," *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 41(4), 430–438.

Costiniuk, C.T., Salahuddin, S., Farnos, O., Olivenstein, R., Pagliuzza, A., Orlova, M., Schurr, E., Castro, C. de, Bourbeau, J., Routy, J.-P., Ancuta, P., Chomont, N. & Jenabian, M.-A., 2018, "HIV persistence in mucosal CD4+ T-cells within the lungs of adults receiving long-term suppressive antiretroviral therapy," *AIDS*, 32(16), 2279–2289.

Cotton, M.F., Violari, A., Otwombe, K., Panchia, R., Dobbels, E., Rabie, H., Josipovic, D., Liberty, A., Lazarus, E., Innes, S., Rensburg, A.J. van, Pelser, W., Truter, H., Madhi, S.A., Handelsman, E., Jean-Philippe, P., McIntyre, J.A., Gibb, D.M. & Babiker, A.G., 2013, "Early time-limited antiretroviral therapy versus deferred therapy in South African infants infected with HIV: Results from the children with HIV early antiretroviral (CHER) randomised trial," *The Lancet*, 382(9904), 1555–1563.

Couturier, J., Suliburk, J.W., Brown, J.M., Luke, D.J., Agarwal, N., Yu, X., Nguyen, C., Iyer, D., Kozinetz, C.A., Overbeek, P.A., Metzker, M.L., Balasubramanyam, A. & Lewis, D.E., 2015, "Human adipose tissue as a reservoir for memory CD4+ T cells and HIV," *AIDS*, 29(6), 667–674.

Craigie, R. & Bushman, F.D., 2012, "HIV DNA Integration," *Cold Spring Harbor Perspectives in Medicine*, 2(7), a006890–a006890.

Crooks, A.M., Bateson, R., Cope, A.B., Dahl, N.P., Griggs, M.K., Kuruc, J.D., Gay, C.L., Eron, J.J., Margolis, D.M., Bosch, R.J. & Archin, N.M., 2015, "Precise Quantitation of the Latent HIV-1 Reservoir: Implications for Eradication Strategies," *Journal of Infectious Diseases*, 212(9), 1361–1365.

Dam, E., Quercia, R., Glass, B., Descamps, D., Launay, O., Duval, X., Kräusslich, H.-G., Hance, A.J. & Clavel, F., 2009, "Gag Mutations Strongly Contribute to HIV-1 Resistance to Protease Inhibitors in Highly Drug-Experienced Patients besides Compensating for Fitness Loss," *PLoS Pathogens*, 5(3), e1000345.

Damond, F., Roquebert, B., Benard, A., Collin, G., Miceli, M., Yeni, P., Brun-Vezinet, F. & Descamps, D., 2007, "Human Immunodeficiency Virus Type 1 (HIV-1) Plasma Load Discrepancies between the Roche COBAS AMPLICOR HIV-1 MONITOR Version 1.5 and the Roche COBAS AmpliPrep/COBAS TaqMan HIV-1 Assays," *Journal of Clinical Microbiology*, 45(10), 3436–3438.

Damouche, A., Lazure, T., Avettand-Fènoël, V., Huot, N., Dejucq-Rainsford, N., Satie, A.-P., Mélard, A., David, L., Gommet, C., Ghosn, J., Noel, N., Pourcher, G., Martinez, V., Benoist, S., Béréziat, V., Cosma, A., Favier, B., Vaslin, B., Rouzioux, C., Capeau, J., Müller-Trutwin, M., Dereuddre-Bosquet, N., Grand, R. le, Lambotte, O. & Bourgeois, C., 2015, "Adipose Tissue Is a Neglected Viral Reservoir and an Inflammatory Site during Chronic HIV and SIV Infection," *PLOS Pathogens*, 11(9), e1005153.

Dash, P.K., Kaminski, R., Bella, R., Su, H., Mathews, S., Ahooyi, T.M., Chen, C., Mancuso, P., Sariyer, R., Ferrante, P., Donadoni, M., Robinson, J.A., Sillman, B., Lin, Z., Hilaire, J.R., Banoub, M., Elango, M., Gautam, N., Mosley, R.L., Poluektova, L.Y., McMillan, J., Bade, A.N., Gorantla, S., Sariyer, I.K., Burdo, T.H., Young, W.-B., Amini, S., Gordon, J., Jacobson, J.M., Edagwa, B., Khalili, K. & Gendelman, H.E., 2019, "Sequential LASER ART and CRISPR Treatments Eliminate HIV-1 in a Subset of Infected Humanized Mice," *Nature Communications*, 10(1), 2753.

Davenport, M.P., Khoury, D.S., Cromer, D., Lewin, S.R., Kelleher, A.D. & Kent, S.J., 2019, "Functional cure of HIV: the scale of the challenge," *Nature Reviews Immunology*, 19(1), 45–54.

Daw, M.A., El-Bouzedi, A., Ahmed, M.O. & Dau, A.A., 2017, "Molecular and epidemiological characterization of HIV-1 subtypes among Libyan patients," *BMC Research Notes*, 10(1), 170.

De Marco, A., Biancotto, C., Knezevich, A., Maiuri, P., Vardabasso, C. & Marcello, A., 2008, "Intragenic transcriptional cis-activation of the human immunodeficiency virus 1 does not result in allele-specific inhibition of the endogenous gene," *Retrovirology*, 5, 98.

Deeks, S.G., 2012, "Shock and kill," *Nature*, 487(7408), 439–440.

Deere, J.D., Kauffman, R.C., Cannavo, E., Higgins, J., Villalobos, A., Adamson, L., Schinazi, R.F., Luciw, P.A. & North, T.W., 2014, "Analysis of Multiply Spliced Transcripts in Lymphoid Tissue Reservoirs of Rhesus Macaques Infected with RT-SHIV during HAART," *PLoS ONE*, 9(2), e87914.

Delchambre, M., Gheysen, D., Thines, D., Thiriart, C., Jacobs, E., Verdin, E., Horth, M., Burny, A. & Bex, F., 1989, "The GAG precursor of simian immunodeficiency virus assembles into virus-like particles.," *The EMBO Journal*, 8(9), 2653–2660.

Deng, K., Pertea, M., Rongvaux, A., Wang, L., Durand, C.M., Ghiaur, G., Lai, J., McHugh, H.L., Hao, H., Zhang, H., Margolick, J.B., Gurer, C., Murphy, A.J., Valenzuela, D.M., Yancopoulos, G.D., Deeks, S.G., Strowig, T., Kumar, P., Siliciano, J.D., Salzberg, S.L., Flavell, R.A., Shan, L. & Siliciano, R.F., 2015, "Broad CTL response is required to clear latent HIV-1 due to dominance of escape mutations," *Nature*, 517(7534), 381–385.

Dimonte, S., 2017, "Different HIV-1 *env* frames: gp120 and ASP (antisense protein) biosynthesis, and their co-variation tropic amino acid signatures in X4- and R5-viruses," *Journal of Medical Virology*, 89(1), 112–122.

Dinoso, J.B., Kim, S.Y., Wiegand, A.M., Palmer, S.E., Gange, S.J., Cranmer, L., O'Shea, A., Callender, M., Spivak, A., Brennan, T., Kearney, M.F., Proschan, M.A., Mican, J.M., Rehm, C.A., Coffin, J.M., Mellors, J.W., Siliciano, R.F. & Maldarelli, F., 2009, "Treatment intensification does not reduce residual HIV-1 viremia in patients on highly active antiretroviral therapy," *Proceedings of the National Academy of Sciences*, 106(23), 9403–9408.

Doitsh, G., Galloway, N.L.K., Geng, X., Yang, Z., Monroe, K.M., Zepeda, O., Hunt, P.W., Hatano, H., Sowinski, S., Muñoz-Arias, I. & Greene, W.C., 2014, "Cell death by pyroptosis drives CD4 T-cell depletion in HIV-1 infection," *Nature*, 505(7484), 509–514.

Doms, R.W. & Trono, D., 2000, "The plasma membrane as a combat zone in the HIV battlefield," *Genes & Development*, 14(21), 2677–2688.

Dorfman, T., Mammano, F., Haseltine, W.A. & Göttlinger, H.G., 1994, "Role of the matrix protein in the virion association of the human immunodeficiency virus type 1 envelope glycoprotein.," *Journal of Virology*, 68(3), 1689.

Doyle, T., Smith, C., Vitiello, P., Cambiano, V., Johnson, M., Owen, A., Phillips, A.N. & Geretti, A.M., 2012, "Plasma HIV-1 RNA Detection Below 50 Copies/mL and Risk of Virologic Rebound in Patients Receiving Highly Active Antiretroviral Therapy," *Clinical Infectious Diseases*, 54(5), 724–732.

EACS, 2018, "Guidelines Version 9.1 European AIDS Clinical Society," *Available at https://www.eacsociety.org/files/2018_guidelines-9.1-english.pdf*.

Ebina, H., Misawa, N., Kanemura, Y. & Koyanagi, Y., 2013, "Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus," *Scientific Reports*, 3(1), 2510.

Einkauf, K.B., Lee, G.Q., Gao, C., Sharaf, R., Sun, X., Hua, S., Chen, S.M.Y., Jiang, C., Lian, X., Chowdhury, F.Z., Rosenberg, E.S., Chun, T.-W., Li, J.Z., Yu, X.G. & Lichterfeld, M., 2019, "Intact HIV-1 proviruses accumulate at distinct chromosomal positions during prolonged antiretroviral therapy," *Journal of Clinical Investigation*, 129(3), 988–998.

Endalamaw, A., Mekonnen, M., Geremew, D., Yehualashet, F.A., Tesera, H. & Habtewold, T.D., 2020, "HIV/AIDS treatment failure and associated factors in Ethiopia: meta-analysis," *BMC Public Health*, 20(1), 82.

Engelman, A. & Cherepanov, P., 2012, "The structural biology of HIV-1: mechanistic and therapeutic insights," *Nature Reviews Microbiology*, 10(4), 279–290.

Eriksson, S., Graf, E.H., Dahl, V., Strain, M.C., Yukl, S.A., Lysenko, E.S., Bosch, R.J., Lai, J., Chioma, S., Emad, F., Abdel-Mohsen, M., Hoh, R., Hecht, F., Hunt, P., Somsouk, M., Wong, J., Johnston, R., Siliciano, R.F., Richman, D.D., O'Doherty, U., Palmer, S., Deeks, S.G. & Siliciano, J.D., 2013, "Comparative Analysis of Measures of Viral Reservoirs in HIV-1 Eradication Studies," *PLoS Pathogens*, 9(2).

Eron, J.J., Clotet, B., Durant, J., Katlama, C., Kumar, P., Lazzarin, A., Poizot-Martin, I., Richmond, G., Soriano, V., Ait-Khaled, M., Fujiwara, T., Huang, J., Min, S., Vavro, C., Yeo, J., Walmsley, S.L., Cox, J., Reynes, J., Morlat, P., Vittecoq, D., Livrozet, J.-M., Fernández, P.V., Gatell, J.M., DeJesus, E., DeVente, J., Lalezari, J.P., McCurdy, L.H., Sloan, L.A., Young, B., LaMarca, A. & Hawkins, T., 2013, "Safety and Efficacy of Dolutegravir in Treatment-Experienced Subjects With Raltegravir-Resistant HIV Type 1 Infection: 24-Week Results of the VIKING Study," *The Journal of Infectious Diseases*, 207(5), 740–748.

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," *Molecular Biology International*, 2012, 1–23.

Evering, T.H., Kamau, E., Bernard, L. st., Farmer, C.B., Kong, X.-P. & Markowitz, M., 2014, "Single genome analysis reveals genetic characteristics of Neuroadaptation across HIV-1 envelope," *Retrovirology*, 11(1), 65.

Farnet, C.M. & Bushman, F.D., 1997, "HIV-1 cDNA Integration: Requirement of HMG I(Y) Protein for Function of Preintegration Complexes In Vitro," *Cell*, 88(4), 483–492.

Feng, S. & Holland, E.C., 1988, "HIV-1 tat trans-activation requires the loop sequence within tar," *Nature*, 334(6178), 165–167.

Finzi, D., Hermankova, M., Pierson, T., Carruth, L.M., Buck, C., Chaisson, R.E., Quinn, T.C., Chadwick, K., Margolick, J., Brookmeyer, R., Gallant, J., Markowitz, M., Ho, D.D., Richman, D.D. & Siliciano, R.F., 1997, "Identification of a Reservoir for HIV-1 in Patients on Highly Active Antiretroviral Therapy," *Science*, 278(5341), 1295–1300.

Fiscus, S.A., Cu-Uvin, S., Eshete, A.T., Hughes, M.D., Bao, Y., Hosseinipour, M., Grinsztejn, B., Badal-Faesen, S., Dragavon, J., Coombs, R.W., Braun, K., Moran, L., Hakim, J., Flanigan, T., Kumarasamy, N., Campbell, T.B., Klingman, K.L., Nair, A., Walawander, A., Smeaton, L.M., Gruttola, V. de, Martinez, A.I., Swann, E., Barnett, R.L., Brizz, B., Delph, Y., Gettinger, N., Mitsuyasu, R.T., Eshleman, S., Safren, S., Andrade, A., Haas, D.W., Amod, F., Berthaud, V., Bollinger, R.C., Bryson, Y., Celentano, D., Chilongozi, D., Cohen, M., Collier, A.C., Currier, J.S., Eron, J., Firnhaber, C., Flexner, C., Gallant, J.E., Gulick, R.M., Hammer, S.M., Hoffman, I., Kazembe, P., Kumwenda, J., Kumwenda, N., Lama, J.R., Lawrence, J., Maponga, C., Martinson, F., Mayer, K., Nielsen, K., Pendame, R.B., Ramratnam, B., Rooney, J.F., Sanchez, J., Sanne, I., Schooley, R.T., Snowden, W., Solomon, S., Tabet, S., Taha, T., Uy, J., Horst, C. van der, Wanke, C., Gormley, J., Marcus, C.J., Putnam, B., Ntshele, S., Loeliger, E., Pappa, K.A., Webb, N., Shugarts, D.L., Winters, M.A., Descallar, R.S., Sharma, J., Poongulali, S., Cardoso, S.W., Faria, D.L., Berendes, S., Burke, K., Kanyama, C., Kayoyo, V., Samaneka, W.P., Chisada, A., Santos, Breno, Rosa, A. la, Infante, R., Balfour, H.H., Mullan, B., Kim, G.-Y., Klebert, M.K., Mildvan, D., Revuelta, M., Jan Geiseler, P., Santos, Bartolo, Daar, E.S., Lopez, R., Frarey, L., Currin, D., Haas, D.H., Bailey, V.L., Tebas, P., Zifchak, L., Sha, B.E. & Fritsche, J.M., 2013, "Changes in HIV-1 Subtypes B and C Genital Tract RNA in Women and Men After Initiation of Antiretroviral Therapy," *Clinical Infectious Diseases*, 57(2), 290–297.

Freed, E.O., 1998, "HIV-1 Gag Proteins: Diverse Functions in the Virus Life Cycle," *Virology*, 251(1), 1–15.

Fung, I.C.-H., Gambhir, M., Sighem, A. van, Wolf, F. de & Garnett, G.P., 2012, "The Clinical Interpretation of Viral Blips in HIV Patients Receiving Antiviral Treatment," *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 60(1), 5–11.

Gallay, P., Hope, T., Chin, D. & Trono, D., 1997, "HIV-1 infection of nondividing cells through the recognition of *integrase* by the importin/karyopherin pathway," *Proceedings of the National Academy of Sciences*, 94(18), 9825–9830.

Gandhi, R.T., Kwon, D.S., Macklin, E.A., Shopis, J.R., McLean, A.P., McBrine, N., Flynn, T., Peter, L., Sbrolla, A., Kaufmann, D.E., Porichis, F., Walker, B.D., Bhardwaj, N., Barouch, D.H. & Kavanagh, D.G., 2016, "Immunization of HIV-1-Infected Persons With Autologous Dendritic Cells Transfected With mRNA Encoding HIV-1 Gag and Nef," *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 71(3), 246–253.

Ganser-Pornillos, B.K., Yeager, M. & Sundquist, W.I., 2008, "The structural biology of HIV assembly," *Current Opinion in Structural Biology*, 18(2), 203–217.

García, F., Climent, N., Assoumou, L., Gil, C., González, N., Alcamí, J., León, A., Romeu, J., Dalmau, J., Martínez-Picado, J., Lifson, J., Autran, B., Costagliola, D., Clotet, B., Gatell, J.M., Plana, M. & Gallart, T., 2011, "A Therapeutic Dendritic Cell-Based Vaccine for HIV-1 Infection," *The Journal of Infectious Diseases*, 203(4), 473–478.

Garcia, J.V. & Miller, A.D., 1992, "Downregulation of cell surface CD4 by nef," *Research in Virology*, 143, 52–55.

Garcia-Gasco, P., Maida, I., Blanco, F., Barreiro, P., Martin-Carbonero, L., Vispo, E., Gonzalez-Lahoz, J. & Soriano, V., 2008, "Episodes of low-level viral rebound in HIV-infected patients on antiretroviral therapy: frequency, predictors and outcome," *Journal of Antimicrobial Chemotherapy*, 61(3), 699–704.

Garg, H., Viard, M., Jacobs, A. & Blumenthal, R., 2011, "Targeting HIV-1 gp41-induced Fusion and Pathogenesis for Anti-viral Therapy," *Current Topics in Medical Chemistry*, 11(24), 2947–2958.

Garnier, L., Parent, L.J., Rovinski, B., Cao, S.-X. & Wills, J.W., 1999, "Identification of Retroviral Late Domains as Determinants of Particle Size," *Journal of Virology*, 73(3), 2309–2320.

Gatanaga, H., Tsukada, K., Honda, H., Tanuma, J., Yazaki, H., Watanabe, T., Honda, M., Teruya, K., Kikuchi, Y. & Oka, S., 2009, "Detection of HIV Type 1 Load by the Roche Cobas TaqMan Assay in Patients with Viral Loads Previously Undetectable by the Roche Cobas Amplicor Monitor," *Clinical Infectious Diseases*, 48(2), 260–262.

Gay, C.L., DeBenedette, M.A., Tcherepanova, I.Y., Gamble, A., Lewis, W.E., Cope, A.B., Kuruc, J.D., McGee, K.S., Kearney, M.F., Coffin, J.M., Archin, N.M., Hicks, C.B., Eron, J.J., Nicolette, C.A. & Margolis, D.M., 2018, "Immunogenicity of AGS-004 Dendritic Cell Therapy in Patients Treated During Acute HIV Infection," *AIDS Research and Human Retroviruses*, 34(1), 111–122.

Gelderblom, H.R., 1991, "Assembly and morphology of HIV: potential effect of structure on viral function," *AIDS*, (5), 617–638.

Gottlieb, M.S., Schroff, R., Schanker, H.M., Weisman, J.D., Fan, P.T., Wolf, R.A. & Saxon, A., 1981, "Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency.," *The New England journal of medicine*, 305(24), 1425–31.

Göttlinger, H.G., 2010, HIV-1 Gag: A Molecular Machine Driving Viral Particle Assembly and Release, *Retrieved July 14, 2020, from Los Alamos, https://www.hiv.lanl.gov/content/sequence/HIV/REVIEWS/GOTTLINGER2001/Gottlinger.html*.

Göttlinger, H.G., Dorfman, T., Sodroski, J.G. & Haseltine, W.A., 1991, "Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release.," *Proceedings of the National Academy of Sciences*, 88(8), 3195–3199.

Göttlinger, H.G., Sodroski, J.G. & Haseltine, W.A., 1989, "Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1.," *Proceedings of the National Academy of Sciences*, 86(15), 5781–5785.

Gounder, K., Oyaro, M., Padayachi, N., Zulu, T.M., Oliveira, T. de, Wylie, J. & Ndung'u, T., 2017, "Complex Subtype Diversity of HIV-1 Among Drug Users in Major Kenyan Cities," *AIDS Research and Human Retroviruses*, 33(5), 500–510.

Greene, W.C. & Peterlin, B.M., 2006, Molecular insights into HIV biology, *Retrieved July 14, 2020, from HIV InSite, http://hivinsite.ucsf.edu/InSite?page=kb-02-01-01#S5X*.

Greub, G., Cozzi-Lepri, A., Ledergerber, B., Staszewski, S., Perrin, L., Miller, V., Francioli, P., Furrer, H., Battegay, M., Vernazza, P., Bernasconi, E., Günthard, H.F., Hirschel, B., Phillips, A.N. & Telenti, A., 2002, "Intermittent and sustained low-level HIV viral rebound in patients receiving potent antiretroviral therapy," *AIDS*, 16(14), 1967–1969.

Günthard, H.F., Wong, J.K., Spina, C.A., Ignacio, C., Kwok, S., Christopherson, C., Hwang, J., Haubrich, R., Havlir, D. & Richman, D.D., 2000, "Effect of Influenza Vaccination on Viral Replication and Immune Response in Persons Infected with Human Immunodeficiency Virus Receiving Potent Antiretroviral Therapy," *The Journal of Infectious Diseases*, 181(2), 522–531.

Gupta, R.K., Abdul-Jawad, S., McCoy, L.E., Mok, H.P., Peppa, D., Salgado, M., Martinez-Picado, J., Nijhuis, M., Wensing, A.M.J., Lee, H., Grant, P., Nastouli, E., Lambert, J., Pace, M., Salasc, F., Monit, C., Innes, A.J., Muir, L., Waters, L., Frater, J., Lever, A.M.L., Edwards, S.G., Gabriel, I.H. & Olavarria, E., 2019, "HIV-1 remission following CCR5Δ32/Δ32 haematopoietic stem-cell transplantation," *Nature*, 568(7751), 244–248.

Halfon, P., Giorgetti, C., Khiri, H., Pénaranda, G., Terriou, P., Porcu-Buisson, G. & Chabert-Orsini, V., 2010, "Semen May Harbor HIV Despite Effective HAART: Another Piece in the Puzzle," *PLoS ONE*, 5(5), e10569.

Halvas, E.K., Joseph, K., Brandt, L.D., Botha, J.C., Sobolewski, M., Jacobs, J.L., Keele, B.F., Kearney, M.F., Coffin, J.M., Rausch, J.W., Guo, S., Wu, X., Hughes, S.H. & Mellors, J.W., 2019, "Nonsuppressible viremia on ART from large cell clones carrying intact proviruses," *CROI 2019,* Abstract 23.

Hardie, D., Korsman, S., Ameer, S., Vojnov, L. & Hsiao, N.Y., 2019, "Reliability of plasma HIV viral load testing beyond 24 hours: Insights gained from a study in a routine diagnostic laboratory," *PLoS ONE*, 14(7).

Harrich, D., Ulich, C. & Gaynor, R.B., 1996, "A Critical Role for the TAR Element in Promoting Efficient Human Immunodeficiency Virus Type 1 Reverse Transcription," *Journal of Virology*, 70(6), 4017–4027.

Harrigan, P.R., Hogg, R.S., Dong, W.W.Y., Yip, B., Wynhoven, B., Woodward, J., Brumme, C.J., Brumme, Z.L., Mo, T., Alexander, C.S. & Montaner, J.S.G., 2005, "Predictors of HIV Drug-Resistance Mutations in a Large Antiretroviral-Naive Cohort Initiating Triple Antiretroviral Therapy," *The Journal of Infectious Diseases*, 191(3), 339–347.

Harrington, P.R., Schnell, G., Letendre, S.L., Ritola, K., Robertson, K., Hall, C., Burch, C.L., Jabara, C.B., Moore, D.T., Ellis, R.J., Price, R.W. & Swanstrom, R., 2009, "Cross-sectional characterization of HIV-1 env compartmentalization in cerebrospinal fluid over the full disease course," *AIDS*, 23(8), 907–915.

Harris, R.S., Bishop, K.N., Sheehy, A.M., Craig, H.M., Petersen-Mahrt, S.K., Watt, I.N., Neuberger, M.S. & Malim, M.H., 2003, "DNA Deamination Mediates Innate Immunity to Retroviral Infection," *Cell*, 113(6), 803–809.

Hassan, A.S., Nabwera, H.M., Mwaringa, S.M., Obonyo, C.A., Sanders, E.J., Rinke de Wit, T.F., Cane, P.A. & Berkley, J.A., 2014, "HIV-1 virologic failure and acquired drug resistance among first-line antiretroviral experienced adults at a rural HIV clinic in coastal Kenya: a cross-sectional study," *AIDS Research and Therapy*, 11(1), 9.

Hatano, H., Somsouk, M., Sinclair, E., Harvill, K., Gilman, L., Cohen, M., Hoh, R., Hunt, P.W., Martin, J.N., Wong, J.K., Deeks, S.G. & Yukl, S.A., 2013, "Comparison of HIV DNA and RNA in gut-associated lymphoid tissue of HIV-infected controllers and noncontrollers," *AIDS*, 27(14), 2255–2260.

Havlir, D. v., 2001, "Prevalence and Predictive Value of Intermittent Viremia with Combination HIV Therapy," *JAMA*, 286(2), 171.

Havlir, D. v, Bassett, R., Diane Levitan, M., Gilbert, P., Tebas, P., Collier, A.C., Hirsch, M.S., Ignacio, C., Jon Condra, B., Gü nthard, H.F., Richman, D.D. & Wong, J.K., 2001, "Prevalence and Predictive Value of Intermittent Viremia With Combination HIV Therapy," *JAMA*, 286(2), 171–179.

Hebberecht, L., Vancoillie, L., Schauvliege, M., Staelens, D., Dauwe, K., Mortier, V. & Verhofstede, C., 2018, "Frequency of occurrence of HIV-1 dual infection in a Belgian MSM population," *PLOS ONE*, 13(4), e0195679.

Heinzinger, N.K., Bukinsky, M.I., Haggerty, S.A., Ragland, A.M., Kewalramani, V., Lee, M.A., Gendelman, H.E., Ratner, L., Stevenson, M. & Emerman, M., 1994, "The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells.," *Proceedings of the National Academy of Sciences*, 91(15), 7311–7315.

Hemelaar, J., 2012, "The origin and diversity of the HIV-1 pandemic.," *Trends in Molecular Medicine*, 18(3), 182–192.

Hemelaar, J., Gouws, E., Ghys, P.D., Osmanov, S. & WHO-UNAIDS, 2011, "Global trends in molecular epidemiology of HIV-1 during 2000– 2007," *AIDS (London, England)*, 25(5), 679–689.

Henderson, L.E., Bowers, M.A., Sowder Ii, R.C., Serabyn, S.A., Johnson, D.G., Bess, J.W., Arthur, L. 0, Bryant, D.K. & 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," *Journal of Virology*, 66(4), 1856–1865.

Henderson, L.J., Reoma, L.B., Kovacs, J.A. & Nath, A., 2019, "Advances toward Curing HIV-1 Infection in Tissue Reservoirs," *Journal of Virology*, 94(3), e00375.

Henrich, T.J., Wood, B.R. & Kuritzkes, D.R., 2012, "Increased Risk of Virologic Rebound in Patients on Antiviral Therapy with a Detectable HIV Load <48 Copies/mL," *PLoS ONE*, 7(11), e50065.

Hermankova, M., Siliciano, J.D., Zhou, Y., Monie, D., Chadwick, K., Margolick, J.B., Quinn, T.C. & Siliciano, R.F., 2003, "Analysis of Human Immunodeficiency Virus Type 1 Gene Expression in Latently Infected Resting CD4+ T Lymphocytes In Vivo," *Journal of Virology*, 77(13), 7383–7392.

Hermans, L.E., Moorhouse, M., Carmona, S., Grobbee, D.E., Hofstra, L.M., Richman, D.D., Tempelman, H.A., Venter, W.D.F. & Wensing, A.M.J., 2018, "Effect of HIV-1 low-level viraemia during antiretroviral therapy on treatment outcomes in WHO-guided South African treatment programmes: a multicentre cohort study," *The Lancet Infectious Diseases*, 18(2), 188–197.

Hernandez-Sanchez, P.G., Guerra-Palomares, S.E., Ramirez-GarciaLuna, J.L., Arguello, J.R., Noyola, D.E. & Garcia-Sepulveda, C.A., 2018, "Prevalence of Drug Resistance Mutations in Protease, *Reverse transcriptase*, and *Integrase* Genes of North Central Mexico HIV Isolates," *AIDS Research and Human Retroviruses*, 34(6), 498–506.

Hiener, B., Horsburgh, B.A., Eden, J.-S., Barton, K., Schlub, T.E., Lee, E., Stockenstrom, S. von, Odevall, L., Milush, J.M., Liegler, T., Sinclair, E., Hoh, R., Boritz, E.A., Douek, D., Fromentin, R., Chomont, N., Deeks, S.G., Hecht, F.M. & Palmer, S., 2017, "Identification of Genetically Intact HIV-1 Proviruses in Specific CD4 + T Cells from Effectively Treated Participants," *Cell Reports*, 21(3), 813–822.

Hill, M., Tachedjian, G. & Mak, J., 2005, "The Packaging and Maturation of the HIV-1 Pol Proteins," *Current HIV Research*, 3(1), 73–85.

Ho, D.D., Rota, T.R., Schooley, R.T., Kaplan, J.C., Allan, J.D., Groopman, J.E., Resnick, L., Felsenstein, D., Andrews, C.A. & Hirsch, M.S., 1985, "Isolation of HTLV-III from Cerebrospinal Fluid and Neural Tissues of Patients with Neurologic Syndromes Related to the Acquired Immunodeficiency Syndrome," *New England Journal of Medicine*, 313(24), 1493–1497.

Ho, Y.C., Shan, L., Hosmane, N.N., Wang, J., Laskey, S.B., Rosenbloom, D.I.S., Lai, J., Blankson, J.N., Siliciano, J.D. & Siliciano, R.F., 2013, "Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure," *Cell*, 155(3), 540–551.

Hocqueloux, L., Avettand-Fènoël, V., Jacquot, S., Prazuck, T., Legac, E., Mélard, A., Niang, M., Mille, C., Moal, G. le, Viard, J.-P. & Rouzioux, C., 2013, "Long-term antiretroviral therapy initiated during primary HIV-1 infection is key to achieving both low HIV reservoirs and normal T cell counts," *Journal of Antimicrobial Chemotherapy*, 68(5), 1169–1178.

Hong, F., Aga, E., Cillo, A., Yates, A.L., Besson, G., Fyne, E., Koontz, D.L., Jennings, C., Zheng, L. & Mellors, J.W., 2016, "Novel assays to measure total cell-associated HIV-1 DNA and RNA.," *Journal of Clinical Microbiology*, 54(4), 902–911.

Hope, T.J. & Trono, D., 2000, Structure, expression and regulation of the HIV genome, *Retrieved July 14, 2020, from HIV InSite, http://hivinsite.ucsf.edu/InSite?page=kb-00&doc=kb-02-01-02*.

Hosmane, N.N., Kwon, K.J., Bruner, K.M., Capoferri, A.A., Beg, S., Rosenbloom, D.I.S., Keele, B.F., Ho, Y.C., Siliciano, J.D. & Siliciano, R.F., 2017, "Proliferation of latently infected CD4+ T cells carrying replication-competent HIV-1: Potential role in latent reservoir dynamics," *Journal of Experimental Medicine*, 214(4), 959–972.

Huber, H.E. & Richardsong, C.C., 1990, "Processing of the Primer for Plus Strand DNA Synthesis by Human Immunodeficiency Virus 1 *Reverse transcriptase*"," *The Journal of Biological Chemistry*, 265(16), 10565–10573.

Hütter, G., Nowak, D., Mossner, M., Ganepola, S., Müßig, A., Allers, K., Schneider, T., Hofmann, J., Kücherer, C., Blau, O., Blau, I.W., Hofmann, W.K. & Thiel, E., 2009, "Long-Term Control of HIV by CCR5 Delta32/Delta32 Stem-Cell Transplantation," *New England Journal of Medicine*, 360(7), 692–698.

Imamichi, H., Dewar, R.L., Adelsberger, J.W., Rehm, C.A., O'Doherty, U., Paxinos, E.E., Fauci, A.S. & Lane, H.C., 2016, "Defective HIV-1 proviruses produce novel protein-coding RNA species in HIV-infected patients on combination antiretroviral therapy," *Proceedings of the National Academy of Sciences*, 113(31), 8783–8788.

Imamichi, H., Natarajan, V., Adelsberger, J.W., Rehm, C.A., Lempicki, R.A., Das, B., Hazen, A., Imamichi, T. & Lane, H.C., 2014, "Lifespan of effector memory CD4+ T cells determined by replication-incompetent integrated HIV-1 provirus," *AIDS*, 28(8), 1091–1099.

Iwabu, Y., Mizuta, H., Kawase, M., Kameoka, M., Goto, T. & Ikuta, K., 2008, "Superinfection of defective human immunodeficiency virus type 1 with different subtypes of wild-type virus efficiently produces infectious variants with the initial viral phenotypes by complementation followed by recombination," *Microbes and Infection*, 10(5), 504–513.

Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J. & Varmus, H.E., 1988, "Characterization of ribosomal frameshifting in HIV-1 gag-pol expression," *Nature*, 331(6153), 280–283.

Jacobs, J.L., Halvas, E.K., Tosiano, M.A. & Mellors, J.W., 2019, "Persistent HIV-1 Viremia on Antiretroviral Therapy: Measurement and Mechanisms," *Frontiers in Microbiology*, 10, 2383.

Jambo, K.C., Banda, D.H., Kankwatira, A.M., Sukumar, N., Allain, T.J., Heyderman, R.S., Russell, D.G. & Mwandumba, H.C., 2014, "Small alveolar macrophages are infected preferentially by HIV and exhibit impaired phagocytic function," *Mucosal Immunology*, 7(5), 1116–1126.

Jarmuz, A., Chester, A., Bayliss, J., Gisbourne, J., Dunham, I., Scott, J. & Navaratnam, N., 2002, "An Anthropoid-Specific Locus of Orphan C to U RNA-Editing Enzymes on Chromosome 22," *Genomics*, 79(3), 285–296.

Jones, K.A. & Peterlin, M.B., 1994, "Control of RNA Initiation and Elongation at the HIV-1 Promoter," *Annual Review of Biochemistry*, 63(1), 717–743.

Jones, L.E. & Perelson, A.S., 2002, "Modeling the Effects of Vaccination on Chronically Infected HIV-Positive Patients," *Journal of Acquired Immune Deficiency Syndromes*, 31(4), 369–377.

Jones, L.E. & Perelson, A.S., 2005, "Opportunistic infection as a cause of transient viremia in chronically infected HIV patients under treatment with HAART," *Bulletin of Mathematical Biology*, 67(6), 1227–1251.

Jones, R.B., Pertea, M., Bruner, K., Pollack, R., Capoferri, A., Siliciano, R. & Ho, Y., 2017, "Cytotoxic lymphocytes may shape the HIV-1 proviral landscape," *CROI 2017*, (Abstract 296).

Jordan, A., Bisgrove, D. & Verdin, E., 2003, "HIV reproducibly establishes a latent infection after acute infection of T cells in vitro," *The EMBO Journal*, 22(8), 1868–1877.

Jordan, A., Defechereux, P. & Verdin, E., 2001, "The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation," *The EMBO Journal*, 20(7), 1726–1738.

Josefsson, L., King, M.S., Makitalo, B., Brännström, J., Shao, W., Maldarelli, F., Kearney, M.F., Hu, W.-S., Chen, J., Gaines, H., Mellors, J.W., Albert, J., Coffin, J.M. & Palmer, S.E., 2011, "Majority of CD4+ T cells from peripheral blood of HIV-1-infected individuals contain only one HIV DNA molecule.," *Proceedings of the National Academy of Sciences of the United States of America*, 108(27), 11199–204.

Josefsson, L., Palmer, S., Faria, N.R., Lemey, P., Casazza, J., Ambrozak, D., Kearney, M., Shao, W., Kottilil, S., Sneller, M., Mellors, J., Coffin, J.M. & Maldarelli, F., 2013, "Single Cell Analysis of Lymph Node Tissue from HIV-1 Infected Patients Reveals that the Majority of CD4+ T-cells Contain One HIV-1 DNA Molecule," *PLoS Pathogens*, 9(6), e1003432.

Josefsson, L., Stockenstrom, S. von, Faria, N.R., Sinclair, E., Bacchetti, P., Killian, M., Epling, L., Tan, A., Ho, T., Lemey, P., Shao, W., Hunt, P.W., Somsouk, M., Wylie, W., Douek, D.C., Loeb, L., Custer, J., Hoh, R., Poole, L., Deeks, S.G., Hecht, F. & Palmer, S., 2013, "The HIV-1 reservoir in eight patients on long-term suppressive antiretroviral therapy is stable with few genetic changes over time.," *Proceedings of the National Academy of Sciences of the United States of America*, 110(51), E4987-96.

Joya, C., Won, S.H., Schofield, C., Lalani, T., Maves, R.C., Kronmann, K., Deiss, R., Okulicz, J., Agan, B.K. & Ganesan, A., 2019, "Persistent Low-level Viremia While on Antiretroviral Therapy Is an Independent Risk Factor for Virologic Failure," *Clinical Infectious Diseases*, 69(12), 2145–2152.

Julg, B., Dee, L., Ananworanich, J., Barouch, D.H., Bar, K., Caskey, M., Colby, D.J., Dawson, L., Dong, K.L., Dubé, K., Eron, J., Frater, J., Gandhi, R.T., Geleziunas, R., Goulder, P., Hanna, G.J., Jefferys, R., Johnston, R., Kuritzkes, D., Li, J.Z., Likhitwonnawut, U., Lunzen, J. van, Martinez-Picado, J., Miller, V., Montaner, L.J., Nixon, D.F., Palm, D., Pantaleo, G., Peay, H., Persaud, D., Salzwedel, J., Salzwedel, K., Schacker, T., Sheikh, V., Søgaard, O.S., Spudich, S., Stephenson, K., Sugarman, J., Taylor, J., Tebas, P., Tiemessen, C.T., Tressler, R., Weiss, C.D., Zheng, L., Robb, M.L., Michael, N.L., Mellors, J.W., Deeks, S.G. & Walker, B.D., 2019, "Recommendations for analytical antiretroviral treatment interruptions in HIV research trials—report of a consensus meeting," *Lancet HIV*, 6(4), e259–e268.

Kaiser, P., Joshi, S.K., Kim, P., Li, P., Liu, H., Rice, A.P., Wong, J.K. & Yukl, S.A., 2017, "Assays for precise quantification of total (including short) and elongated HIV-1 transcripts," *Journal of Virological Methods*, 242, 1–8.

Kandathil, A.J., Durand, C.M., Quinn, J., Cameron, A.M., Thomas, D.L. & Balagopal, A., 2015, "Liver macrophages and HIV persistence," *CROI 2015*, (Abstract 380).

Karin, M. & Ben-Neriah, Y., 2000, "Phosphorylation Meets Ubiquitination: The Control of NF-κB Activity," *Annual Review of Immunology*, 18(1), 621–663.

Katzourakis, A., Tristem, M., Pybus, O.G. & Gifford, R.J., 2007, "Discovery and analysis of the first endogenous lentivirus," *Proceedings of the National Academy of Sciences*, 104(15), 6261–6265.

Kauder, S.E., Bosque, A., Lindqvist, A., Planelles, V. & Verdin, E., 2009, "Epigenetic Regulation of HIV-1 Latency by Cytosine Methylation," *PLoS Pathogens*, 5(6), e1000495.

Kearney, M., Palmer, S., Maldarelli, F., Shao, W., Polis, M.A., Mican, J., Rock-Kress, D., Margolick, J.B., Coffin, J.M. & Mellors, J.W., 2008, "Frequent polymorphism at drug resistance sites in HIV-1 protease and *reverse transcriptase*," *AIDS*, 22(4), 497–501.

Kearney, M.F., Spindler, J., Shao, W., Yu, S., Anderson, E.M., O'Shea, A., Rehm, C., Poethke, C., Kovacs, N., Mellors, J.W., Coffin, J.M. & Maldarelli, F., 2014, "Lack of Detectable HIV-1 Molecular Evolution during Suppressive Antiretroviral Therapy," *PLoS Pathogens*, 10(3), e1004010.

Kearney, M.F., Wiegand, A., Shao, W., Coffin, J.M., Mellors, J.W., Lederman, M., Gandhi, R.T., Keele, B.F. & Li, J.Z., 2016, "Origin of Rebound Plasma HIV Includes Cells with Identical Proviruses That Are Transcriptionally Active before Stopping of Antiretroviral Therapy," *Journal of Virology*, 90(3), 1369–1376.

Kearney, M.F., Wiegand, A., Shao, W., McManus, W.R., Bale, M.J., Luke, B., Maldarelli, F., Mellors, J.W. & Coffin, J.M., 2017, "Ongoing HIV Replication During ART Reconsidered," *Open Forum Infectious Diseases*, 4(3).

Kelley, C.F., Sullivan, S.T., Lennox, J.L., Evans-Strickfaden, T. & Hart, C.E., 2010, "Lack of effect of compartmentalized drug resistance mutations on HIV-1 pol divergence in antiretroviral-experienced women," *AIDS*, 24(9), 1361–1366.

Kemnic, T.R. & Gulick, P.G., 2020, "HIV Antiretroviral Therapy," *StatPearls, Treasure Island, Available from: https://www.ncbi.nlm.nih.gov/books/NBK513308/*.

Khoury, G., Fromentin, R., Solomon, A., Hartogensis, W., Killian, M., Hoh, R., Somsouk, M., Hunt, P.W., Girling, V., Sinclair, E., Bacchetti, P., Anderson, J.L., Hecht, F.M., Deeks, S.G., Cameron, P.U., Chomont, N. & Lewin, S.R., 2017, "Human Immunodeficiency Virus Persistence and T-Cell Activation in Blood, Rectal, and Lymph Node Tissue in Human Immunodeficiency Virus–Infected Individuals Receiving Suppressive Antiretroviral Therapy," *The Journal of Infectious Diseases*, 215(6), 911–919.

Kim, S., Byrn, R., Groopman, J. & Baltimore12, D., 1989, "Temporal Aspects of DNA and RNA Synthesis during Human Immunodeficiency Virus Infection: Evidence for Differential Gene Expression," *Journal of Virology*, 63(9), 3708–3713.

Kiweewa, F., Esber, A., Musingye, E., Reed, D., Crowell, T.A., Cham, F., Semwogerere, M., Namagembe, R., Nambuya, A., Kafeero, C., Tindikahwa, A., Eller, L.A., Millard, M., Gelderblom, H.C., Keshinro, B., Adamu, Y., Maswai, J., Owuoth, J., Sing'oei, V.C., Maganga, L., Bahemana, E., Khamadi, S., Robb, M.L., Ake, J.A., Polyak, C.S. & Kibuuka, H., 2019, "HIV virologic failure and its predictors among HIV-infected adults on antiretroviral therapy in the African Cohort Study," *PLOS ONE*, 14(2), e0211344.

Klatt, N.R., Bosinger, S.E., Peck, M., Richert-Spuhler, L.E., Heigele, A., Gile, J.P., Patel, N., Taaffe, J., Julg, B., Camerini, D., Torti, C., Martin, J.N., Deeks, S.G., Sinclair, E., Hecht, F.M., Lederman, M.M., Paiardini, M., Kirchhoff, F., Brenchley, J.M., Hunt, P.W. & Silvestri, G., 2014, "Limited HIV Infection of Central Memory and Stem Cell Memory CD4+ T Cells Is Associated with Lack of Progression in Viremic Individuals," *PLoS Pathogens*, 10(8), e1004345.

Klaver, B. & Berkhout, B., 1994, "Comparison of 5' and 3' Long Terminal Repeat Promoter Function in Human Immunodeficiency Virus," *Journal of Virology*, 68(6), 3830–3840.

Kolber, M.A., Gabr, A.H., La Rosa, A. de, Glock, J.A., Jayaweera, D., Miller, N. & Dickinson, G.M., 2002, "Genotypic analysis of plasma HIV-1 RNA after influenza vaccination of patients with previously undetectable viral loads," *AIDS*, 16(4), 537–542.

Korber, B.T.M., Kunstman, K.J., Patterson, B.K., Furtado, M., Mcevilly, M.M., Levy, R. & Wolinsky3, S.M., 1994, "Genetic Differences between Blood-and Brain-Derived Viral Sequences from Human Immunodeficiency Virus Type 1-Infected Patients: Evidence of Conserved Elements in the V3 Region of the Envelope Protein of Brain-Derived Sequences," *Journal of Virology*, 68(11), 7467–7481.

Kovacs, A., Wasserman, S.S., Burns, D., Wright, D.J., Cohn, J., Landay, A., Weber, K., Cohen, M., Levine, A., Minkoff, H., Miotti, P., Palefsky, J., Young, M. & Reichelderfer, P., 2001, "Determinants of HIV-1 shedding in the genital tract of women," *The Lancet*, 358(9293), 1593–1601.

Koyanagi, Y., Miles, S., Mitsuyasu, R., Merrill, J., Vinters, H. & Chen, I., 1987, "Dual infection of the central nervous system by AIDS viruses with distinct cellular tropisms," *Science*, 236(4803), 819–822.

Kumar, S., Stecher, G. & Tamura, K., 2016, "MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets," *Molecular Biology and Evolution*, 33(7), 1870–1874.

Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J. & Hendrickson, W.A., 1998, "Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody," *Nature*, 393(6686), 648–659.

Laanani, M., Ghosn, J., Essat, A., Melard, A., Seng, R., Gousset, M., Panjo, H., Mortier, E., Girard, P.-M., Goujard, C., Meyer, L. & Rouzioux, C., 2015, "Impact of the Timing of Initiation of Antiretroviral Therapy During Primary HIV-1 Infection on the Decay of Cell-Associated HIV-DNA," *Clinical Infectious Diseases*, 60(11), 1715–1721.

Lafeuillade, A., Solas, C., Halfon, P., Chadapaud, S., Hittinger, G. & Lacarelle, B., 2002, "Differences in the detection of three HIV-1 protease inhibitors in non-blood compartments: Clinical correlations," *HIV Clinical Trials*, 3(1), 27–35.

Landau, N.R., Warton, M. & Littman, D.R., 1988, "The envelope glycoprotein of the human immunodeficiency virus binds to the immunoglobulin-like domain of CD4," *Nature*, 334(6178), 159–162.

Laprise, C., Pokomandy, A. de, Baril, J.-G., Dufresne, S. & Trottier, H., 2013, "Virologic Failure Following Persistent Low-level Viremia in a Cohort of HIV-Positive Patients: Results From 12 Years of Observation," *Clinical Infectious Diseases*, 57(10), 1489–1496.

Laskey, S.B. & Siliciano, R.F., 2014, "A mechanistic theory to explain the efficacy of antiretroviral therapy," *Nature Reviews Microbiology*, 12(11), 772–780.

Lassen, K., Han, Y., Zhou, Y., Siliciano, J. & Siliciano, R.F., 2004, "The multifactorial nature of HIV-1 latency," *Trends in Molecular Medicine*, 10(11), 525–531.

Launay, O., Tod, M., Tschöpe, I., Si-Mohamed, A., Bélarbi, L., Charpentier, C., Goujard, C., Taburet, A.-M., Lortholary, O., Leroy, V. & Bélec, L., 2011, "Residual HIV-1 RNA and HIV-1 DNA production in the genital tract reservoir of women treated with HAART: the prospective ANRS EP24 GYNODYN study," *Antiviral Therapy*, 16(6), 843–852.

Lecatsas, G., Houff, S., Macher, A., Gelman, E., Steis, R., Reichert, C., Masur, H. & Sever, J.L., 1985, "Retrovirus-Like Particles in Salivary Glands, Prostate and Testes of Aids Patients," *Experimental Biology and Medicine*, 178(4), 653–655.

Lecossier, D., Bouchonnet, F., Clavel, F. & Hance, A.J., 2003, "Hypermutation of HIV-1 DNA in the Absence of the Vif Protein," *Science*, 300(5622), 1112–1112.

Lee, M.S. & Craigie, R., 1998, "A previously unidentified host protein protects retroviral DNA from autointegration," *Proceedings of the National Academy of Sciences*, 95(4), 1528–1533.

Lee, P.K., Kieffer, T.L., Siliciano, R.F. & Nettles, R.E., 2006, "HIV-1 viral load blips are of limited clinical significance," *Journal of Antimicrobial Chemotherapy*, 57(5), 803–805.

Lenasi, T., Contreras, X. & Peterlin, B.M., 2008, "Transcriptional Interference Antagonizes Proviral Gene Expression to Promote HIV Latency," *Cell Host & Microbe*, 4(2), 123–133.

Levy, J.A., 1993, "Pathogenesis of human immunodeficiency virus infection.," *Microbiological reviews*, 57(1), 183–289.

Levy, J.A., 2011, "Virus-Host Interactions in HIV Pathogenesis," *Advances in Dental Research*, 23(1), 13–18.

Levy, J.A., Shimabukuro, J., Hollander, H., Mills, J. & Kaminsky, L., 1985, "Isolation of AIDS-associated retroviruses from cerebrospinal fluid and brain of patients with neurological symptoms," *Lancet*, 2(8455), 586–588.

Lévy, Y., Durier, C., Lascaux, A.-S., Meiffrédy, V., Gahéry-Ségard, H., Goujard, C., Rouzioux, C., Resch, M., Guillet, J.-G., Kazatchkine, M., Delfraissy, J.-F. & Aboulker, J.-P., 2006, "Sustained control of viremia following therapeutic immunization in chronically HIV-1-infected individuals," *AIDS*, 20(3), 405–413.

Lévy, Y., Gahéry-Ségard, H., Durier, C., Lascaux, A.-S., Goujard, C., Meiffrédy, V., Rouzioux, C., Habib, R. el, Beumont-Mauviel, M., Guillet, J.-G., Delfraissy, J.-F., Aboulker, J.-P. & ANRS 093 Study Group, 2005, "Immunological and virological efficacy of a therapeutic immunization combined with interleukin-2 in chronically HIV-1 infected patients," *AIDS*, 19(3), 279–286.

Li, B., Gladden, A.D., Altfeld, M., Kaldor, J.M., Cooper, D.A., Kelleher, A.D. & Allen, T.M., 2007, "Rapid Reversion of Sequence Polymorphisms Dominates Early Human Immunodeficiency Virus Type 1 Evolution," *Journal of Virology*, 81(1), 193–201.

Li, J.Z., Etemad, B., Ahmed, H., Aga, E., Bosch, R.J., Mellors, J.W., Kuritzkes, D.R., Lederman, M.M., Para, M. & Gandhi, R.T., 2016, "The size of the expressed HIV reservoir predicts timing of viral rebound after treatment interruption," *AIDS*, 30(3), 343–353.

Li, L., Liang, S., Chen, L., Liu, W., Li, H., Liu, Y., Bao, Z., Wang, Z., Zhuang, D., Liu, S. & Li, J., 2010, "Genetic characterization of 13 subtype CRF01_AE near full-length genomes in Guangxi, China," *AIDS*, 26(6), 699–704.

Li, L., Yoder, K., Hansen, M.S.T., Olvera, J., Miller, M.D. & Bushman, F.D., 2000, "Retroviral cDNA Integration: Stimulation by HMG I Family Proteins," *Journal of Virology*, 74(23), 10965–10974.

Li, M. & Craigie, R., 2009, "Nucleoprotein complex intermediates in HIV-1 integration," *Methods*, 47(4), 237–242.

Li, S., Hill, C.P., Sundquist, W.I. & Finch, J.T., 2000, "Image reconstructions of helical assemblies of the HIV-1 CA protein," *Nature*, 407(6802), 409–413.

Liao, H.-K., Gu, Y., Diaz, A., Marlett, J., Takahashi, Y., Li, M., Suzuki, K., Xu, R., Hishida, T., Chang, C.-J., Esteban, C.R., Young, J. & Belmonte, J.C.I., 2015, "Use of the CRISPR/Cas9 system as an intracellular defense against HIV-1 infection in human cells," *Nature Communications*, 6(1), 6413.

Liddament, M.T., Brown, W.L., Schumacher, A.J. & Harris, R.S., 2004, "APOBEC3F Properties and Hypermutation Preferences Indicate Activity against HIV-1 In Vivo," *Current Biology*, 14(15), 1385–1391.

Lima, K., Leal, É., Cavalcanti, A.M.S., Salustiano, D.M., Medeiros, L.B. de, Silva, S.P. da & Lacerda, H.R., 2017, "Increase in human immunodeficiency virus 1 diversity and detection of various subtypes and recombinants in north-eastern Brazil," *Journal of Medical Microbiology*, 66(4), 526–535.

Lin, C.-W. & Engelman, A., 2003, "The Barrier-to-Autointegration Factor Is a Component of Functional Human Immunodeficiency Virus Type 1 Preintegration Complexes," *Journal of Virology*, 77(8), 5030–5036.

Lin, X., Irwin, D., Kanazawa, S., Huang, L., Romeo, J., Yen, T.S.B. & Peterlin, B.M., 2003, "Transcriptional Profiles of Latent Human Immunodeficiency Virus in Infected Individuals: Effects of Tat on the Host and Reservoir," *Journal of Virology*, 77(15), 8227–8236.

Liu, R., Paxton, W.A., Choe, S., Ceradini, D., Martin, S.R., Horuk, R., MacDonald, M.E., Stuhlmann, H., Koup, R.A. & Landau, N.R., 1996, "Homozygous Defect in HIV-1 Coreceptor Accounts for Resistance of Some Multiply-Exposed Individuals to HIV-1 Infection," *Cell*, 86(3), 367–377.

Liu, T.F. & Shafer, R.W., 2006, "Web Resources for HIV Type 1 Genotypic-Resistance Test Interpretation," *Clinical Infectious Diseases*, 42(11), 1608–1618.

Lorenzi, J.C.C., Cohen, Y.Z., Cohn, L.B., Kreider, E.F., Barton, J.P., Learn, G.H., Oliveira, T., Lavine, C.L., Horwitz, J.A., Settler, A., Jankovic, M., Seaman, M.S., Chakraborty, A.K., Hahn, B.H., Caskey, M. & Nussenzweig, M.C., 2016, "Paired quantitative and qualitative assessment of the replication-competent HIV-1 reservoir and comparison with integrated proviral DNA.," *Proceedings of the National Academy of Sciences of the United States of America*, 113(49), E7908–E7916.

Los Alamos, 2017, HIV and SIV nomenclature, *Retrieved July 13, 2020, from https://www.hiv.lanl.gov/content/sequence/HIV/MAP/landmark.html*.

Lu, X., Yu, H., Liu, S.-H., Brodsky, F.M. & Peterlin, B.M., 1998, "Interactions between HIV1 Nef and Vacuolar ATPase Facilitate the Internalization of CD4," *Immunity*, 8(5), 647–656.

Luciw, P.A., 1996, Human immunodeficiency viruses and their replication, 3rd edn., Lippencott-Raven, Philadelphia.

Luria, S., Chambers, I. & Berg, P., 1991, "Expression of the type 1 human immunodeficiency virus Nef protein in T cells prevents antigen receptor-mediated induction of interleukin 2 mRNA.," *Proceedings of the National Academy of Sciences*, 88(12), 5326–5330.

Maggiolo, F., Ravasio, L., Ripamonti, D., Gregis, G., Quinzan, G., Arici, C., Airoldi, M. & Suter, F., 2005, "Similar Adherence Rates Favor Different Virologic Outcomes for Patients Treated with Nonnucleoside Analogues or Protease Inhibitors," *Clinical Infectious Diseases*, 40(1), 158–163.

Maldarelli, F., 2016, "The role of HIV integration in viral persistence: no more whistling past the proviral graveyard," *Journal of Clinical Investigation*, 126(2), 438–447.

Maldarelli, F., Wu, X., Su, L., Simonetti, F.R., Shao, W., Hill, S., Spindler, J., Ferris, A.L., Mellors, J.W., Kearney, M.F., Coffin, J.M. & Hughes, S.H., 2014, "Specific HIV integration sites are linked to clonal expansion and persistence of infected cells," *Science*, 345(6193), 179–183.

Malet, I., Subra, F., Charpentier, C., Collin, G., Descamps, D., Calvez, V., Marcelin, A.-G. & Delelis, O., 2017, "Mutations Located outside the Integrase Gene Can Confer Resistance to HIV-1 Integrase Strand Transfer Inhibitors," *mBio*, 8(5), e00922-17.

Malim, M.H., Hauber, J., Le, S.-Y., Maizel, J. v. & Cullen, B.R., 1989, "The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA," *Nature*, 338(6212), 254–257.

Malnati, M.S., Scarlatti, G., Gatto, F., Salvatori, F., Cassina, G., Rutigliano, T., Volpi, R. & Lusso, P., 2008, "A universal real-time PCR assay for the quantification of group-M HIV-1 proviral load," *Nature Protocols*, 3(7), 1240–1248.

Mangeat, B., Turelli, P., Caron, G., Friedli, M., Perrin, L. & Trono, D., 2003, "Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts," *Nature*, 424(6944), 99–103.

Marcelin, A.-G., Tubiana, R., Lambert-Niclot, S., Lefebvre, G., Dominguez, S., Bonmarchand, M., Vauthier-Brouzes, D., Marguet, F., Mousset-Simeon, N., Peytavin, G. & Poirot, C., 2008, "Detection of HIV-1 RNA in seminal plasma samples from treated patients with undetectable HIV-1 RNA in blood plasma," *AIDS*, 22(13), 1677–1679.

Mariani, R., Chen, D., Schröfelbauer, B., Navarro, F., König, R., Bollman, B., Münk, C., Nymark-McMahon, H. & Landau, N.R., 2003, "Species-Specific Exclusion of APOBEC3G from HIV-1 Virions by Vif," *Cell*, 114(1), 21–31.

Marin, M., Rose, K.M., Kozak, S.L. & Kabat, D., 2003, "HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation," *Nature Medicine*, 9(11), 1398–1403.

Marras, D., Bruggeman, L.A., Gao, F., Tanji, N., Mansukhani, M.M., Cara, A., Ross, M.D., Gusella, G.L., Benson, G., D'Agati, V.D., Hahn, B.H., Klotman, M.E. & Klotman, P.E., 2002, "Replication and compartmentalization of HIV-1 in kidney epithelium of patients with HIV-associated nephropathy," *Nature Medicine*, 8(5), 522–526.

Martin, M., Cacho, E. del, Codina, C., Tuset, M., Lazzari, E. de, Mallolas, J., Miró, J.-M., Gatell, J.M. & Ribas, J., 2008, "Relationship between Adherence Level, Type of the Antiretroviral Regimen, and Plasma HIV Type 1 RNA Viral Load: A Prospective Cohort Study," *AIDS Research and Human Retroviruses*, 24(10), 1263–1268.

Martinson, J.J., Chapman, N.H., Rees, D.C., Liu, Y.-T. & Clegg, J.B., 1997, "Global distribution of the CCR5 gene 32-basepair deletion," *Nature Genetics*, 16(1), 100–103.

Massanella, M. & Richman, D.D., 2016, "Measuring the latent reservoir in vivo," *Journal of Clinical Investigation*, 126(2), 464–472.

McKinnon, J.E., Delgado, R., Pulido, F., Shao, W., Arribas, J.R. & Mellors, J.W., 2011, "Single genome sequencing of HIV-1 gag and protease resistance mutations at virologic failure during the OK04 trial of simplified versus standard maintenance therapy," *Antiviral Therapy*, 16(5), 725–732.

McManus, W.R., Bale, M.J., Spindler, J., Wiegand, A., Musick, A., Patro, S.C., Sobolewski, M.D., Musick, V.K., Anderson, E.M., Cyktor, J.C., Halvas, E.K., Shao, W., Wells, D., Wu, X., Keele, B.F., Milush, J.M., Hoh, R., Mellors, J.W., Hughes, S.H., Deeks, S.G., Coffin, J.M. & Kearney, M.F., 2019, "HIV-1 in lymph nodes is maintained by cellular proliferation during antiretroviral therapy," *Journal of Clinical Investigation*, 129(11), 4629–4642.

McNamara, L.A., Onafuwa-Nuga, A., Sebastian, N.T., Riddell, J., Bixby, D. & Collins, K.L., 2013, "CD133+ Hematopoietic Progenitor Cells Harbor HIV Genomes in a Subset of Optimally Treated People With Long-Term Viral Suppression," *Journal of Infectious Diseases*, 207(12), 1807–1816.

Méndez, C., Ledger, S., Petoumenos, K., Ahlenstiel, C. & Kelleher, A.D., 2018, "RNA-induced epigenetic silencing inhibits HIV-1 reactivation from latency," *Retrovirology*, 15(1), 67.

Miller, M., Jaskólski, M., Rao, J.K.M., Leis, J. & Wlodawer, A., 1989, "Crystal structure of a retroviral protease proves relationship to aspartic protease family," *Nature*, 337(6207), 576–579.

Miller, M.D., Farnet, C.M. & Bushman, F.D., 1997, "Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition," *Journal of virology*, 71(7), 5382–5390.

Miller, M.D., Warmerdam, M.T., Gaston, I., Greene, W.C. & Feinberg, M.B., 1994, "The human immunodeficiency virus-1 nef gene product: a positive factor for viral infection and replication in primary lymphocytes and macrophages," *The Journal of Experimental Medicine*, 179(1), 101–113.

Mitsuya, Y., Winters, M. A., Fessel, W. J., Rhee, S. Y., Slome, S., Flamm, J., Horberg, M., Hurley, L., Klein, D., & Shafer, R. W., 2006, "HIV-1 drug resistance genotype results in patients with plasma samples with HIV-1 RNA levels less than 75 copies/ml," *Journal of Acquired Immune Deficiency Syndromes*, 43(1), 56–59.

Mousseau, G., Kessing, C.F., Fromentin, R., Trautmann, L., Chomont, N. & Valente, S.T., 2015, "The Tat Inhibitor Didehydro-Cortistatin A Prevents HIV-1 Reactivation from Latency," *mBio*, 6(4).

Muciaccia, B., Filippini, A., Ziparo, E., Colelli, F., Baroni, C.D. & Stefanini, M., 1998, "Testicular germ cells of HIV-seropositive asymptomatic men are infected by the virus," *Journal of Reproductive Immunology*, 41(1–2), 81–93.

Muciaccia, B., Uccini, S., Filippini, A., Ziparo, E., Paraire, F., Baroni, C.D. & Stefanini, M., 1998, "Presence and cellular distribution of HIV in the testes of seropositive subjects: an evaluation by in situ PCR hybridization," *The FASEB Journal*, 12(2), 151–163.

Muir, D., White, D., King, J., Verlander, N., & Pillay, D., 2000, "Predictive value of the ultrasensitive HIV viral load assay in clinical practice," *Journal of Medical Virology*, 61(4), 411–416.

Nachega, J.B., Hislop, M., Dowdy, D.W., Chaisson, R.E., Regensberg, L. & Maartens, G., 2007, "Adherence to Nonnucleoside *Reverse transcriptase* Inhibitor–Based HIV Therapy and Virologic Outcomes," *Annals of Internal Medicine*, 146(8), 564.

Nachega, J.B., Marconi, V.C., Zyl, G.U. van, Gardner, E.M., Preiser, W., Hong, S.Y., Mills, E.J. & Gross, R., 2011, "HIV Treatment Adherence, Drug Resistance, Virologic Failure," *Infectious Disorders - Drug Targets*, 11(2), 167–174.

Nasta, P., Castelnuovo, F. & Paraninfo, G., 2005, "To maintain or to switch to HAART in heavily pre-treated patients with lowlevel viremia," *12th Conference on Retroviruses and Opportunistic Infections (CROI 2005)*, (Abstract 605).

Navia, M.A., Fitzgerald, P.M.D., McKeever, B.M., Leu, C.-T., Heimbach, J.C., Herber, W.K., Sigal, I.S., Darke, P.L. & Springer, J.P., 1989, "Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1," *Nature*, 337(6208), 615–620.

Nettles, R.E., Kieffer, T.L., Kwon, P., Daphne Monie, B., Han, Y., Teresa Parsons, B., Cofrancesco, J., Gallant, J.E., Thomas Quinn, M.C., Jackson, B., Flexner, C., Carson, K., Stuart Ray, S., Persaud, D. & Siliciano, R.F., 2005, "Intermittent HIV-1 Viremia (Blips) and Drug Resistance in Patients Receiving HAART," *JAMA*, 293(7), 817–829.

Ngcapu, S., Theys, K., Libin, P., Marconi, V., Sunpath, H., Ndung'u, T. & Gordon, M., 2017, "Characterization of Nucleoside *Reverse transcriptase* Inhibitor-Associated Mutations in the RNase H Region of HIV-1 Subtype C Infected Individuals," *Viruses*, 9(11), 330.

NIAID, 2018, HIV Replication Cycle, *Retrieved July 15, 2020, from National Institue of Allergy and Infectious Diseases, https://www.niaid.nih.gov/diseases-conditions/hiv-replication-cycle*.

Niedrig, M., Gelderblom, H.R., Pauli, G., Marz, J., Bickhard, H., Wolf, H. & Modrow, S., 1994, "Inhibition of infectious human immunodeficiency virus type 1 particle formation by Gag protein-derived peptides," *Journal of General Virology*, 75(6), 1469–1474.

Nolan, D., Lamers, S., Rose, R., Dollar, J., Salemi, M. & McGrath, M., 2017, "Single Genome Sequencing of Expressed and Proviral HIV-1 Envelope Glycoprotein 120 (gp120) and nef Genes," *Bio-protocol*, 7(12), e2334.

Nottet, H.S.L.M., Dijk, S.J. van, Fanoy, E.B., Goedegebuure, I.W., Jong, D. de, Vrisekoop, N., Baarle, D. van, Boltz, V., Palmer, S., Borleffs, J.C.C. & Boucher, C.A.B., 2009, "HIV-1 Can Persist in Aged Memory CD4+ T Lymphocytes With Minimal Signs of Evolution After 8.3 Years of Effective Highly Active Antiretroviral Therapy," *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 50(4), 345–353.

Nunnari, G., Sullivan, J., Xu, Y., Nyirjesy, P., Kulkosky, J., Cavert, W., Frank, I. & Pomerantz, R.J., 2005, "HIV Type 1 Cervicovaginal Reservoirs in the Era of HAART," *AIDS Research and Human Retroviruses*, 21(8), 714–718.

Ochodo, E.A., Kakourou, A., Mallett, S. & Deeks, J.J., 2018, "Point-of-care tests detecting HIV nucleic acids for diagnosis of HIV infection in infants and children aged 18 months or less," *Cochrane Database of Systematic Reviews*, (11).

Öhagen, Åsa & Gabuzda, D., 2000, "Role of Vif in Stability of the Human Immunodeficiency Virus Type 1 Core," *Journal of Virology*, 74(23), 11055–11066.

O'Neil, P.K., Sun, G., Yu, H., Ron, Y., Dougherty, J.P. & Preston, B.D., 2002, "Mutational Analysis of HIV-1 Long Terminal Repeats to Explore the Relative Contribution of *Reverse transcriptase* and RNA Polymerase II to Viral Mutagenesis," *Journal of Biological Chemistry*, 277(41), 38053–38061.

Ono, A., 2009, "HIV-1 assembly at the plasma membrane: Gag trafficking and localization," *Future Virology*, 4(3), 241–257.

Ono, A. & Freed, E.O., 2001, "Plasma membrane rafts play a critical role in HIV-1 assembly and release," *Proceedings of the National Academy of Sciences*, 98(24), 13925–13930.

Ooms, M., Abbink, T.E.M., Pham, C. & Berkhout, B., 2007, "Circularization of the HIV-1 RNA genome," *Nucleic Acids Research*, 35(15), 5253–5261.

O'Shea, S., Chrystie, I., Cranston, R., Mullen, J., Corbett, K., Murphy, G., Parry, J. V., De Ruiter, A., & Banatvala, J., 2000, "Problems in the interpretation of HIV-1 vital load assays using commercial reagents," *Journal of Medical Virology*, 61(2), 187–194.

Oster, A.M., Switzer, W.M., Hernandez, A.L., Saduvala, N., Wertheim, J.O., Nwangwu-Ike, N., Ocfemia, M.C., Campbell, E. & Hall, H.I., 2017, "Increasing HIV-1 subtype diversity in seven states, United States, 2006–2013," *Annals of Epidemiology*, 27(4), 244-251.

Overbaugh, J., Anderson, R.J., Ndinya-Achola, J.O. & Kreiss, J.K., 1996, "Distinct but Related Human Immunodeficiency Virus Type 1 Variant Populations in Genital Secretions and Blood," *AIDS Research and Human Retroviruses*, 12(2), 107–115.

Özen, A., Lin, K.-H., Kurt Yilmaz, N. & Schiffer, C.A., 2014, "Structural basis and distal effects of Gag substrate coevolution in drug resistance to HIV-1 protease," *Proceedings of the National Academy of Sciences*, 111(45), 15993–15998.

Pallikkuth, S., Sharkey, M., Babic, D.Z., Gupta, S., Stone, G.W., Fischl, M.A., Stevenson, M. & Pahwa, S., 2016, "Peripheral T Follicular Helper Cells Are the Major HIV Reservoir within Central Memory CD4 T Cells in Peripheral Blood from Chronically HIV-Infected Individuals on Combination Antiretroviral," 90(6), 2718–2728.

Palmer, S., Kearney, M., Maldarelli, F., Halvas, E.K., Bixby, C.J., Bazmi, H., Rock, D., Falloon, J., Davey, R.T., Dewar, R.L., Metcalf, J.A., Hammer, S., Mellors, J.W. & Coffin, J.M., 2005, "Multiple, Linked Human Immunodeficiency Virus Type 1 Drug Resistance Mutations in Treatment-Experienced Patients Are Missed by Standard Genotype Analysis," *Journal of Clinical Microbiology*, 43(1), 406–413.

Palmer, S., Maldarelli, F., Wiegand, A., Bernstein, B., Hanna, G.J., Brun, S.C., Kempf, D.J., Mellors, J.W., Coffin, J.M. & King, M.S., 2008, "Low-level viremia persists for at least 7 years in patients on suppressive antiretroviral therapy," *Proceedings of the National Academy of Sciences*, 105(10), 3879–3884.

Pandori, M.W., Fitch, N.J.S., Craig, H.M., Richman, D.D., Spina, C.A. & Guatelli, J.C., 1996, "Producer-Cell Modification of Human Immunodeficiency Virus Type 1: Nef Is a Virion Protein," *Journal of Virology*, 70(7), 4283–4290.

Paranjpe, S., Craigo, J., Patterson, B., Ding, M., Barroso, P., Harrison, L., Montelaro, R. & Gupta, P., 2002, "Subcompartmentalization of HIV-1 Quasispecies between Seminal Cells and Seminal Plasma Indicates Their Origin in Distinct Genital Tissues," *AIDS Research and Human Retroviruses*, 18(17), 1271–1280.

Parkin, N.T., Chamorro, M. & Varmus, H.E., 1992, "Human Immunodeficiency Virus Type 1 gag-pol Frameshifting Is Dependent on Downstream mRNA Secondary Structure: Demonstration by Expression In Vivo," *Journal of Virology*, 66(8), 5147–5151.

Pasternak, A.O., Adema, K.W., Bakker, M., Jurriaans, S., Berkhout, B., Cornelissen, M. & Lukashov, V. v., 2008, "Highly Sensitive Methods Based on Seminested Real-Time Reverse Transcription-PCR for Quantitation of Human Immunodeficiency Virus Type 1 Unspliced and Multiply Spliced RNA and Proviral DNA," *Journal of Clinical Microbiology*, 46(7), 2206–2211.

Patro, S.C., Brandt, L.D., Bale, M.J., Halvas, E.K., Joseph, K.W., Shao, W., Wu, X., Guo, S., Murrell, B., Wiegand, A., Spindler, J., Raley, C., Hautman, C., Sobolewski, M., Fennessey, C.M., Hu, W.-S., Luke, B., Hasson, J.M., Niyongabo, A., Capoferri, A.A., Keele, B.F., Milush, J., Hoh, R., Deeks, S.G., Maldarelli, F., Hughes, S.H., Coffin, J.M., Rausch, J.W., Mellors, J.W. & Kearney, M.F., 2019, "Combined HIV-1 sequence and integration site analysis informs viral dynamics and allows reconstruction of replicating viral ancestors," *Proceedings of the National Academy of Sciences*, 116(51), 25891–25899.

Peel, S., Macheboeuf, P., Martinelli, N. & Weissenhorn, W., 2011, "Divergent pathways lead to ESCRT-III-catalyzed membrane fission," *Trends in Biochemical Sciences*, 36(4), 199–210.

Peluso, M.J., Bacchetti, P., Ritter, K.D., Beg, S., Lai, J., Martin, J.N., Hunt, P.W., Henrich, T.J., Siliciano, J.D., Siliciano, R.F., Laird, G.M. & Deeks, S.G., 2020, "Differential decay of intact and defective proviral DNA in HIV-1–infected individuals on suppressive antiretroviral therapy," *JCI Insight*, 5(4), e132997.

Pereira, L.A., Bentley, K., Peeters, A., Churchill, M.J. & Deacon, N.J., 2000, "A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter," *Nucleic Acids Research*, 28(3), 663–668.

Perelson, A.S., Essunger, P., Cao, Y., Vesanen, M., Hurley, A., Saksela, K., Markowitz, M. & Ho, D.D., 1997, "Decay characteristics of HIV-1-infected compartments during combination therapy," *Nature*, 387(6629), 188–191.

Phillips, A.N., Venter, F., Havlir, D., Pozniak, A., Kuritzkes, D., Wensing, A., Lundgren, J.D., Luca, A. de, Pillay, D., Mellors, J., Cambiano, V., Bansi-Matharu, L., Nakagawa, F., Kalua, T., Jahn, A., Apollo, T., Mugurungi, O., Clayden, P., Gupta, R.K., Barnabas, R., Revill, P., Cohn, J., Bertagnolio, S. & Calmy, A., 2019, "Risks and benefits of dolutegravir-based antiretroviral drug regimens in sub-Saharan Africa: a modelling study," *The Lancet HIV*, 6(2), e116–e127.

Piketty, C., Weiss, L., Assoumou, L., Burgard, M., Mélard, A., Ragnaud, J.-M., Bentata, M., Girard, P.-M., Rouzioux, C. & Costagliola, D., 2010, "A high HIV DNA level in PBMCs at antiretroviral treatment interruption predicts a shorter time to treatment resumption, independently of the CD4 nadir," *Journal of Medical Virology*, 82(11), 1819–1828.

Pinzone, M.R., VanBelzen, D.J., Weissman, S., Bertuccio, M.P., Cannon, L., Venanzi-Rullo, E., Migueles, S., Jones, R.B., Mota, T., Joseph, S.B., Groen, K., Pasternak, A.O., Hwang, W.-T., Sherman, B., Vourekas, A., Nunnari, G. & O'Doherty, U., 2019, "Longitudinal HIV sequencing reveals reservoir expression leading to decay which is obscured by clonal expansion," *Nature Communications*, 10(1), 728.

Plantier, J.-C.C., Leoz, M., Dickerson, J.E., Oliveira, F. de, Cordonnier, F., Lemée, V., Damond, F., Robertson, D.L. & Simon, F., 2009, "A new human immunodeficiency virus derived from gorillas.," *Nat Med*, 15(8), 871–872.

Politch, J.A., Mayer, K.H., Welles, S.L., O'Brien, W.X., Xu, C., Bowman, F.P. & Anderson, D.J., 2012, "Highly active antiretroviral therapy does not completely suppress HIV in semen of sexually active HIV-infected men who have sex with men," *AIDS*, 26(12), 1535–1543.

Pollack, R.A., Jones, R.B., Pertea, M., Bruner, K.M., Martin, A.R., Thomas, A.S., Capoferri, A.A., Beg, S.A., Huang, S.H., Karandish, S., Hao, H., Halper-Stromberg, E., Yong, P.C., Kovacs, C., Benko, E., Siliciano, R.F. & Ho, Y.C., 2017, "Defective HIV-1 Proviruses Are Expressed and Can Be Recognized by Cytotoxic T Lymphocytes, which Shape the Proviral Landscape," *Cell Host and Microbe*, 21(4), 494-506.e4.

Pollard, R.B., Rockstroh, J.K., Pantaleo, G., Asmuth, D.M., Peters, B., Lazzarin, A., Garcia, F., Ellefsen, K., Podzamczer, D., Lunzen, J. van, Arastéh, K., Schürmann, D., Clotet, B., Hardy, W.D., Mitsuyasu, R., Moyle, G., Plettenberg, A., Fisher, M., Fätkenheuer, G., Fischl, M., Taiwo, B., Baksaas, I., Jolliffe, D., Persson, S., Jelmert, Ø., Hovden, A.-O., Sommerfelt, M.A., Wendel-Hansen, V. & Sørensen, B., 2014, "Safety and efficacy of the peptide-based therapeutic vaccine for HIV-1, Vacc-4×: a phase 2 randomised, double-blind, placebo-controlled trial," *The Lancet Infectious Diseases*, 14(4), 291–300.

Pudney, J. & Anderson, D., 1991, "Orchitis and human immunodeficiency virus type 1 infected cells in reproductive tissues from men with the acquired immune deficiency syndrome," *The American journal of pathology*, 139, 149–160.

Rassler, S., Ramirez, R., Khoury, N., Skowron, G. & Sahu, G.K., 2016, "Prolonged persistence of a novel replication-defective HIV-1 variant in plasma of a patient on suppressive therapy," *Virology Journal*, 13(1), 157.

Rausch, J. & Grice, S.F.J. le, 2004, "'Binding, bending and bonding': polypurine tract-primed initiation of plus-strand DNA synthesis in human immunodeficiency virus," *The International Journal of Biochemistry & Cell Biology*, 36(9), 1752–1766.

Reeves, D.B., Duke, E.R., Wagner, T.A., Palmer, S.E., Spivak, A.M. & Schiffer, J.T., 2018, "A majority of HIV persistence during antiretroviral therapy is due to infected cell proliferation," *Nature Communications*, 9(1), 4811.

Reil, H., Bukovsky, A.A., Gelderblom, H.R. & Göttlinger, H.G., 1998, "Efficient HIV-1 replication can occur in the absence of the viral matrix protein," *The EMBO Journal*, 17(9), 2699–2708.

Rein, A., Henderson, L.E. & Levin, J.G., 1998, "Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: significance for viral replication," *Trends in Biochemical Sciences*, 23(8), 297–301.

Restifo, N.P. & Gattinoni, L., 2013, "Lineage relationship of effector and memory T cells", *Current Opinion in Immunology*, 25(5), 556–563.

Robertson, D.L., Anderson, J.P., Bradac, J. a, Carr, J.K., Foley, B., Funkhouser, R.K., Gao, F., Hahn, B.H., Kalish, M.L., Kuiken, C., Learn, G.H., Leitner, T., McCutchan, F., Osmanov, S., Peeters, M., Pieniazek, D., Salminen, M., Sharp, P.M., Wolinsky, S. & Korber, B., 2000, "HIV-1 nomenclature proposal.," *Science (New York, N.Y.)*, 288(5463), 55–56.

Rosenbloom, D.I.S., Hill, A.L., Laskey, S.B. & Siliciano, R.F., 2017, "Re-evaluating evolution in the HIV reservoir," *Nature*, 551(7681), E6–E9.

Rouzioux, C. & Avettand-Fenoël, V., 2018, "Total HIV DNA: a global marker of HIV persistence," *Retrovirology*, 15(1), 30.

Rouzioux, C. & Richman, D., 2013, "How to best measure HIV reservoirs?" *Current Opinion in HIV and AIDS*, 8(3), 170–175.

Roy, S., Delling, U., Chen, C.H., Rosen, C.A. & Sonenberg, N., 1990, "A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated trans-activation.," *Genes & Development*, 4(8), 1365–1373.

Rozera, G., Abbate, I., Bruselles, A., Bartolini, B., D'Offizi, G., Nicastri, E., Tommasi, C. & Capobianchi, M.R., 2010, "Comparison of real-time PCR methods for measurement of HIV-1 proviral DNA," *Journal of Virological Methods*, 164(1–2), 135–138.

Ryscavage, P., Kelly, S., Li, J.Z., Harrigan, P.R. & Taiwo, B., 2014, "Significance and Clinical Management of Persistent Low-Level Viremia and Very-Low-Level Viremia in HIV-1-Infected Patients," *Antimicrobial Agents and Chemotherapy*, 58(7), 3585–3598.

Saksena, N.K., Lau, K.A., Dwyer, D.E. & Wang, B., (October 26th, 2011). HIV Recombination and Pathogenesis – Biological and Epidemiological Implications, HIV and AIDS - Updates on Biology, Immunology, Epidemiology and Treatment Strategies, Nancy Dumais, IntechOpen, Available from: https://www.intechopen.com/books/hiv-and-aids-updates-on-biology-immunology-epidemiology-and-treatment-strategies/hiv-recombination-and-pathogenesis-biological-and-epidemiological-implications

Sallam, M., Esbjörnsson, J., Baldvinsdóttir, G., Indriðason, H., Björnsdóttir, T.B., Widell, A., Gottfreðsson, M., Löve, A. & Medstrand, P., 2017, "Molecular epidemiology of HIV-1 in Iceland: Early introductions, transmission dynamics and recent outbreaks among injection drug users," *Infection, Genetics and Evolution*, 49, 157–163.

Sallam, M., Şahin, G.Ö., Ingman, M., Widell, A., Esbjörnsson, J. & Medstrand, P., 2017, "Genetic characterization of human immunodeficiency virus type 1 transmission in the Middle East and North Africa," *Heliyon*, 3(7), e00352.

Sato, A., Igarashi, H., Adachi, A. & Hayami, M., 1990, "Identification and localization ofvpr gene product of human immunodeficiency virus type 1," *Virus Genes*, 4(4), 303–312.

Schaeffer, E., Geleziunas, R. & Greene, W.C., 2001, "Human Immunodeficiency Virus Type 1 Nef Functions at the Level of Virus Entry by Enhancing Cytoplasmic Delivery of Virions," *Journal of Virology*, 75(6), 2993–3000.

Scheerder, M.-A. de, Vrancken, B., Dellicour, S., Schlub, T., Lee, E., Shao, W., Rutsaert, S., Verhofstede, C., Kerre, T., Malfait, T., Hemelsoet, D., Coppens, M., Dhondt, A., Looze, D. de, Vermassen, F., Lemey, P., Palmer, S. & Vandekerckhove, L., 2019, "HIV Rebound Is Predominantly Fueled by Genetically Identical Viral Expansions from Diverse Reservoirs," *Cell Host & Microbe*, 26(3), 347-358.e7.

Schnell, G., Spudich, S., Harrington, P., Price, R.W. & Swanstrom, R., 2009, "Compartmentalized Human Immunodeficiency Virus Type 1 Originates from Long-Lived Cells in Some Subjects with HIV-1–Associated Dementia," *PLoS Pathogens*, 5(4), e1000395.

Schooley, R.T., Spritzler, J., Wang, H., Lederman, M.M., Havlir, D., Kuritzkes, D.R., Pollard, R., Battaglia, C., Robertson, M., Mehrotra, D., Casimiro, D., Cox, K. & Schock, B., 2010, "AIDS Clinical Trials Group 5197: A Placebo-Controlled Trial of Immunization of HIV-1–Infected Persons with a Replication-Deficient Adenovirus Type 5 Vaccine Expressing the HIV-1 Core Protein," *The Journal of Infectious Diseases*, 202(5), 705–716.

Schröder, A.R.W., Shinn, P., Chen, H., Berry, C., Ecker, J.R. & Bushman, F., 2002, "HIV-1 Integration in the Human Genome Favors Active Genes and Local Hotspots," *Cell*, 110(4), 521–529.

Schubert, U., Bour, S., Ferrer-Montiel, A. v, Montal, M., Maldarelli, F. & Strebel, K., 1996, "Env and Vpu Proteins of Human Immunodeficiency Virus Type 1 Are Produced from Multiple Bicistronic mRNAs," *Journal of Virology*, 70(2), 809–819.

Schwartz, S., Felber, B.K., Fenyo, E.-M. & Pavlakisl, G.N., 1990, "Env and Vpu Proteins of Human Immunodeficiency Virus Type 1 Are Produced from Multiple Bicistronic mRNAs," *Journal of Virology*, 64(11), 5448–5456.

Seitz, R., 2016, "Human Immunodeficiency Virus (HIV)," *Transfusion Medicine and Hemotherapy*, 43(3), 203–222.

Semenzato, G., Agostini, C., Ometto, L., Zambello, R., Trentin, L., Chieco-Bianchi, L. & Rossi, A. de, 1995, "CD8+ T lymphocytes in the lung of acquired immunodeficiency syndrome patients harbor human immunodeficiency virus type 1," *Blood*, 85(9), 2308–2314.

Sengupta, S. & Siliciano, R.F., 2018, "Targeting the Latent Reservoir for HIV-1," *Immunity*, 48(5), 872–895.

Sgadari, C., Monini, P., Tripiciano, A., Picconi, O., Casabianca, A., Orlandi, C., Moretti, S., Francavilla, V., Arancio, A., Paniccia, G., Campagna, M., Bellino, S., Meschiari, M., Nozza, S., Sighinolfi, L., Latini, A., Muscatello, A., Saracino, A., Pietro, M. di, Galli, M., Cafaro, A., Magnani, M., Ensoli, F. & Ensoli, B., 2019, "Continued Decay of HIV Proviral DNA Upon Vaccination With HIV-1 Tat of Subjects on Long-Term ART: An 8-Year Follow-Up Study," *Frontiers in Immunology*, 10, 233.

Shan, L., Yang, H.-C., Rabi, S.A., Bravo, H.C., Shroff, N.S., Irizarry, R.A., Zhang, H., Margolick, J.B., Siliciano, J.D. & Siliciano, R.F., 2011, "Influence of Host Gene Transcription Level and Orientation on HIV-1 Latency in a Primary-Cell Model," *Journal of Virology*, 85(11), 5384–5393.

Sharaf, R.R. & Li, J.Z., 2017, "The Alphabet Soup of HIV Reservoir Markers," *Current HIV/AIDS Reports*, 14(2), 72–81.

Sharma, A.L., Singh, T.R., Devi, K.R. & Singh, L.S., 2017, "Molecular epidemiology of HIV-1 among the HIV infected people of Manipur, Northeastern India: Emergence of unique recombinant forms," *Journal of Medical Virology*, 89(6), 989–999.

Shaw, G.M. & Hunter, E., 2012, "HIV Transmission," *Cold Spring Harbor Perspectives in Medicine*, 2(11), a006965–a006965.

Shen, L. & Siliciano, R.F., 2008, "Viral reservoirs, residual viremia, and the potential of highly active antiretroviral therapy to eradicate HIV infection," *Journal of Allergy and Clinical Immunology*, 122(1), 22–28.

Sherman, E., Nobles, C., Berry, C.C., Six, E., Wu, Y., Dryga, A., Malani, N., Male, F., Reddy, S., Bailey, A., Bittinger, K., Everett, J.K., Caccavelli, L., Drake, M.J., Bates, P., Hacein-Bey-Abina, S., Cavazzana, M. & Bushman, F.D., 2017, "INSPIIRED: A Pipeline for Quantitative Analysis of Sites of New DNA Integration in Cellular Genomes," *Molecular Therapy - Methods and Clinical Development*, 4, 39–49.

Sheth, P.M., Kovacs, C., Kemal, K.S., Jones, R.B., Raboud, J.M., Pilon, R., Porte, C. la, Ostrowski, M., Loutfy, M., Burger, H., Weiser, B. & Kaul, R., 2009, "Persistent HIV RNA shedding in semen despite effective antiretroviral therapy," *AIDS*, 23(15), 2050–2054.

Shevchuk, M.M., Nuovo, G.J. & Khalife, G., 1998, "HIV in testis: quantitative histology and HIV localization in germ cells," *Journal of Reproductive Immunology*, 41(1–2), 69–79.

Sierra-Madero, J.G., Toossi, Z., Hom, D.L., Finegan, C.K., Hoenig, E. & Rich, E.A., 1994, "Relationship between Load of Virus in Alveolar Macrophages from Human Immunodeficiency Virus Type 1-Infected Persons, Production of Cytokines, and Clinical Status," *Journal of Infectious Diseases*, 169(1), 18–27.

Siliciano, J.D., Kajdas, J., Finzi, D., Quinn, T.C., Chadwick, K., Margolick, J.B., Kovacs, C., Gange, S.J. & Siliciano, R.F., 2003, "Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells," *Nature Medicine*, 9(6), 727–728.

Siliciano, J.D. & Siliciano, R.F., 2005, "Enhanced culture assay for detection and quantitation of latently infected, resting CD4+ T-cells carrying replication-competent virus in HIV-1-infected individuals," *Methods in Molecular Biology*, (304), 3–15.

Siliciano, J.D. & Siliciano, R.F., 2014, Recent developments in the search for a cure for HIV-1 infection: Targeting the latent reservoir for HIV-1, *Journal of Allergy and Clinical Immunology*, 134(1), 12–19.

Siliciano, R.F. & Greene, W.C., 2011, "HIV Latency," *Cold Spring Harbor Perspectives in Medicine*, 1(1), a007096–a007096.

Silva, M. da, Shevchuk, M.M., Cronin, W.J., Armenakas, N.A., Tannenbaum, M., Fracchia, J.A. & Ioachim, H.L., 1990, "Detection of HIV-Related Protein in Testes and Prostates of Patients with AIDS," *American Journal of Clinical Pathology*, 93(2), 196–201.

Simon, F., Mauclère, P., Roques, P., Loussert-Ajaka, I., Müller-Trutwin, M.C., Saragosti, S., Georges-Courbot, M.C., Barré-Sinoussi, F. & Brun-Vézinet, F., 1998, "Identification of a new human immunodeficiency virus type 1 distinct from group M and group O," *Nature medicine*, 4(9), 1032–7.

Simonetti, F.R., Sobolewski, M.D., Fyne, E., Shao, W., Spindler, J., Hattori, J., Anderson, E.M., Watters, S.A., Hill, S., Wu, X., Wells, D., Su, L., Luke, B.T., Halvas, E.K., Besson, G., Penrose, K.J., Yang, Z., Kwan, R.W., Waes, C. van, Uldrick, T., Citrin, D.E., Kovacs, J., Polis, M.A., Rehm, C.A., Gorelick, R., Piatak, M., Keele, B.F., Kearney, M.F., Coffin, J.M., Hughes, S.H., Mellors, J.W. & Maldarelli, F., 2016, "Clonally expanded CD4+ T cells can produce infectious HIV-1 in vivo," *Proceedings of the National Academy of Sciences*, 113(7), 1883–1888.

Simonetti, F.R., Sobolowski, M.D., Hill, S., Shao, W., Fyne, E., Wu, X., Coffin, J.M., Hughes, S.H., Mellors, J.W. & Maldarelli, F., 2015, "Residual Viremia Caused by Clonally Expanded Tumor-Infiltrating CD4+ Cells," *CROI 2015*, Abstract 105.

Simonetti, F.R., Zhang, H., Soroosh, G., Beg, S.A., Duan, J., Rhodehouse, K., Nobles, C.L., Lai, J., Hoh, R., Deeks, S.G., Bushman, F., Siliciano, J. & Siliciano, R., 2020, "Antigen-driven clonal selection shapes the fate of HIV-infected CD4+ T cells in vivo," *CROI 2020*, Abstract 73.

Sivay, M. v., Hudelson, S.E., Wang, J., Agyei, Y., Hamilton, E.L., Selin, A., Dennis, A., Kahn, K., Gomez-Olive, F.X., MacPhail, C., Hughes, J.P., Pettifor, A., Eshleman, S.H. & Grabowski, M.K., 2018, "HIV-1 diversity among young women in rural South Africa: HPTN 068," *PLOS ONE*, 13(7), e0198999.

Smith, D.M., Kingery, J.D., Wong, J.K., Ignacio, C.C., Richman, D.D. & Little, S.J., 2004, "The prostate as a reservoir for HIV-1," *AIDS*, 18(11), 1600–1602.

Soriano-Sarabia, N., Bateson, R.E., Dahl, N.P., Crooks, A.M., Kuruc, J.D., Margolis, D.M. & Archin, N.M., 2014, "Quantitation of replication-competent HIV-1 in populations of resting CD4+ T cells.," *Journal of Virology*, 88(24), 14070–14077.

Sörstedt, E., Nilsson, S., Blaxhult, A., Gisslén, M., Flamholc, L., Sönnerborg, A. & Yilmaz, A., 2016, "Viral blips during suppressive antiretroviral treatment are associated with high baseline HIV-1 RNA levels," *BMC Infectious Diseases*, 16(1), 305.

Spach, D.H., 2019, Antiretroviral Medications and Initial Therapy, *Retrieved July 14, 2020, from National HIV curriculum, https://www.hiv.uw.edu/go/antiretroviral-therapy/general-information/core-concept/all*.

Spiegelaere, W. de, Malatinkova, E., Lynch, L., Nieuwerburgh, F. van, Messiaen, P., O'Doherty, U. & Vandekerckhove, L., 2014, "Quantification of Integrated HIV DNA by Repetitive-Sampling Alu-HIV PCR on the Basis of Poisson Statistics," *Clinical Chemistry*, 60(6), 886–895.

Stockenstrom, S. von, Odevall, L., Lee, E., Sinclair, E., Bacchetti, P., Killian, M., Epling, L., Shao, W., Hoh, R., Ho, T., Faria, N.R., Lemey, P., Albert, J., Hunt, P., Loeb, L., Pilcher, C., Poole, L., Hatano, H., Somsouk, M., Douek, D., Boritz, E., Deeks, S.G., Hecht, F.M. & Palmer, S., 2015, "Longitudinal Genetic Characterization Reveals That Cell Proliferation Maintains a Persistent HIV Type 1 DNA Pool during Effective HIV Therapy," *Journal of Infectious Diseases*, 212(4), 596–607.

Strack, B., Calistri, A., Accola, M.A., Palu, G. & Göttlinger, H.G., 2000, "A role for ubiquitin ligase recruitment in retrovirus release," *Proceedings of the National Academy of Sciences*, 97(24), 13063–13068.

Strain, M.C., Gunthard, H.F., Havlir, D. v., Ignacio, C.C., Smith, D.M., Leigh-Brown, A.J., Macaranas, T.R., Lam, R.Y., Daly, O.A., Fischer, M., Opravil, M., Levine, H., Bacheler, L., Spina, C.A., Richman, D.D. & Wong, J.K., 2003, "Heterogeneous clearance rates of long-lived lymphocytes infected with HIV: Intrinsic stability predicts lifelong persistence," *Proceedings of the National Academy of Sciences*, 100(8), 4819–4824.

Strain, M.C., Lada, S.M., Luong, T., Rought, S.E., Gianella, S., Terry, V.H., Spina, C.A., Woelk, C.H. & Richman, D.D., 2013, "Highly Precise Measurement of HIV DNA by Droplet Digital PCR," *PLoS ONE*, 8(4), e55943.

Strain, M.C., Letendre, S., Pillai, S.K., Russell, T., Ignacio, C.C., Günthard, H.F., Good, B., Smith, D.M., Wolinsky, S.M., Furtado, M., Marquie-Beck, J., Durelle, J., Grant, I., Richman, D.D., Marcotte, T., McCutchan, J.A., Ellis, R.J. & Wong, J.K., 2005, "Genetic Composition of Human Immunodeficiency Virus Type 1 in Cerebrospinal Fluid and Blood without Treatment and during Failing Antiretroviral Therapy," *Journal of Virology*, 79(3), 1772–1788.

Strebel, K., Daugherty, D., Clouse, K., Cohen, D., Folks, T. & Martin, M.A., 1987, "The HIV A (sor) gene product is essential for virus infectivity," *Nature*, 328(6132), 728–730.

Stys, D., Blaha, I. & Strop, P., 1993, "Structural and functional studies in vitro on the p6 protein from the HIV-1 gag open reading frame," *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1182(2), 157–161.

Subramanian, R.P., Wildschutte, J.H., Russo, C. & Coffin, J.M., 2011, "Identification, characterization, and comparative genomic distribution of the HERV-K (HML-2) group of human endogenous retroviruses," *Retrovirology*, 8, 90.

Sundquist, W.I. & Krausslich, H.-G., 2012, "HIV-1 Assembly, Budding, and Maturation," *Cold Spring Harbor Perspectives in Medicine*, 2(7), a006924–a006924.

Sundstrom, J.B., Ellis, J.E., Hair, G.A., Kirshenbaum, A.S., Metcalfe, D.D., Yi, H., Cardona, A.C., Lindsay, M.K. & Ansari, A.A., 2007, "Human tissue mast cells are an inducible reservoir of persistent HIV infection," *Blood*, 109(12), 5293–5300.

Sunshine, S., Kirchner, R., Amr, S.S., Mansur, L., Shakhbatyan, R., Kim, M., Bosque, A., Siliciano, R.F., Planelles, V., Hofmann, O., Ho Sui, S. & Li, J.Z., 2016, "HIV Integration Site Analysis of Cellular Models of HIV Latency with a Probe-Enriched Next-Generation Sequencing Assay," *Journal of Virology*, 90(9), 4511–4519.

Suzuki, K., Juelich, T., Lim, H., Ishida, T., Watanebe, T., Cooper, D.A., Rao, S. & Kelleher, A.D., 2008, "Closed Chromatin Architecture Is Induced by an RNA Duplex Targeting the HIV-1 Promoter Region," *Journal of Biological Chemistry*, 283(34), 23353–23363.

Symons, J., Chopra, A., Malantinkova, E., Spiegelaere, W. de, Leary, S., Cooper, D., Abana, C.O., Rhodes, A., Rezaei, S.D., Vandekerckhove, L., Mallal, S., Lewin, S.R. & Cameron, P.U., 2017, "HIV integration sites in latently infected cell lines: evidence of ongoing replication," *Retrovirology*, 14(1), 2.

Tang, M. W. and Shafer, R. W., 2012, "HIV-1 antiretroviral resistance: Scientific principles and clinical applications", *Drugs*, 72(9).

Taube, R., Fujinaga, K., Wimmer, J., Barboric, M. & Peterlin, B.M., 1999, "Tat Transactivation: A Model for the Regulation of Eukaryotic Transcriptional Elongation," *Virology*, 264(2), 245–253.

Templer, S.P., Seiverth, B., Baum, P., Stevens, W., Seguin-Devaux, C. & Carmona, S., 2016, "Improved Sensitivity of a Dual-Target HIV-1 Qualitative Test for Plasma and Dried Blood Spots," *Journal of Clinical Microbiology*, 54(7), 1877–1882.

Thomas, J., Perron, H. & Feschotte, C., 2018, "Variation in proviral content among human genomes mediated by LTR recombination," *Mobile DNA*, 9(1), 36.

Thomas, J., Ruggiero, A., Paxton, W.A. & Pollakis, G., 2020, "Measuring the Success of HIV-1 Cure Strategies," *Frontiers in Cellular and Infection Microbiology*, 10, 134.

Tierney, C., Lathey, J.L., Christopherson, C., Bettendorf, D.M., D'Aquila, R.T., Hammer, S.M. & Katzenstein, D.A., 2003, "Prognostic Value of Baseline Human Immunodeficiency Virus Type 1 DNA Measurement for Disease Progression in Patients Receiving Nucleoside Therapy," *The Journal of Infectious Diseases*, 187(1), 144–148.

Tirado, G., Jove, G., Kumar, R., Noel, R.J., Reyes, E., Sepulveda, G., Yamamura, Y. & Kumar, A., 2004, "Differential virus evolution in blood and genital tract of HIV-infected females: evidence for the involvement of drug and non-drug resistance-associated mutations," *Virology*, 324(2), 577–586.

Tobin, N.H., Learn, G.H., Holte, S.E., Wang, Y., Melvin, A.J., McKernan, J.L., Pawluk, D.M., Mohan, K.M., Lewis, P.F., Mullins, J.I. & Frenkel, L.M., 2005, "Evidence that Low-Level Viremias during Effective Highly Active Antiretroviral Therapy Result from Two Processes: Expression of Archival Virus and Replication of Virus," *Journal of Virology*, 79(15), 9625–9634.

Torresilla, C., Mesnard, J.-M. & Barbeau, B., 2015, "Reviving an Old HIV-1 Gene: The HIV-1 Antisense Protein," *Current HIV Research*, 13(2), 117–124.

Tumiotto, C., Bellecave, P., Recordon-Pinson, P., Groppi, A., Nikolski, M. & Fleury, H., 2018, "Diversity of HIV-1 in Aquitaine, Southwestern France, 2012–2016," *AIDS Research and Human Retroviruses*, 34(5), 471–473.

UNAIDS, 2020, UNAIDS Data, *Retrieved July 16, 2020, from UNAIDS, https://aids2020.unaids.org/report/*.

Usami, Y., Popov, S., Popova, E., Inoue, M., Weissenhorn, W. & G. Göttlinger, H., 2009, "The ESCRT pathway and HIV-1 budding," *Biochemical Society Transactions*, 37(1), 181–184.

Vandegraaff, N., Kumar, R., Burrell, C.J. & Li, P., 2001, "Kinetics of Human Immunodeficiency Virus Type 1 (HIV) DNA Integration in Acutely Infected Cells as Determined Using a Novel Assay for Detection of Integrated HIV DNA," *Journal of Virology*, 75(22), 11253–11260.

Vanhamel, J., Bruggemans, A. & Debyser, Z., 2019, "Establishment of latent HIV-1 reservoirs: what do we really know?" *Journal of Virus Eradication*, 5, 3–9.

van Duyne, R., Kuo, L.S., Pham, P., Fujii, K. & Freed, E.O., 2019, "Mutations in the HIV-1 envelope glycoprotein can broadly rescue blocks at multiple steps in the virus replication cycle," *Proceedings of the National Academy of Sciences*, 116(18), 9040–9049.

van Zyl, G.U. van, Katusiime, M.G., Wiegand, A., McManus, W.R., Bale, M.J., Halvas, E.K., Luke, B., Boltz, V.F., Spindler, J., Laughton, B., Engelbrecht, S., Coffin, J.M., Cotton, M.F., Shao, W., Mellors, J.W. & Kearney, M.F., 2017, "No evidence of HIV replication in children on antiretroviral therapy," *Journal of Clinical Investigation*, 127(10), 3827–3834.

van't Wout, A.B., Ran, L.J., Kuiken, C.L., Kootstra, N.A., Pals, S.T. & Schuitemaker, H., 1998, "Analysis of the Temporal Relationship between Human Immunodeficiency Virus Type 1 Quasispecies in Sequential Blood Samples and Various Organs Obtained at Autopsy," *Journal of Virology*, 72(1), 488–496.

Veldsman, K.A., Rensburg, A., Isaacs, S., Naidoo, S., Laughton, B., Lombard, C., Cotton, M.F., Mellors, J.W. & Zyl, G.U., 2019, "HIV-1 DNA decay is faster in children who initiate ART shortly after birth than later," *Journal of the International AIDS Society*, 22(8).

Viard, J.-P., Burgard, M., Hubert, J.-B., Aaron, L., Rabian, C., Pertuiset, N., Lourenço, M., Rothschild, C. & Rouzioux, C., 2004, "Impact of 5 years of maximally successful highly active antiretroviral therapy on CD4 cell count and HIV-1 DNA level," *AIDS*, 18(1), 45–49.

Violari, A., Cotton, M.F., Gibb, D.M., Babiker, A.G., Steyn, J., Madhi, S.A., Jean-Philippe, P. & McIntyre, J.A., 2008, "Early Antiretroviral Therapy and Mortality among HIV-Infected Infants," *The New England Journal of Medicine*, 359(21), 2233–2244.

Volz, E.M., Vu, S. le, Ratmann, O., Tostevin, A., Dunn, D., Orkin, C., O'Shea, S., Delpech, V., Brown, A., Gill, N. & Fraser, C., 2018, "Molecular Epidemiology of HIV-1 Subtype B Reveals Heterogeneous Transmission Risk: Implications for Intervention and Control," *The Journal of Infectious Diseases*, 217(10), 1522–1529.

Wagner, T.A., McKernan, J.L., Tobin, N.H., Tapia, K.A., Mullins, J.I. & Frenkel, L.M., 2013, "An increasing proportion of monotypic HIV-1 DNA sequences during antiretroviral treatment suggests proliferation of HIV-infected cells.," *Journal of virology*, 87(3), 1770–1778.

Wagner, T.A., McLaughlin, S., Garg, K., Cheung, C.Y.K., Larsen, B.B., Styrchak, S., Huang, H.C., Edlefsen, P.T., Mullins, J.I. & Frenkel, L.M., 2014, "Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection," *Science*, 345(6196), 570–573.

Wahl, S.M., Redford, M., Christensen, S., Mack, W., Cohn, J., Janoff, E.N., Mestecky, J., Jenson, H.B., Navazesh, M., Cohen, M., Reichelderfer, P. & Kovacs, A., 2011, "Systemic and Mucosal Differences in HIV Burden, Immune, and Therapeutic Responses," *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 56(5), 401–411.

Wang, G.P., Ciuffi, A., Leipzig, J., Berry, C.C. & Bushman, F.D., 2007, "HIV integration site selection: Analysis by massively parallel pyrosequencing reveals association with epigenetic modifications," *Genome Research*, 17(8), 1186–1194.

Wang, W., Ye, C., Liu, J., Zhang, D., Kimata, J.T. & Zhou, P., 2014, "CCR5 Gene Disruption via Lentiviral Vectors Expressing Cas9 and Single Guided RNA Renders Cells Resistant to HIV-1 Infection," *PLoS ONE*, 9(12), e115987.

Wang, Z., Gurule, E.E., Brennan, T.P., Gerold, J.M., Kwon, K.J., Hosmane, N.N., Kumar, M.R., Beg, S.A., Capoferri, A.A., Ray, S.C., Ho, Y.C., Hill, A.L., Siliciano, J.D. & Siliciano, R.F., 2018, "Expanded cellular clones carrying replication-competent HIV-1 persist, wax, and wane," *Proceedings of the National Academy of Sciences of the United States of America*, 115(11), E2575–E2584.

Wei, D.G., Chiang, V., Fyne, E., Balakrishnan, M., Barnes, T., Graupe, M., Hesselgesser, J., Irrinki, A., Murry, J.P., Stepan, G., Stray, K.M., Tsai, A., Yu, H., Spindler, J., Kearney, M., Spina, C.A., McMahon, D., Lalezari, J., Sloan, D., Mellors, J., Geleziunas, R. & Cihlar, T., 2014, "Histone Deacetylase Inhibitor Romidepsin Induces HIV Expression in CD4 T Cells from Patients on Suppressive Antiretroviral Therapy at Concentrations Achieved by Clinical Dosing," *PLoS Pathogens*, 10(4), e1004071.

Weinberger, L.S., Burnett, J.C., Toettcher, J.E., Arkin, A.P. & Schaffer, D. v., 2005, "Stochastic Gene Expression in a Lentiviral Positive-Feedback Loop: HIV-1 Tat Fluctuations Drive Phenotypic Diversity," *Cell*, 122(2), 169–182.

WHO, 2020, "Mother-to-child transmission of HIV," *Retrieved 16 July, 2020, from World Health Organization, https://www.who.int/hiv/topics/mtct/about/en/*.

Wilk, T., Gross, I., Gowen, B.E., Rutten, T., Haas, F. de, Welker, R., Kräusslich, H.-G., Boulanger, P. & Fuller, S.D., 2001, "Organization of Immature Human Immunodeficiency Virus Type 1," *Journal of Virology*, 75(2), 759–771.

Williams, J.P., Hurst, J., Stöhr, W., Robinson, N., Brown, H., Fisher, M., Kinloch, S., Cooper, D., Schechter, M., Tambussi, G., Fidler, S., Carrington, M., Babiker, A., Weber, J., Koelsch, K.K., Kelleher, A.D., Phillips, R.E. & Frater, J., 2014, "HIV-1 DNA predicts disease progression and post-treatment virological control," *eLife*, 3, e03821.

Willig, J.H., Nevin, C.R., Raper, J.L., Saag, M.S., Mugavero, M.J., Willig, A.L., Burkhardt, J.H., Schumacher, J.E. & Johnson, V.A., 2010, "Cost Ramifications of Increased Reporting of Detectable Plasma HIV-1 RNA Levels by the Roche COBAS AmpliPrep/COBAS TaqMan HIV-1 Version 1.0 Viral Load Test," *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 54(4), 442–443.

Winston, J.A., Bruggeman, L.A., Ross, M.D., Jacobson, J., Ross, L., D'Agati, V.D., Klotman, P.E. & Klotman, M.E., 2001, "Nephropathy and Establishment of a Renal Reservoir of HIV Type 1 during Primary Infection," *New England Journal of Medicine*, 344(26), 1979–1984.

Wolff, H., Mayer, K., Seage, G., Politch, J., Horsburgh, C.R. & Anderson, D., 1992, "A comparison of HIV-1 antibody classes, titers, and specificities in paired semen and blood samples from HIV-1 seropositive men," *Journal of Acquired Immune Deficiency Syndromes*, 5, 65–69.

Wong, J.K., Ignacio, C.C., Torriani, F., Havlir, D., Fitch, N.J.S. & Richman, D.D., 1997, "In Vivo Compartmentalization of Human Immunodeficiency Virus: Evidence from the Examination of pol Sequences from Autopsy Tissues," *Journal of Virology*, 71(3), 2059–2071.

Wong, J.K. & Yukl, S.A., 2016, "Tissue reservoirs of HIV," *Current Opinion in HIV and AIDS*, 11(4), 362–370.

World Health Organisation, 2016, "Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection", *Available at https://www.who.int/hiv/pub/arv/arv-2016/en/*.

Worobey, M., Gemmel, M., Teuwen, D.E., Haselkorn, T., Kunstman, K., Bunce, M., Muyembe, J.-J., Kabongo, J.-M.M., Kalengayi, R.M., Marck, E. van, Gilbert, M.T.P. & Wolinsky, S.M., 2008, "Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960.," *Nature*, 455(7213), 661–664.

Xiao, Q., Guo, D. & Chen, S., 2019, "Application of CRISPR/Cas9-Based Gene Editing in HIV-1/AIDS Therapy," *Frontiers in Cellular and Infection Microbiology*, 9, 69.

Xu, L., Yang, H., Gao, Y., Chen, Z., Xie, L., Liu, Yulin, Liu, Ying, Wang, X., Li, H., Lai, W., He, Y., Yao, A., Ma, L., Shao, Y., Zhang, B., Wang, C., Chen, H. & Deng, H., 2017, "CRISPR/Cas9-Mediated CCR5 Ablation in Human Hematopoietic Stem/Progenitor Cells Confers HIV-1 Resistance In Vivo," *Molecular Therapy*, 25(8), 1782–1789.

Yang, C., McNulty, A., Diallo, K., Zhang, J., Titanji, B., Kassim, S., Wadonda-Kabondo, N., Aberle-Grasse, J., Kibuka, T., Ndumbe, P. M., Vedapuri, S., Zhou, Z., Chilima, B., & Nkengasong, J. N., 2010, "Development and application of a broadly sensitive dried-blood-spot-based genotyping assay for global surveillance of HIV-1 drug resistance," *Journal of Clinical Microbiology*, 48(9).

Yerly, S., Günthard, H.F., Fagard, C., Joos, B., Perneger, T. v, Hirschel, B. & Perrin, L., 2004, "Proviral HIV-DNA predicts viral rebound and viral setpoint after structured treatment interruptions," *AIDS*, 18(14), 1951–1953.

Young, J., Rickenbach, M., Calmy, A., Bernasconi, E., Staehelin, C., Schmid, P., Cavassini, M., Battegay, M., Günthard, H.F. & Bucher, H.C., 2015, "Transient detectable viremia and the risk of viral rebound in patients from the Swiss HIV Cohort Study," *BMC Infectious Diseases*, 15(1), 382.

Yu, J.J., Wu, T.L., Liszewski, M.K., Dai, J., Swiggard, W.J., Baytop, C., Frank, I., Levine, B.L., Yang, W., Theodosopoulos, T. & O'Doherty, U., 2008, "A more precise HIV integration assay designed to detect small differences finds lower levels of integrated DNA in HAART treated patients," *Virology*, 379(1), 78–86.

Yu, X., Yuan, X., Matsuda, Z., Lee, T.-H. & Essex, M., 1992, "The Matrix Protein of Human Immunodeficiency Virus Type 1 Is Required for Incorporation of Viral Envelope Protein into Mature Virions," *Journal of Virology*, 66(8), 4966–4971.

Yuan, X., Yu, X., Lee, T.-H. & Essex, M., 1993, "Mutations in the N-Terminal Region of Human Immunodeficiency Virus Type 1 Matrix Protein Block Intracellular Transport of the Gag Precursor," *Journal of Virology*, 67(11), 6387–6394.

Yukl, S.A., Gianella, S., Sinclair, E., Epling, L., Li, Q., Duan, L., Choi, A.L.M., Girling, V., Ho, T., Li, P., Fujimoto, K., Lampiris, H., Hare, C.B., Pandori, M., Haase, A.T., Günthard, H.F., Fischer, M., Shergill, A.K., McQuaid, K., Havlir, D.V. & Wong, J.K., 2010, "Differences in HIV Burden and Immune Activation within the Gut of HIV-Positive Patients Receiving Suppressive Antiretroviral Therapy," *The Journal of Infectious Diseases*, 202(10), 1553–1561.

Yukl, S.A., Kaiser, P., Kim, P., Telwatte, S., Joshi, S.K., Vu, M., Lampiris, H. & Wong, J.K., 2018, "HIV latency in isolated patient CD4+T cells may be due to blocks in HIV transcriptional elongation, completion, and splicing," *Science Translational Medicine*, 10(430), e9927.

Yukl, S.A., Shergill, A.K., Ho, T., Killian, M., Girling, V., Epling, L., Li, P., Wong, L.K., Crouch, P., Deeks, S.G., Havlir, D. v., McQuaid, K., Sinclair, E. & Wong, J.K., 2013, "The Distribution of HIV DNA and RNA in Cell Subsets Differs in Gut and Blood of HIV-Positive Patients on ART: Implications for Viral Persistence," *The Journal of Infectious Diseases*, 208(8), 1212–1220.

Yukl, S.A., Sinclair, E., Somsouk, M., Hunt, P.W., Epling, L., Killian, M., Girling, V., Li, P., Havlir, D. v., Deeks, S.G., Wong, J.K. & Hatano, H., 2014, "A comparison of methods for measuring rectal HIV levels suggests that HIV DNA resides in cells other than CD4+ T cells, including myeloid cells," *AIDS*, 28(3), 439–442.

Zack, J.A., Arrigo, S.J., Weitsman, S.R., Go, A.S., Haislip, A. & Chen, I.S.Y., 1990a, "HIV-1 entry into quiescent primary lymphocytes: Molecular analysis reveals a labile, latent viral structure," *Cell*, 61(2), 213–222.

Zack, J.A., Arrigo, S.J., Weitsman, S.R., Go, A.S., Haislip, A. & Chen, I.S.Y., 1990b, "HIV-1 entry into quiescent primary lymphocytes: Molecular analysis reveals a labile, latent viral structure," *Cell*, 61(2), 213–222.

Zapp, M.L. & Green, M.R., 1989, "Sequence-specific RNA binding by the HIV-1 Rev protein," *Nature*, 342(6250), 714–716.

Zhou, W., Parent, L.J., Wills, J.W. & Reshl2, M.D., 1994, "Identification of a Membrane-Binding Domain within the Amino-Terminal Region of Human Immunodeficiency Virus Type 1 Gag Protein Which Interacts with Acidic Phospholipids," *Journal of Virology*, 68(4), 2556–2569.

Zimmerman, C., Klein, K.C., Kiser, P.K., Singh, A.R., Firestein, B.L., Riba, S.C. & Lingappa, J.R., 2002, "Identification of a host protein essential for assembly of immature HIV-1 capsids," *Nature*, 415(6867), 88–92.

Addendum A – Research output materials

HIV Pseudo viremia from large lysed defective CD4 cell clone, Stellenbosch University Annual Academic Year Day 2019

Botha, J.C., Steegen, K., Hans L., Karstaedt A., Carmona S., Reddy D., van Zyl G.U.

HIV viral load and drug resistance testing for patient management is based on the accurate and reliable analysis of circulating replication competent virus. The presence of cellular DNA in plasma can cause an overestimation of the HIV viral load and influence drug resistance results. A vast majority of the HIV reservoir is defective, partially due to APOBEC mediated hypermutations. These hypermutated proviruses do not produce progeny virus but can be detected in plasma as a result of cell lysis due to improper sample processing. Drug resistance associated mutations (M184I and E183K) have been identified in proviral DNA of treatment naïve patients as a result of APOBEC editing. Investigating unusual cases of unsuppressed viremia is essential to understand the mechanism and cause of the detectable viral load. Diagnostic error due to pre-analytical sample processing could impact patient management through overestimation of viral loads and detection of drug resistance mutations arising from defective proviruses. It is therefore imperative to avoid cell lysis when processing plasma samples.

In this study a patient on cART presented with an unsuppressed viral load for over 3 months. Drug resistance nucleotide sequencing revealed identical APOBEC mutated sequences in three plasma samples taken over this time period. A whole blood EDTA sample (sample 4) was collected for further investigation. Peripheral blood mononuclear cells and plasma were isolated from the EDTA sample for HIV cell associated DNA analysis and single genome sequencing (SGS). SGS targeting the p6, protease and reverse-transcription (p6ProRT) region was performed on cell debris and virus isolated from plasma samples. Identical p6ProRT hypermutated sequences were detected in centrifuged cell debris from multiple plasma samples suggesting that lysed cells rather than free virus are the predominant species detected by viral load testing and sequencing.

HIV pseudo viremia from large defective cell population

JC Botha[1], K Steegen[2,3], L Hans[3], A Karstaedt[2], S Carmona[2], D Reddy[2,4], GU van Zyl[1,3]

1 Stellenbosch University, Cape Town, South Africa; 2 University of the Witwatersrand, Johannesburg, South Africa; 3 National Health Laboratory Service, Johannesburg, South Africa; 4 Chris Hani Baragwanath Hospital, Johannesburg, South Africa

Background and preliminary data

- HIV-1 viral load (VL) monitoring is premised on an elevated VL indicating HIV-1 replication due to drug resistance or poor adherence
- An HIV-1 positive adult female patient on long term cART presented with unsuppressed viral loads
- HIV-1 drug resistance testing by Sanger population sequencing performed on 3 separate samples over a 3 month period, all yielded identical hypermutated sequences (Figure 1)
- Residual plasma from samples 1-3 taken on dates (red arrows, Figure 2):
  - Sample 1: 12-02-2018 (1.5 ml), VL: 4230 copies/ml
  - Sample 2: 06-04-2018 (0.5 ml), VL: 3880 copies/ml
  - Sample 3: 02-05-2018 (excluded due to limited sample)
- EDTA sample 4 taken on 15-03-2019 (black arrow, Figure 2), VL: 2388 copies/ml
- ART regimen (indicated by bar in Figure 2):
  - 2013 – 10/2014 (blue): FTC/TDF/EFV
  - 10/2014 – 14/03/2019 (red): AZT/FTC/TDF/LPV/R
- The three plasma samples and an EDTA sample (taken more than a year after the first plasma sample) were sent to our lab for further investigation

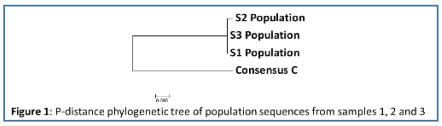


Figure 1: P-distance phylogenetic tree of population sequences from samples 1, 2 and 3

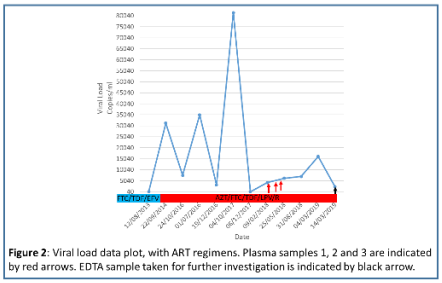
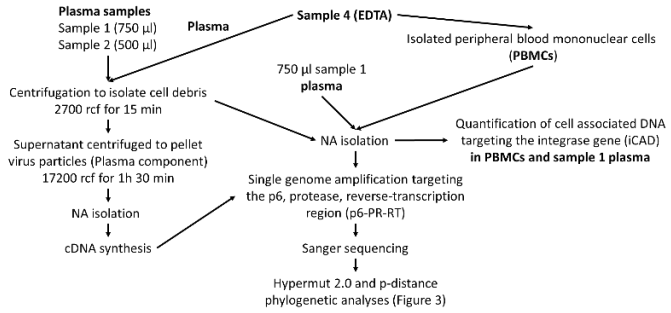


Figure 2: Viral load data plot, with ART regimens. Plasma samples 1, 2 and 3 are indicated by red arrows. EDTA sample taken for further investigation is indicated by black arrow.

Aim

- To investigate the source of apparent elevated VL by determining the source of HIV genomic material: distinguishing cell debris and virus particles

Methods



Results - Continued

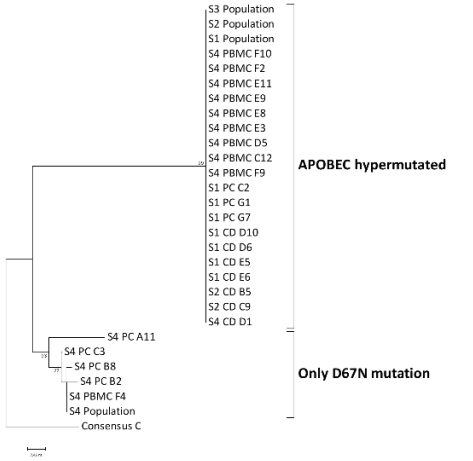


Figure 3: P-distance phylogenetic tree of all nucleotide sequences generated for the p6-PR-RT region

Results

- iCAD results indicated the presence of 66,3 HIV DNA copies per million PBMCs tested
- iCAD results from Sample 1 plasma was negative for HIV but CCR5 results indicate the presence of 46525 cells/ml plasma
- p6ProRT SGS yielded limited positive results for plasma compartment and cell debris samples
- APOBEC population detected in all PCR positive components, except 4 plasma component
- Plasma population sequences from drug resistance testing included in phylogenetic analysis for comparison (Figure 3)

Table 1: Summary of PCR and sequencing results indicating detected mutations

Sample	Component	Sequence results	Mutations detected
Sample 1	Plasma	Population sequence	APOBEC
	Cell debris (CD)	p6-PR-RT SGS	APOBEC
	Plasma component (PC)	p6-PR-RT SGS	APOBEC
Sample 2	Plasma	Population sequence	APOBEC
	Cell debris	p6-PR-RT SGS	APOBEC
	Plasma component	RT-PCR negative	RT-PCR negative
Sample 4 plasma	Plasma	Population sequence	D67N
	Cell debris	p6-PR-RT SGS	APOBEC
	Plasma component	p6-PR-RT SGS	D67N
Sample 4	PBMC	p6-PR-RT SGS	1/10 – D67N 9/10 – APOBEC

Discussion and Conclusions

- Defective HIV-1 sequences rather than a replication competent viral population was the main source of "plasma" viraemia
  - 9/10 sequences from PBMCs are identical to APOBEC sequences from cell debris and plasma component from "plasma samples"
  - This suggests a major APOBEC mutated population consisting of identical p6-PR-RT sequences (a possible large CD4 clone with a defective provirus)
  - The presence of these sequences in pre-spin cell debris indicate that these viral sequences in plasma most likely originate from cell lysis
  - The iCAD results from PBMCs is likely an underestimation of the total number of infected cells: APOBEC hypermutated population may be poorly detectable by this assay as supported by the inability of the iCAD assay to detect HIV-1 DNA in sample 1 plasma, which yielded defective p6-pro-RT sequences
- APOBEC mutated sequences represent the main source of detectable HIV in 'plasma' in 2/3 time points and likely originates from spill-over from cell associated DNA as it was found only once out of 3 times in plasma after excluding cell debris

Acknowledgements

- Study participant, clinical and nursing staff
- Funding received from: South Africa – U.S. Programme for Collaborative Biomedical Research, NIH grant U01CA200441, the South African Polio Research Foundation (PRF) and GvZ from U01CA200441

Nonsuppressible viremia on ART from large cell clones carrying intact proviruses, *CROI 2019*

Halvas, E.K., Joseph, K., Brandt, L.D., **Botha, J.C**., Sobolewski, M., Jacobs, J.L., Keele, B.F., Kearney, M.F., Coffin, J.M., Rausch, J.W., Guo, S., Wu, X., Hughes, S.H. & Mellors, J.W.

Clinically detectable viremia on ART is generally attributed to virus replication from incomplete adherence and/or drug resistance. One case of infectious viremia from a large cell clone with an intergenic intact provirus has been reported in an individual with metastatic cancer (Simonetti, PNAS 2016). We studied individuals referred for clinically detectable viremia despite receiving potent ART, adherence counseling, and in some cases, regimen switches or intensification.

Peripheral blood mononuclear cells (PBMCs) and plasma were collected at two or more time points from donors with plasma HIV RNA >20 copies/ml occurring for >6 months on combination ART. Single-genome sequencing was performed on plasma HIV RNA, cell-associated HIV DNA (CAD), and p24+ culture supernatants from quantitative viral outgrowth assays (qVOA). The clonal cellular origin of viremia was assessed by phylogenetics and integration site analysis (ISA), and confirmed by sequencing the integrated provirus and the flanking host sequences.

Across the 10 individuals referred, median plasma HIV-1 RNA was 97.5 cps/mL (range 40 to 356 cps/mL) after a median of 10 years on ART. One donor (A-04) had phylogenetic evidence of virus evolution and accumulation of resistance mutations and was not analyzed further. Each of the other 9 donors had multiple identical single-genome HIV RNA sequences in plasma that did not change over time and lacked resistance to the current ART regimen. In 6 of 9 donors, HIV sequences from plasma matched proviral sequences in PBMC. Plasma HIV RNA and proviral sequences were identical to HIV RNA in p24+ qVOA wells for 4 donors (C02, C03, R09, T13). The integration sites for the intact proviruses producing viremia were mapped to the human genome for 3 donors (4th pending). Integrations were in introns of the MATR3, ZNF268, and ABCA11P genes for C02, C03, and R09, respectively. The provirus in MATR3 and ZNF268 were in the opposite orientation to the gene, whereas the ABCA11P integrant was

in the same orientation. The intact provirus comprised 4.2-15.4% of all proviruses in PBMC with amplifiable pro/pol sequences.

Large cell clones carrying intact proviruses can produce clinically relevant levels of viremia and should be considered in managing patients. The mechanisms involved in clonal expansion and persistence of cells with intact proviruses that produce viremia need to be understood to effectively target the HIV reservoir.

HIV Pseudo viremia from large defective cell population, *28th International Workshop on HIV Drug Resistance and Treatment Strategies 2019*

Botha, J.C., Steegan, K., Hans, L., Karstaedt, A., Carmona, S., Reddy, D. & van Zyl, G.U.

BACKGROUND: HIV-1 plasma viral load (HIVVL) >1000 copies/mL is the World Health Organization threshold for virological failure (VF). When HIVVL assays detect free plasma virus, it is thought to represent viral replication. However, with delayed sample processing HIV-1 cell associated DNA could "leak" into the plasma compartment. Most HIV-1 DNA has defects, which include large internal deletions, gag-leader mutations and APOBEC mediated hypermutations; the latter associated with the spurious detection of drug resistance associated mutations (E138K and M184I). It is not known, however, to what extent HIV-1 DNA leakage could result in diagnosing "persisting VF". We therefore investigated an unusual case, who had identical APOBEC mediated hypermutations, by bulk sequencing, over a period of 3 months, in three subsequent plasma samples (1-3), while viral loads ranged from 4230 to 6170 copies/mL, using single genome sequencing (SGS) and in addition processed a whole blood EDTA sample (sample 4), collected 10 months later.

METHODS: Plasma from each time point was centrifuged at low speed to separate cell debris from plasma, from which virus was pelleted through high-speed microcentrifugation. SGS targeting the p6, protease and reverse-transcription (p6-PR-RT) region was performed on cell debris and virus pellets from plasma samples with adequate volume (samples 1, 2 and 4). For sample 4 peripheral blood mononuclear cells (PBMCs) were processed to quantify HIV-1 cell associated DNA using a quantitative PCR targeting *integrase* and to perform SGS.

RESULTS: Identical p6-PR-RT hypermutated nucleotide sequences were detected in cell debris of samples 1, 2 and 4, plasma pelleted virus (samples 1) and PBMC (sample 4). Sample 4 had a total HIV-1 DNA load of 66.3 HIV DNA copies per 1 million PBMCs and of 10 p6-PR-RT sequences obtained, 9 were identical to APOBEC mutated sequences detected in sample 1-, 2- and 4 cell debris and sample 1 "virus pellet". Sample 4 plasma virus was non-hypermutated.

CONCLUSION: Our findings showed that over a 3-month period a defective HIV sequence was the predominant sequence in blood "plasma" during VF. All "plasma" samples contained cell debris, which in 2 out of 3 cases had hypermutated sequences, representing the only

detectable sequence at two time points. In sample 1 the hypermutated sequence remained in the ‘virus pellet’ after a pre-spin. A large defective hypermutated HIV-1 DNA population, which may represent a CD4 clone prone to cell lysis, was the main source of detected HIV-1 in “plasma” over the 3 months and remained detectable as the major proviral sequence in PBMC 10 months later. Multiple hypermutated sequences drew our attention to this case, not representing true viremia. It is not known how many cases with “leaked” HIV-1 DNA without such obvious defects may be falsely classified as having VF.

HIV pseudo viremia from large defective cell population

JC Botha[1], K Steegen[2,3], L Hans[3], A Karstaedt[2], S Carmona[2], D Reddy[2,4], GU van Zyl[1,3]

1 Stellenbosch University, Cape Town, South Africa; 2 University of the Witwatersrand, Johannesburg, South Africa; 3 National Health Laboratory Service, Johannesburg, South Africa; 4 Chris Hani Baragwanath Hospital, Johannesburg, South Africa

Background and preliminary data

- HIV-1 viral load (VL) monitoring is premised on an elevated VL indicating HIV-1 replication due to drug resistance or poor adherence
- A 46 year old HIV-1 positive adult female patient on long term cART presented with virological failure
- Drug resistance testing by Sanger population sequencing performed on 3 separate samples over a 3-month period, all yielded identical hypermutated sequences (Figure 1)
- Residual plasma from samples 1-3 taken on dates (red dots, Figure 2):
  - Sample 1: 12-02-2018 (1.5 ml), VL: 4230 copies/ml
  - Sample 2: 06-04-2018 (0.5 ml), VL: 3880 copies/ml
  - Sample 3: 02-05-2018 (0.5 ml), VL: 439 copies/ml
- EDTA sample 4 taken on 14-03-2019 (black dot, Figure 2), VL: 2388 copies/ml
- ART regimen (indicated by bar in Figure 2):
  - 2013 – 10/2014 (blue): FTC/TDF/EFV
  - 10/2014 – 14/03/2019 (red): AZT/FTC/TDF/LPV/R
- The three plasma samples and an EDTA sample (taken 13 months after the first plasma sample) were sent to our lab for further investigation

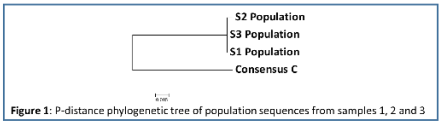


Figure 1: P-distance phylogenetic tree of population sequences from samples 1, 2 and 3

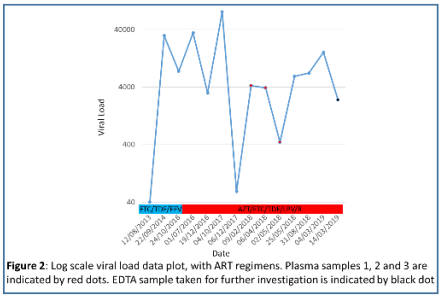
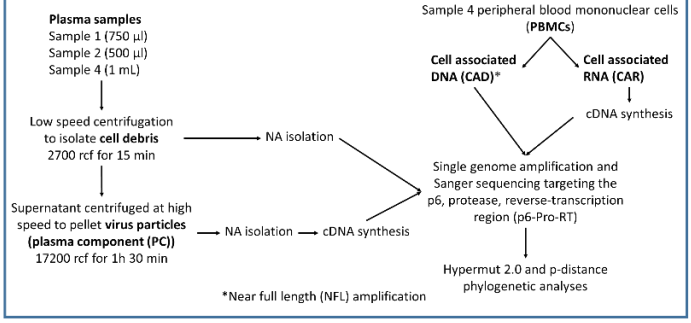


Figure 2: Log scale viral load data plot, with ART regimens. Plasma samples 1, 2 and 3 are indicated by red dots. EDTA sample taken for further investigation is indicated by black dot

Aim

To investigate the apparent virological failure by determining the source of HIV genomic material: distinguishing cell debris and virus particles

Methods



Results - Continued

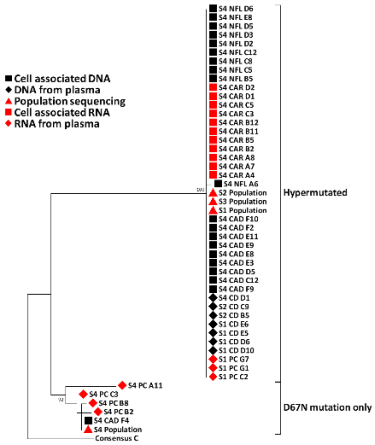


Figure 3: P-distance phylogenetic tree of all nucleotide sequences generated for the p6-Pro-RT region

Results

- p6-Pro-RT CAD-SGS endpoint dilution (1:27)
- p6-Pro-RT CAD-SGS endpoint dilution (1:2187)
- Identical hypermutated sequences detected in all PCR positive components, except sample 2 & 4 plasma component (Table 1)
- Plasma population sequences from drug resistance testing included in phylogenetic analysis for comparison (Figure 3)
- Partial gag sequences of NFL products indicate a stop codon in p24, resulting in truncated p24 capsid protein

Table 1: Summary of PCR and sequencing results indicating detected mutations

Sample	Component	Sequence results	Mutations detected
Sample 1	Plasma	Population drug resistance seq	APOBEC
	Cell debris	p6-Pro-RT SGS	APOBEC
	Virus pellet	cDNA p6-Pro-RT SGS	APOBEC
Sample 2	Plasma	Population drug resistance seq	APOBEC
	Cell debris	p6-Pro-RT SGS	APOBEC
	Virus pellet	cDNA p6-Pro-RT negative	PCR negative
Sample 4 plasma	Plasma	Population drug resistance seq	D67N
	Cell debris	p6-Pro-RT SGS	APOBEC
	Virus pellet	cDNA p6-Pro-RT SGS	D67N
Sample 4	PBMC	CAD p6-Pro-RT SGS	1/10 – D67N 9/10 – APOBEC
		CAD NFL (p6-Pro-RT region)	APOBEC
		CAR p6-Pro-RT SGS	APOBEC

Discussion and Conclusions

- 19/20 CAD and 11/11 CAR sequences from PBMCs are identical to APOBEC hypermutated sequences from cell debris and plasma component
- This suggests a major APOBEC mutated population consisting of identical p6-Pro-RT sequences (possible CD4+ T-cell clone)
- The APOBEC strain is highly transcribed, as detected by CAR-SGS (endpoint dilution 1:2187)
- Production of viral proteins and packaging of viral genomes are highly unlikely given the hypermutated genome with at least one stop codon in gag-p24
- The cause of "virological failure" appears to be a large population of cells with identical hypermutated proviruses undergoing cytolysis, releasing cellular nucleic acid including HIV DNA and mRNA into plasma
- Release of cellular nucleic acids into plasma may be an underappreciated cause of false virological failure and likely influence patient care

Acknowledgements

- Study participant, clinical and nursing staff
- Funding received from: South Africa – U.S. Programme for Collaborative Biomedical Research, NIH grant U01CA200441, the South African Polio Research Foundation (PRF) and GvZ from U01CA200441

'False ART failure' from identical hypermutated HIV nucleic acid in plasma, *CROI 2020*

Botha, J.C., Steegan, K., Hans, L., Karstaedt, A., Carmona, S., Reddy, D., Kearney, M., Mellors, J. & van Zyl, G.U.

BACKGROUND: Plasma HIV-1 RNA above the limit of detection of commercial assays on ART arise from i)complete cycles of viral replication as a consequence of inadequate drug exposure and/or drug resistance; or from ii)virions produced from proviruses in clonally-expanded cells without complete cycles of replication. Most proviruses that persist on ART are defective, including those hypermutated by APOBEC. Although hypermutated proviruses can be transcribed into mRNA and even spliced, their packaging into virions is expected to be very inefficient. Here we report the first instance of false virologic failure on ART arising from cells with hypermutated proviruses.

METHODS: A 46-year-old female presented with detectable HIV on ART ranging from 439 to 4230 copies/mL. Single genome sequencing (SGS) analysis (p6-Pro-RT) of 4 longitudinal plasma samples obtained over 13 months was performed. To characterize the source of viremia, fractions of plasma after low- (2700 g) and high-speed centrifugation (17 200 g) and total nucleic acid from PBMC were analyzed by SGS including cell-associated HIV mRNA and near-full length (NFL) sequencing of proviral DNA.

RESULTS: SGS (p6-Pro-RT) revealed multiple, identical hypermutated sequences in all low- and one high-speed plasma pellet(s), and in PBMC HIV DNA (p6-Pro-RT and NFL) and cell-associated HIV mRNA. The only non-hypermutated sequences were from the high-speed plasma pellet (4 of 4) and PBMC HIV DNA (1 of 10) at the 13-month time point. Sequencing of *gag* revealed a stop codon at amino acid 211 which would prevent capsid (p24) formation.

CONCLUSIONS: This is the first report of false virologic failure of ART resulting from release of defective HIV nucleic acid into plasma. The source appears to be a large population of cells with identical hypermutated proviruses, i.e. an infected CD4+ T-cell clone that is undergoing cytolysis and release of cellular nucleic acid including HIV DNA and mRNA into plasma. Production of viral proteins and packaging of viral genomes is a highly unlikely source given the hypermutated genome with at least one stop codon in gag-p24. Release of cellular nucleic acids into plasma may be an underappreciated cause of false virologic failure.

"FALSE ART FAILURE" FROM IDENTICAL HYPERMUTATED HIV NUCLEIC ACID IN PLASMA

JC Botha[1], K Steegen[2,3], L Hans[3], A Karstaedt[2], S Carmona[2], D Reddy[2,4], MF Kearney[5], JW Mellors[6], GU van Zyl[1,3]

1 Stellenbosch University, Cape Town, South Africa; 2 University of the Witwatersrand, Johannesburg, South Africa; 3 National Health Laboratory Service, Johannesburg, South Africa; 4 Chris Hani Baragwanath Hospital, Johannesburg, South Africa; 5 HIV Dynamics and Replication Program, National Cancer Institute, Frederick, Maryland, United States of America; 6 Division of Infectious Diseases, Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, United States of America

00373

Background and preliminary data

HIV-1 viral load (VL) monitoring is premised on an elevated VL indicating HIV-1 replication due to drug resistance or poor adherence.

A 46 year old HIV-1 positive adult female patient on long term cART presented with virological failure.

Drug resistance testing by Sanger population sequencing performed on 3 separate samples over a 3-month period, all yielded identical hypermutated sequences.

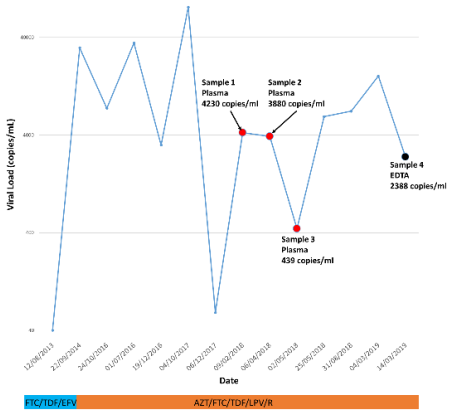
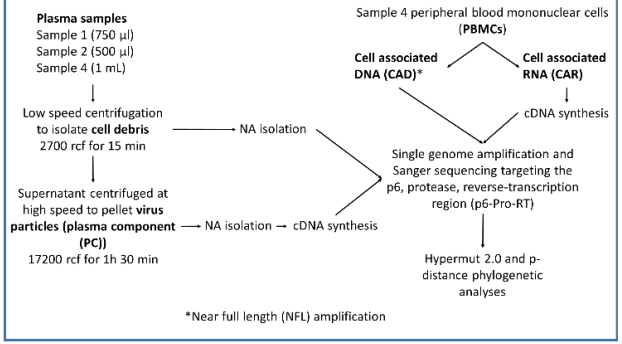


Figure 1: Log scale viral load data plot, with ART regimens. Plasma samples 1, 2 and 3 are indicated by red dots. EDTA sample 4 taken for further investigation is indicated by black dot.

Methods



An HIV-1 positive adult female patient on long term cART presented with virological failure. Investigation revealed this case of false cART failure is caused by a large cell population continuously releasing identical defective HIV DNA and RNA, detectable in plasma for over 13 months.

Aim

To investigate the apparent virological failure by determining the source of HIV genomic material: distinguishing cell debris and virus particles.

Results

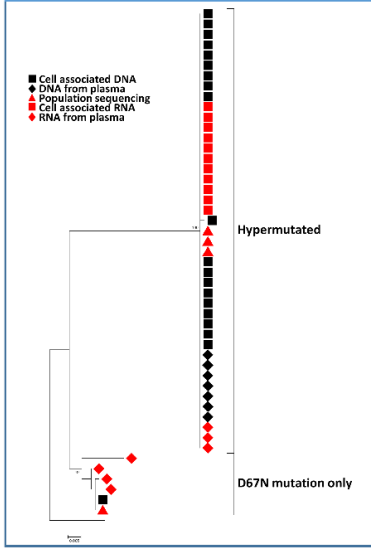


Figure 2: P-distance phylogenetic tree of all nucleotide sequences generated for the p6-Pro-RT region. Plasma population sequences from drug resistance testing is included in phylogenetic tree as reference.

Results - Continued

The APOBEC hypermutated strain is highly transcribed into RNA, based on the endpoint dilutions of single genome sequencing, cell associated RNA (1:2187) was found to be about 81 times more concentrated than cell associated DNA (1:27).

Partial gag sequences of NFL products indicate a stop codon in p24, resulting in truncated p24 capsid protein.

Table 1: Summary of PCR and sequencing results indicating detected mutations

Sample	Component	Sequence results	Mutations detected
Sample 1	Plasma	Population drug resistance seq	APOBEC
	Cell debris	p6-Pro-RT SGS	APOBEC
	Virus pellet	cDNA p6-Pro-RT SGS	APOBEC
Sample 2	Plasma	Population drug resistance seq	APOBEC
	Cell debris	p6-Pro-RT SGS	APOBEC
	Virus pellet	cDNA p6-Pro-RT negative	PCR negative
Sample 4 plasma	Plasma	Population drug resistance seq	D67N
	Cell debris	p6-Pro-RT SGS	APOBEC
	Virus pellet	cDNA p6-Pro-RT SGS	D67N
Sample 4	PBMC	CAD p6-Pro-RT SGS	1/10 – D67N 9/10 – APOBEC
		CAD NFL (p6-Pro-RT region)	APOBEC
		CAR p6-Pro-RT SGS	APOBEC

Discussion and Conclusions

19/20 HIV-1 DNA and 11/11 HIV-1 cellular RNA sequences from PBMCs are identical to APOBEC hypermutated sequences from cell debris and plasma component.

This suggests a major APOBEC mutated population consisting of identical p6-Pro-RT sequences (possible CD4+ T-cell clone).

The APOBEC strain is highly transcribed, as detected by CAR-SGS (endpoint dilution 1:2187).

Production of viral proteins and packaging of viral genomes are highly unlikely given the hypermutated genome with at least one stop codon in gag-p24.

The cause of "virological failure" appears to be a large population of cells with identical hypermutated proviruses undergoing cytolysis, releasing cellular nucleic acid including HIV DNA and mRNA into plasma.

Release of cellular nucleic acids into plasma may be an underappreciated cause of false virological failure and likely influence patient care.

Acknowledgements

Study participant, clinical and nursing staff.

Funding received from: South Africa – U.S. Programme for Collaborative Biomedical Research, NIH grant U01CA200441, the South African Polio Research Foundation (PRF) and GvZ from U01CA200441.

Contact details:

JC Botha
jbotha26@gmail.com

Prof Gert U van Zyl
guvz@sun.ac.za

Addendum B – ISSPA primers

Table of DPO and standard primers used for integration site specific proviral amplification (ISSPA). Primers for ACH2 and all selected patient integration sites are included. Patient primer locations are indicated according to the chromosome (Chr) location of the Hg19 genome. (Inosine = **I**)

Primer name	Orientation	Sequence (5'-3')	Location based on Hg19
ACH2			
DPO_ACH2_5O	Forward	ACT TAC TTG AGC ATA ACG TCA **III IIT** GCC AC	131 bp upstream of integration
ACH2_5O	Forward	CAT AAC GTC AGA TTC TGC C	133 bp upstream of integration
DPO_ACH2_5I	Forward	CAT GTG ATT TAG TAT ACC TGG **III IIA** TGA CC	52 bp upstream of integration
ACH2_5I	Forward	GTG ATT TAG TAT ACC TGG AG	61 bp upstream of integration
ACH2_5J	Forward	AAT TAT TAT CTT TAT TAC TGC TAG	5' Junction
ACH2_3J	Reverse	AAG AAT GCT TTG TAA TTG GAA G	3' Junction
ACH2_3I	Reverse	AAT CCA AAT ATG AGG AAC AG	30 bp downstream of integration
DPO_ACH2_3I	Reverse	GAA CTT CAA TCT CTA AAA TTT AAA TC**I III I**AT GAG G	35 bp downstream of integration
ACH2_3O	Reverse	CAC ATT TTG AGA CAT GAG G	90 bp downstream of integration
DPO_ACH2_3O	Reverse	GTA AGA ACA TGC ATA TCA CAT TTT **III II**A TGA GG	90 bp downstream of integration
338206			
DPO_ C1_24294547_5O	Forward	AGC ACC TTT CCT CTC TCA CTG AA**I III I**AT CTC TCA	Chr1_24 294 391 - 427
C1_24294547_5O	Forward	TCT CTC ACT GAA GCA TTA TCT CTC	Chr1_24 294 403 - 426
DPO_ C1_24294547_5I	Forward	AGG TGA AGT GTG TCA GCC CT**I III I**GA GTG TTG	Chr1_24 294 495 - 528
C1_24294547_5I	Forward	TGT GTC AGC CCT CAT TTG AGT G	Chr1_24 294 503 - 526
C1_24294547_5J	Forward	ACC ATT CTA TTT ATC ATT GAG CTG GAA G	Chr1_24 294 530 - Junction
C1_24294547_3J	Reverse	GTA TGA CTG GGC TCA TGC TAG AG	Junction - Chr1_24 294 562
C1_24294547_3I	Reverse	GAT CCT ACT GAT ACT TGA TGA GG	Chr1_24 294 641 - 664
DPO_ C1_24294547_3I	Reverse	ACC CAT GGT AGT TTG GAT CCT ACT G**II III** TGA TGA GG	Chr1_24 294 641 - 679
C1_24294547_3O	Reverse	ACA CCT AAC AGC TTC AAA AGG AC	Chr1_24 294 694 - 717
DPO_ C1_24294547_3O	Reverse	CCT AAA GCG ACA CAC CTA ACA GCT T**II III** GGA CTA TAT TAC	Chr1_24 294 686 - 728
C19_1423908_5O	Forward	GTT TCA TCA CCG ATG CTG TGT TCC	Chr19_1 423 699 - 722
C19_1423908_5I	Forward	GTT GCT TCT AGA TCC TCC CGG T	Chr19_1 423 770 - 791
C19_1423908_5J	Forward	CTT CGG GGC CTG GAA GG	Chr19_1 423 898 – Junction

Primer name	Orientation	Sequence (5'-3')	Location based on Hg19
C19_1423908_3J	Reverse	CGT GCT GGC CCT GCT AG	Junction - Chr19_1 423 914
C19_1423908_3I	Reverse	GGG ACA CAC TCT CTG GGA GG	Chr19_1 424 010 - 029
C19_1423908_3O	Reverse	ACA GAC CCG TCT CAG AAG ATG C	Chr19_1 424 047 - 068
C11_73469151_5O	Forward	GTT GCA TTC TTC TGT GGG	Chr11_73 468 991 - 9 008
C11_73469151_5I	Forward	CTT TGA AAG CTG AAT TGT GG	Chr11_73 469 097 - 116
C11_73469151_5J	Forward	CAT GAA GAG TTA CTG CTA G	Chr11_73 469 139 - Junction
C11_73469151_3J	Reverse	CTT TTG ATG TAA CTG GAA GG	Junction - Chr11_73 469 163
C11_73469151_3I	Reverse	TGT CTA TAG TTT TCT TAT GCG	Chr11_73 469 184 - 204
C11_73469151_3O	Reverse	ACT ATA TTG TCA AGT AAT GCC	Chr11_73 469 253 - 273
C1_231081054_5O	Forward	GAT CTG GGT GGA AAT CC	Chr1_231 080 956 - 972
C1_231081054_5I	Forward	TTT CTC ATC TGT GAA CTG G	Chr1_231 080 999 - 1 017
C1_231081054_5J	Forward	GGT CAA TGA ACA GTG CTA G	Chr1_231 081 047 - Junction
C1_231081054_3J	Reverse	AAC TGC TAC TGT TTG GAA G	Junction - Chr1_231 081 065
C1_231081054_3I	Reverse	GTT ACG ATT GCC ATA GCG	Chr1_231 081 098 - 115
C1_231081054_3O	Reverse	GGT TAG TGC AAT TTA TGG C	Chr1_231 081 163 - 181
C1_153861271_5O	Forward	ACG TTT ATC TCT TTT GCC AC	Chr1_153 861 152 - 171
C1_153861271_5I	Forward	GGC CTG GCT TTT TGC TG	Chr1_153 861 203 - 219
C1_153861271_5J	Forward	CAA GAG TAG AAA TCT GCT AG	Chr1_153 861 255 - Junction
C1_153861271_3J	Reverse	ATA AGC TGG GAT TTT GGA AG	Junction - Chr1_153 861 283
C1_153861271_3I	Reverse	GAT GAG AAG CTT AAT GTA GAG	Chr1_153 861 313 - 335
C1_153861271_3O	Reverse	AAG ATA AGA GCA ACC TTC AC	Chr1_153 861 347 - 366
340116			
C17_58684892_5O	Forward	CCT CAG CTT CCC GAG TAA CTG G	Chr17_58 684 757 - 778
C17_58684892_5I	Forward	AAC TGG GAT TAC ACG CAT GTG C	Chr17_58 684 773 - 794
C17_58684892_5J	Forward	GTC TTG AAC TCC CGA CTG CTA G	Chr17_58 684 877 - Junction
C17_58684892_3J	Reverse	ATC ACT TGA GGT CGG TGG ATG	Junction - Chr17_58 684 902
C17_58684892_3I	Reverse	GTG CGG TGG CTC ACG C	Chr17_58 684 940 - 955
C17_58684892_3O	Reverse	TCA CAC CAG AGA TCA CCC TTT GC	Chr17_58 685 126 - 148
C6_13641182_5O	Forward	CAA AAT GTA GAA TGG TGG TTT CCA GGG	Chr6_13 640 938 - 946
C6_13641182_5I	Forward	GGA GTC GGA TGG TGG TGA TGG	Chr6_13 641 032 - 052
C6_13641182_5J	Forward	TAG ACA CCC ATC TTC TAA ATA AGT GCT AG	Chr6_13 641 160 - Junction
C6_13641182_3J	Reverse	GTT TAA AAA TGT GAA AAT CTT ATT GGA TGG G	Junction - Chr6_13 641 200
C6_13641182_3I	Reverse	GCA GGA AGA ATG GGA GAA GCC	Chr6_13 641 499 - 520
C6_13641182_3O	Reverse	CAG AAA TTG GTA TTA GCA GGA AGA ATG GG	Chr6_13 641 506 - 534

Primer name	Orientation	Sequence (5'-3')	Location based on Hg19
C18_20529591_5O	Forward	AAG AGC TCT GCA GAG AAA ATT CTG G	Chr18_20 529 398 - 422
C18_20529591_5I	Forward	ACA CGG TGG AGC TCT TAG AAT TAT CC	Chr18_20 529 506 - 531
C18_20529591_5J	Forward	CCT TTA GAG ATG CAC AAA TGG AAG G	Chr18_20 529 574 - Junction
C18_20529591_3J	Reverse	GTG AAG AAT TCT TCT AGT CTT TGT TGC TAG	Junction - Chr18_20 529 610
C18_20529591_3I	Reverse	ACC TAA GTG CCA GAC TCA CCG	Chr18_20 529 675 - 695
C18_20529591_3O	Reverse	AGG TGA GAG AAG GCT GTG GG	Chr18_20 529 779 - 798
C2_73620016_5O	Forward	CTA TTA TGT ATA ATG CTA CCG	Chr2_73 619 909 - 929
C2_73620016_5I	Forward	GGA TTC ATT TCT CTT GGA TG	Chr2_73 619 965 - 984
C2_73620016_5J	Forward	TTT TGG GAC ATA TGT GCT AG	Chr2_73 620 003 - Junction
C2_73620016_3J	Reverse	CAG AAT TCC ATA TTG GAA GG	Junction - Chr2_73 620 024
C2_73620016_3I	Reverse	GCC AGA TTG TTT TCC AAA G	Chr2_73 620 043 - 061
C2_73620016_3O	Reverse	GGA ACT CTC ATA CAC TGC	Chr2_73 620 076 - 093
C18_6467385_5O	Forward	CAC CTC AGC CTC CC	Chr18_6 467 304 - 317
C18_6467385_5I	Forward	GCC ACC ATT CCC AGG	Chr18_6 467 342 - 356
C18_6467385_5J	Forward	GGG AGA CAC TGG AAG G	Chr18_6 467 377 - Junction
C18_6467385_3J	Reverse	AAG ACC GTG TCT GCT AG	Junction - Chr18_6 467 391
C18_6467385_3I	Reverse	CAG GGA GTT CAA GAC C	Chr18_6 467 409 - 424
C18_6467385_3O	Reverse	CAG CAC TTT GCA AGG C	Chr18_6 467 449 - 464
C16_17134060_5O	Forward	CTG TTT TTC ACT CCT GGC	Chr16_17 133 978 - 995
C16_17134060_5I	Forward	AAA CAT TAC CTG TCT CCA C	Chr16_17 134 015 - 033
C16_17134060_5J	Forward	TTC AAG GAA ACT GGA AGG G	Chr16_17 134 055 - Junction
C16_17134060_3J	Reverse	TAG TTT GGG TTT CTG CTA G	Junction - Chr16_17 134 068
C16_17134060_3I	Reverse	CTT TGA ACC CAA ATC CCG	Chr16_17 134 093 - 110
C16_17134060_3O	Reverse	CAG GGT CAC AAG GGC	Chr16_17 134 135 - 149
332406			
C18_18609870_5O	Forward	ATT TGA TGT GGA CCA AGT CTT CCA G	Chr18_18 609 603 - 627
C18_18609870_5I	Forward	GGC TTG AAT GTT ACA AGG GGT ACA TC	Chr18_18 609 726 - 751
C18_18609870_5J	Forward	CAA AAT ACT AAA AAA TTA TTT TAG TGG AAG GG	Chr18_18 609 847 - Junction
C18_18609870_3J	Reverse	AGA ACA TGA GGT TAT TGT ACC TAA ATG CTA G	Junction - Chr18_18 609 890
C18_18609870_3I	Reverse	TGG CTA AAT GTG TTG TTT CGT TCC	Chr18_18 609 991 - 10 014
C18_18609870_3O	Reverse	GAA TGG TGG AAA AGC CCT AAT ATC C	Chr18_18 610 048 - 072
C14_68759177_5O	Forward	GTT GGC TTT CCC AGT GGT CC	Chr14_68 759 006 - 025
C14_68759177_5I	Forward	TGC TGA GTT CTT TTG GAC TAA AAA GCG	Chr14_68 759 050 - 076

Primer name	Orientation	Sequence (5'-3')	Location based on Hg19
C14_68759177_5J	Forward	GTA AAT CTG CAA AAT TTT CTC CAC TGC TAG	Chr14_68 759 154 - Junction
C14_68759177_3J	Reverse	GGG AGG GTG GAT GGA AGG G	Junction - Chr14_68 759 183
C14_68759177_3I	Reverse	CCA GAC CCA CAG CAT TCT GC	Chr14_68 759 341 - 360
3O2_68759177_3O	Reverse	TGG AAG AGC TAT GCA TGT CTA CCC	Chr14_68 759 358 - 382
C6_32786955_5O	Forward	CAT CAT CTT TCA GCA TGA AAT GTG CC	Chr6_32 786 809 - 834
C6_32786955_5I	Forward	TGA AAT GTG CCC CTG ATT TGC C	Chr6_32 786 824 - 845
C6_32786955_5J	Forward	TCC AGG GCC ATG ACT GCT AG	Chr6_32 786 942 - Junction
C6_32786955_3J	Reverse	GCT AGA AGT TCA AGG TCA TTG GAA GG	Junction - Chr6_32 786 969
C6_32786955_3I	Reverse	AGC CTT ATG CTC TCT GTC TTA GTC C	Chr6_32 787 047 - 071
C6_32786955_3O	Reverse	TCT GAA GAA GTG TAC CCT TCT TCC C	Chr6_32 787 075 - 099
339266			
C2_241353495_5O	Forward	ACC TCT GCC CAT TCT TAG AGA AGC	Chr2_241 353 326 - 349
C2_241353495_5I	Forward	CAG ATG GAT CCT GAG TGC TCT GG	Chr2_241 353 393 - 415
C2_241353495_5J	Forward	TCC CCT TCC ATG CTG TGC TAG	Chr2_241 353 481 - Junction
C2_241353495_3J	Reverse	GTT TGG GCT ACA GCA TGG AAG GG	Junction - Chr2_241 353 505
C2_241353495_3I	Reverse	TCA TTT TGG ACA CTG ATG AGG TTT GG	Chr2_241 353 614 - 639
C2_241353495_3O	Reverse	AGC CTG GGC TTC CCT AAG C	Chr2_241 353 661 - 679
C9_137214293_5O	Forward	ACG GGG AAA CCA GAG TCT TGG	Chr9_137 214 095 - 115
C9_137214293_5I	Forward	AGC TCA GGT TCC TCC TCT CCG	Chr9_137 214 190 - 210
C9_137214293_5J	Forward	GGG GCT CTT TGG TGT GGA AGG	Chr9_137 214 280 - Junction
C9_137214293_3J	Reverse	CCA GGG GCC ACC ATG CTA G	Junction - Chr9_ 137 214 301
C9_137214293_3I	Reverse	CCA AAG AGT TCA GGT GAG TGA GGC	Chr9_ 137 214 379 - 402
C9_137214293_3O	Reverse	TAC TTC CAG AAC TCT GGG GAG C	Chr9_ 137 214 419 - 440
334696			
C6_135525218_5O	Forward	GCT TCC CTA AAT TAT TTA AAG C	Chr6_135 525 113 - 134
C6_135525218_5I	Forward	ATA GCC CAA TAC CCA GCC ACC	Chr6_135 525 150 - 170
C6_135525218_5J	Forward	TGC AGA GGG AAA CTT CTG GAA GG	Chr6_135 525 203 - Junction
C6_135525218_3J	Reverse	ACT TCC TCT AGG GTT ATG AAG TTG CTA G	Junction - Chr6_135 525 235
C6_135525218_3I	Reverse	CTT TAA GGT AAT GGG CAG ACT CAA AGG	Chr6_135 525 309 - 335
C6_135525218_3O	Reverse	TGG AAG GAG AAA CTA TCA CAA GAA TGG G	Chr6_135 525 342 - 369

Primer name	Orientation	Sequence (5'-3')	Location based on Hg19
C6_26227488_5O	Forward	GGT CTT GTG AAT TGG AGA TTC AGT GC	Chr6_26 227 338 - 363
C6_26227488_5I	Forward	GCG TGA GGA AAA CAG CCT AAC TAT CC	Chr6_26 227 361 - 387
C6_26227488_5J	Forward	AAT GTT GTT TCC TGT ATC CTG TAT GCT AG	Chr6_26 227 467 - Junction
C6_26227488_3J	Reverse	GCT AGA ATC TTA AAG AAA GTA CAG TGG AAG G	Junction - Chr6_26 227 507
C6_26227488_3I	Reverse	AAA AGA TGC ATT CTG TGC CCA CC	Chr6_26 227 573 - 595
C6_26227488_3O	Reverse	AGA GGC TTT GTT TTA GGC AAA TGA CC	Chr6_26 227 628 - 653
335106			
C6_45580207_5O	Forward	TAT TGA AAG CTC CCC AAG GAG AGG	Chr6_45 579 989 - 80 012
C6_45580207_5I	Forward	CCA AGG AGA GGG GCT ATG C	Chr6_45 580 002 - 020
C6_45580207_5J	Forward	TGG TCC TCC TTG ATT CCT GCT AG	Chr6_45 580 191 - Junction
C6_45580207_3J	Reverse	TAG GGA TGA AGG AGG AAT TGG ATG G	Junction - Chr6_45 580 220
C6_45580207_3I	Reverse	CCA GTG GAA TAA TAC TCA GCA GTA AAG AGG	Chr6_45 580 383 - 412
C6_45580207_3O	Reverse	AGG TGT CAA ATA CCT GTC CAG TGG	Chr6_45 580 406 - 429
C4_54281522_5O	Forward	GTT GTT ATG CTT TTG TTA ATG ATC TAC CC	Chr4_54 281 382 - 410
C4_54281522_5I	Forward	GTT AAT GAT CTA CCC TCC AGT GAA TTT TGG	Chr4_54 281 395 - 424
C4_54281522_5J	Forward	GAA ATA ATA ACT TCA TCT TTA ATG GAT GG	Chr4_54 281 502 - Junction
C4_54281522_3J	Reverse	GTT TCT CAT TGC ATC CCT TAA ATG CTA G	Junction - Chr4_54 281 539
C4_54281522_3I	Reverse	GTC TGT ATT TTG AAG TCA GGT TAT CAA AGC	Chr4_54 281 595 - 624
C4_54281522_3O	Reverse	CCT CAA ATG ATA AAT TAG GAA AGT ATC TTG G	Chr4_54 281 624 - 654
337286			
C17_40421775_5O	Forward	CAA AGA AAA TTC AAC CCT GGT CAA AGG	Chr17_40 421 619 - 645
C17_40421775_5I	Forward	GTT GAC AAT CCT GGA AGC TAA CAT CC	Chr17_40 421 651 - 676
C17_40421775_5J	Forward	CAC TAC GCT TTA GGG TTG TGC TAG	Chr17_40 421 758 - Junction
C17_40421775_3J	Reverse	GGA AAC AGA ATT CAC AAC CTG GAA G	Junction - Chr17_40 421 789
C17_40421775_3I	Reverse	TGT ATG TCT AGA GCT CTG CAA AGG C	Chr17_40 421 888 - 912
C17_40421775_3O	Reverse	GCC TAT TTG GGC GAA TTT TTA AGG C	Chr17_40 421 924 - 948
C19_53455758_5O	Forward	GCT TGA ACT CAG GAT GCA GAG GTT GC	Chr19_53 455 538 - 563
C19_53455758_5I	Forward	GTC TGT TCA ACA GAG GGA GAC TCT GTC TG	Chr19_53 455 594 - 622

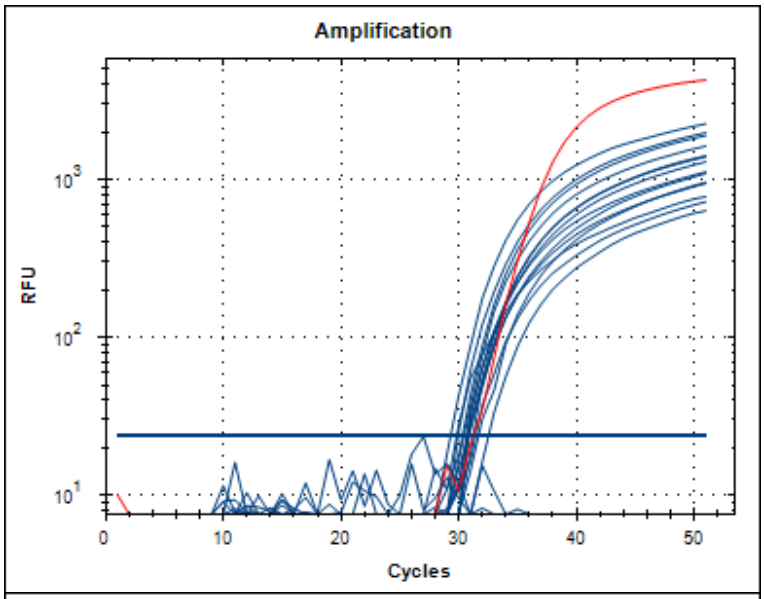
Primer name	Orientation	Sequence (5'-3')	Location based on Hg19
C19_53455758_5J	Forward	CCA GGG ACA TCT CTT CCT AGT GGA AGG	Chr19_53 455 739 - Junction
C19_53455758_3J	Reverse	CAA TTG AAC ATC ATC CTG CTC TAG GTG CTA G	Junction - Chr19_53 455 778
C19_53455758_3I	Reverse	CAT TGT TTT CTG CGG TGA TGA CCC TCC	Chr19_53 455 864 - 890
C19_53455758_3O	Reverse	ATT CTC TTT TGT GAT TTT GCC CCA TAC ATG C	Chr19_53 455 906 - 936
337916			
C16_47550997_5O	Forward	AAG TGG GAT TAT TCA AGC ATT TGG G	Chr16_47 550 811 - 835
C16_47550997_5I	Forward	AAC AAA TAA GGG AAA GGA GCT GGG	Chr16_47 550 898 - 921
C16_47550997_5J	Forward	CAT GAC TCT GAC TGA GTG CTA GAG	Chr16_47 550 982 - Junction
C16_47550997_3J	Reverse	GCC TCC TCT CCT CAG TGG AAG	Junction - Chr16_47 551 007
C16_47550997_3I	Reverse	CGG TTA TCA TAA GAT GTT CCA CAG GG	Chr16_47 551 086 - 111
C16_47550997_3O	Reverse	CAA TAT TAG TTG CCA CTA ATT GTA AAC GG	Chr16_47 551 108 - 137
C3_141257185_5O	Forward	TCC CTG ATC TCA AGG GAG GC	Chr3_141 256 874 - 893
C3_141257185_5I	Forward	CAC ACA CAC AAA CTT GTT TTG CTG C	Chr3_141 256 964 - 988
C3_141257185_5J	Forward	CTT TGA AAA TTA AAT TGG TAG TAT GGA AGG G	Chr3_141 257 163 - Junction
C3_141257185_3J	Reverse	AAT TTT CAG ATT TTA AAA GTA TAT ACT ATG CTA GAG	Junction - Chr3_141 257 208
C3_141257185_3I	Reverse	AAG ATA CCT TTG GTT GGT GAA ATT TCA G	Chr3_141 257 262 - 289
C3_141257185_3O	Reverse	GAC ACA AAA CAC ATA CTG TTA GGG TGG	Chr3_141 257 505 - 531
C8_43028875_5O	Forward	TTT CTA TAC CAG TGT GTA ATG AAG TCC	Chr8_43 028 763 - 789
C8_43028875_5I	Forward	GTA ATG AAG TCC ACA CTA GAT TTA GGA GG	Chr8_43 028 778 - 806
C8_43028875_5J	Forward	GAC AGT GGC TGA CCT CTG CTA G	Chr8_43 028 860 - Junction
C8_43028875_3J	Reverse	CGG GAA CAC GAG GTT GGA AG	Junction - Chr8_43 028 884
C8_43028875_3I	Reverse	ACT TCC TAA CCA CAT AAC ACA TGC C	Chr8_43 028 975 - 999
C8_43028875_3O	Reverse	TCT TTT ATA CAC ACG TGT AGG CAC TGG	Chr8_43 029 041 - 068
335836			
C10_103786725_5O	Forward	ACT AAT GCC ACA CAG TTG TGG TG	Chr10_103 786 562 - 584
C10_103786725_5I	Forward	AGT TGT GGT GTC AGT GTG GAT GAT GG	Chr10_103 786 575 - 600
C10_103786725_5J	Forward	CTG GCA TAT AGG GGC CTG CTA GAG	Chr10_103 786 710 - Junction
C10_103786725_3J	Reverse	CAA ATA TTT GTT GGG CCC CTG GAT GG	Junction - Chr10_103 786 739

Primer name	Orientation	Sequence (5'-3')	Location based on Hg19
C10_103786725_3I	Reverse	GGT CTC GAA CTC CTA ACC TCA GG	Chr10_103 786 818 - 840
C10_103786725_3O	Reverse	AGA CGG AAT CTC ACT CTG TCA CC	Chr10_103 787 013 - 035
C3_49116408_5O	Forward	GAG TTG GAG TCT CCC TCT GTC GC	Chr3_49 116 149 - 171
C3_49116408_5I	Forward	TGC AGT GGC ACG ATC TCG	Chr3_49 116 183 - 200
C3_49116408_5J	Forward	GGG ATT ATA GAC GTG ATG GAT GGG	Chr3_49 116 393 - Junction
C3_49116408_3J	Reverse	GGG TGG CTC ACG TGC TAG	Junction - Chr3_49 116 415
C3_49116408_3I	Reverse	GTA GTA GCT GGG ACT ACA GGT GC	Chr3_49 116 539 - 561
C3_49116408_3O	Reverse	AGT GCA TTC GAG TGA TCT CGG G	Chr3_49 116 613 - 634
341862			
C17_42550940_5O	Forward	GGC CTA AAA TCA GTG ATT TTT AAC TTA GC	Chr17_42 550 799 - 827
C17_42550940_5I	Forward	CAG TGA TTT TTA ACT TAG CAA AGT GGG	Chr17_42 550 809 - 835
C17_42550940_5J	Forward	CGT CAT CAA TTA CTG CTG CTA GAG	Chr17_42 550 925 - Junction
C17_42550940_3J	Reverse	CTT AGG CAT GGC AGT TGG ATA G	Junction - Chr17_42 550 950
C17_42550940_3I	Reverse	GGT TTG TAT TGC ACG GTC TGC	Chr17_42 551 073 - 093
C17_42550940_3O	Reverse	GCC CAT ATT AGC TAA GCA GCC	Chr17_42 551 101 - 121
C11_116877612_5O	Forward	GCA AAT CTA GGT ATT TAG CTC AGG C	Chr11_116 877 444 - 468
C11_116877612_5I	Forward	GGA TAA CAA AAA TGC ACA CTG GC	Chr11_116 877 495 - 517
C11_116877612_5J	Forward	CTT GTT ATT AAA TAC TAT TAA TAC ATG GAT AGG	Chr11_116 877 588 - Junction
C11_116877612_3J	Reverse	CCT TTA TGG TAA GTG TAT TGC TAG AG	Junction - Chr11_116 877 625
C11_116877612_3I	Reverse	GGG TGA TAA TCC AGT TTT GAG GC	Chr11_116 877 707 - 729
C11_116877612_3O	Reverse	GCA CGA GCA TGT CCA AAG G	Chr11_116 877 757 - 775
C19_10287814_5O	Forward	GGA CAA AGT AAT CCA GGT GGG	Chr19_10 287 639 - 659
C19_10287814_5I	Forward	GGC TGT AAC TTG ATA ACT GTT GGC	Chr19_10 287 697 - 720
C19_10287814_5J	Forward	GTA AAA ACA GGT AAC GTT TTT GCT AGA G	Chr19_10 287 795 - Junction
C19_10287814_3J	Reverse	ATG AAT AGC ATT TCT GAA AAC TGG ATA G	Junction - Chr19_10 287 830
C19_10287814_3I	Reverse	TGT TTT GGG ATG GAA TTG GTA GG	Chr19_10 287 901 - 923
C19_10287814_3O	Reverse	GGA TTA CAA GGA AAA GCA CCA GG	Chr19_10 288 000 - 022

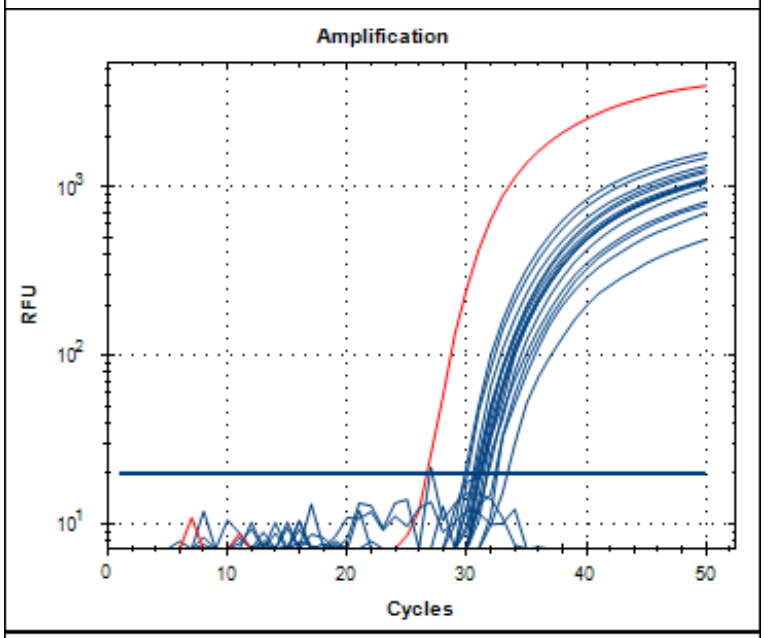
Addendum C – ISSAQ amplification curves

ISSAQ log scale amplification curves obtained from each time point of selected patient integration sites. Positive controls are indicated in red.

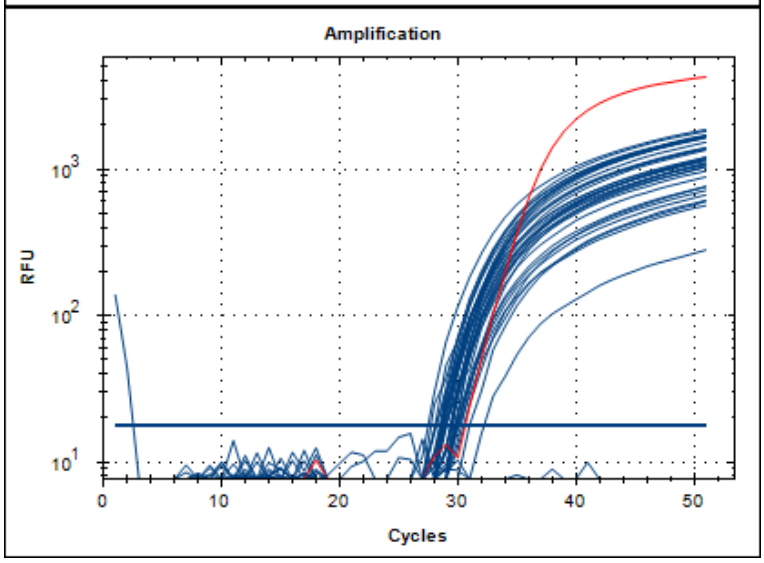
338206
Chr1_24294547



Time point 1
15 Positives

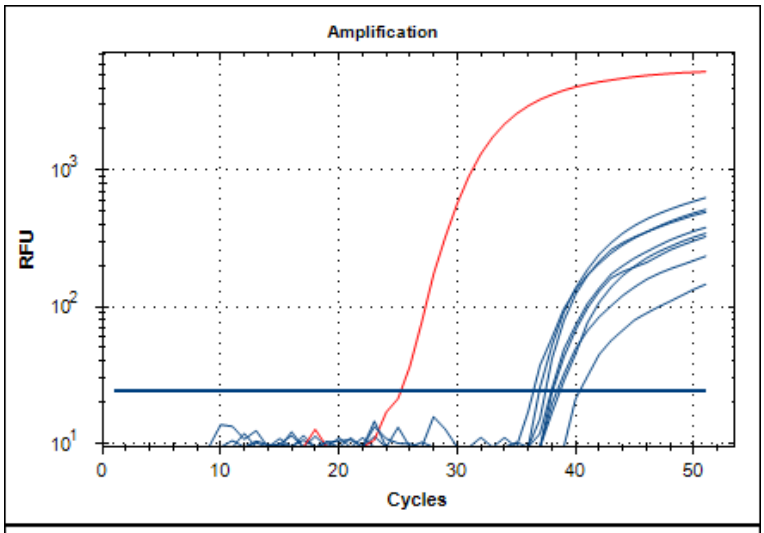


Time point 2
15 Positives



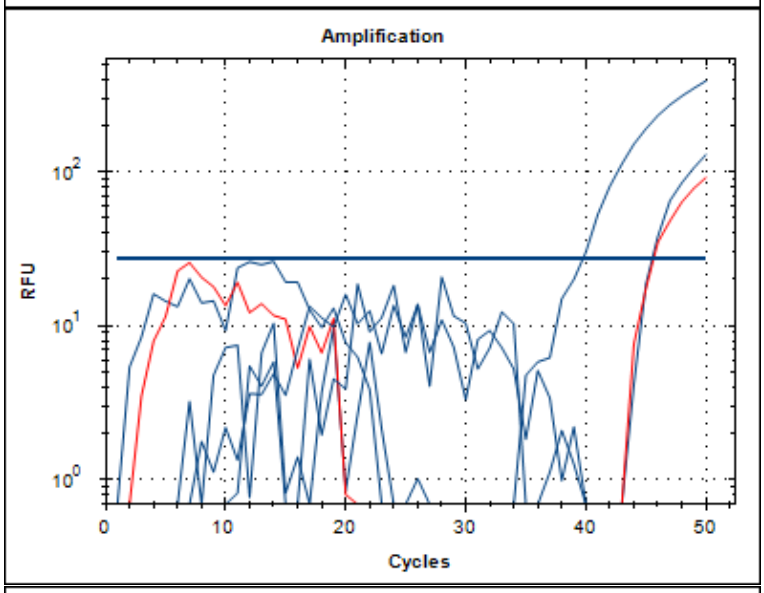
Time point 3
35 Positives

338206 Chr11_73469151



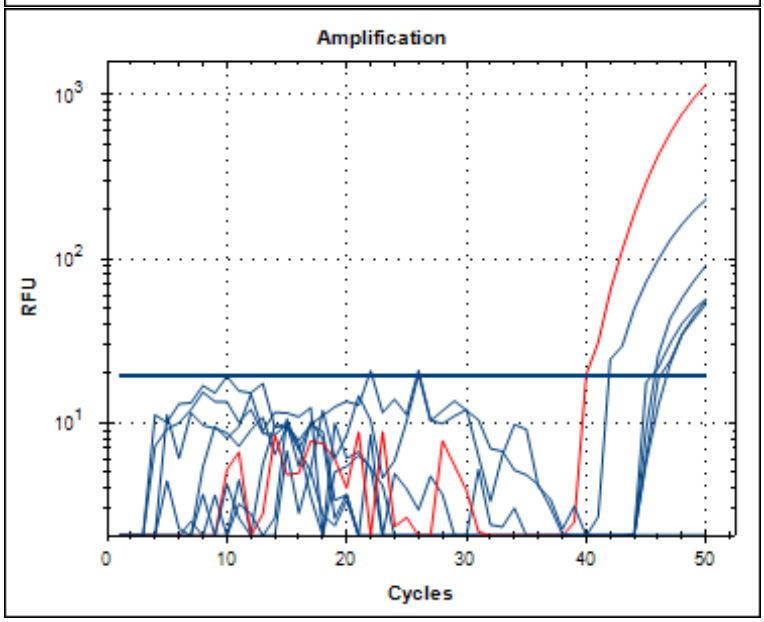
Time point 1

8 Positives



Time point 2

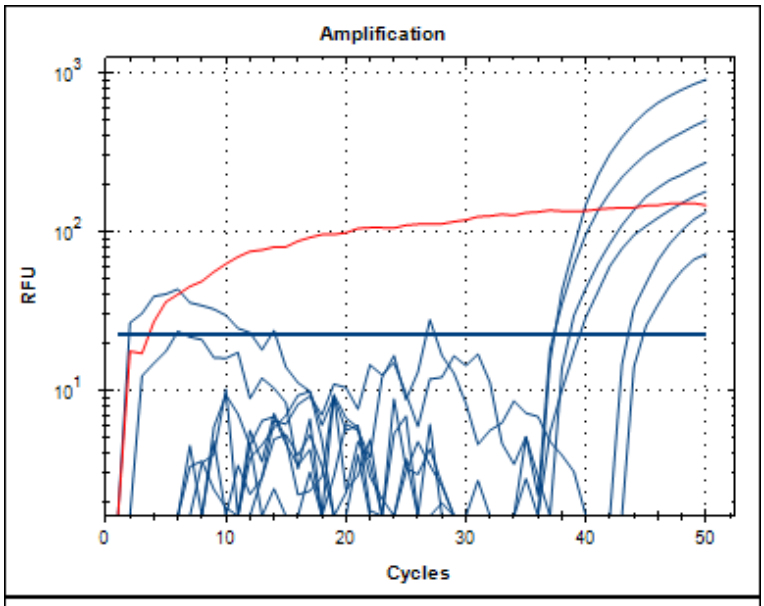
2 Positives



Time point 3

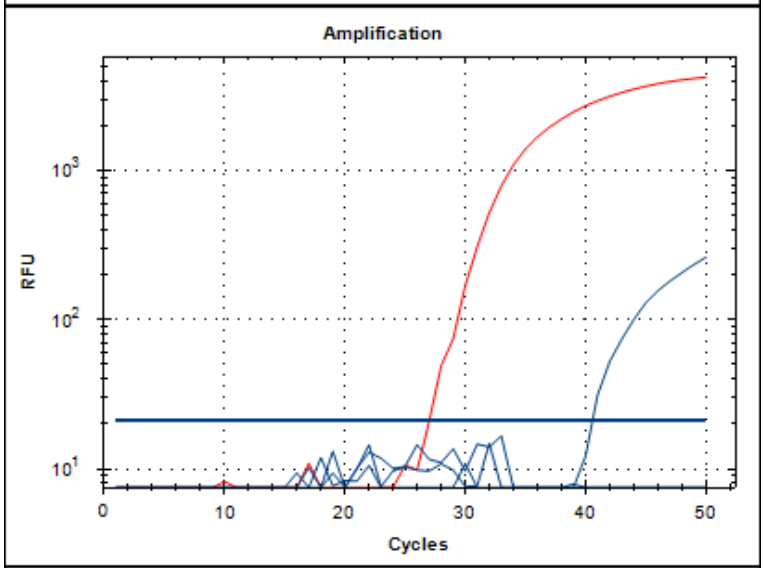
5 Positives

338206 Chr1_231081054



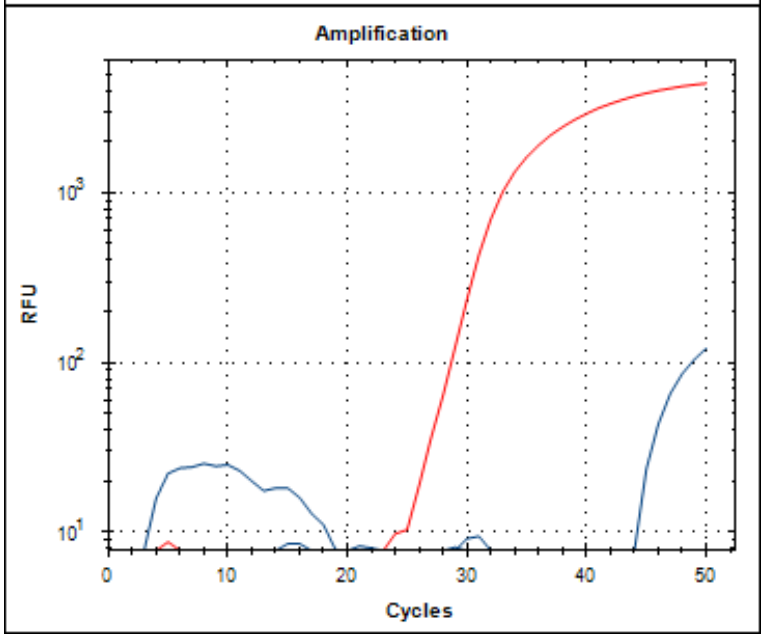
Time point 1

6 Positives



Time point 2

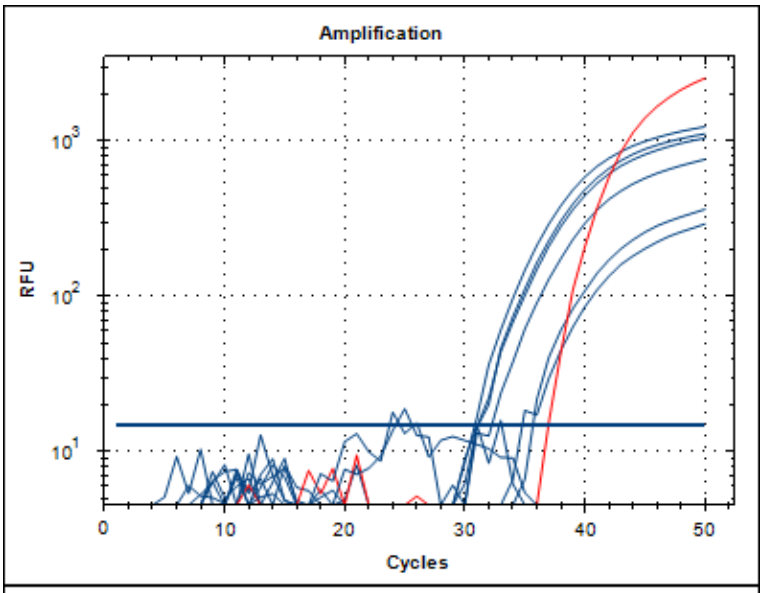
1 Positive



Time point 3

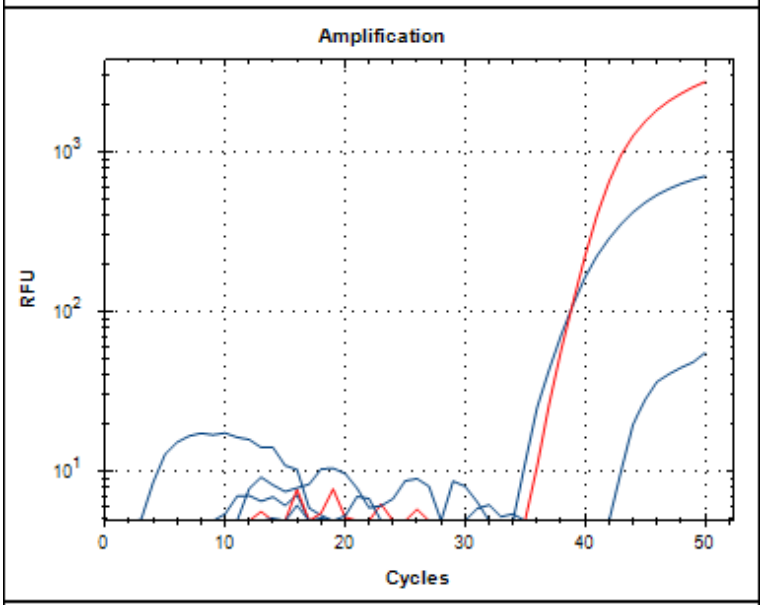
1 Positive

340116 Chr6_13641182



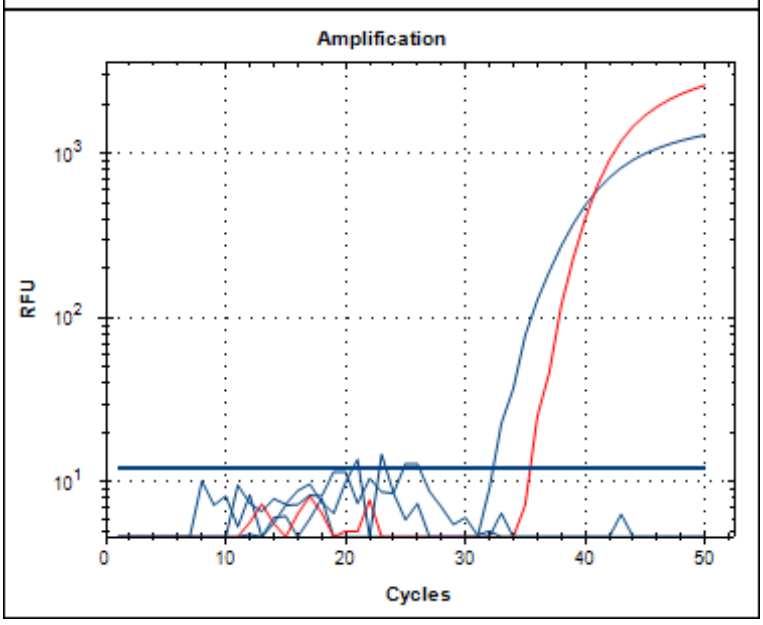
Time point 1

6 Positives



Time point 2

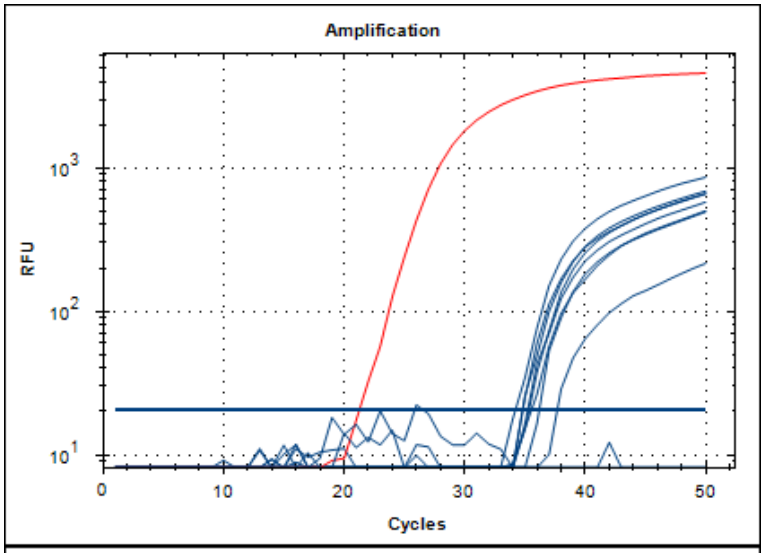
2 Positives



Time point 3

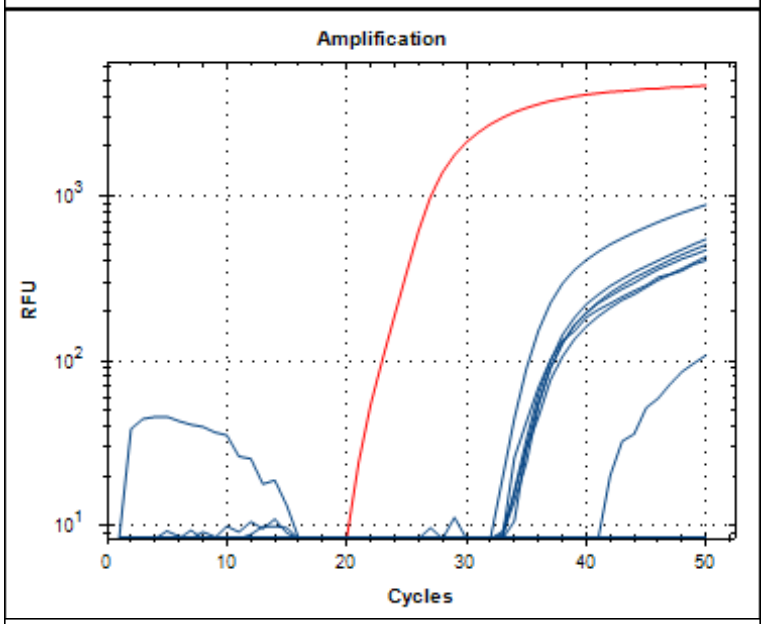
1 Positive

335106 Chr6_45580207



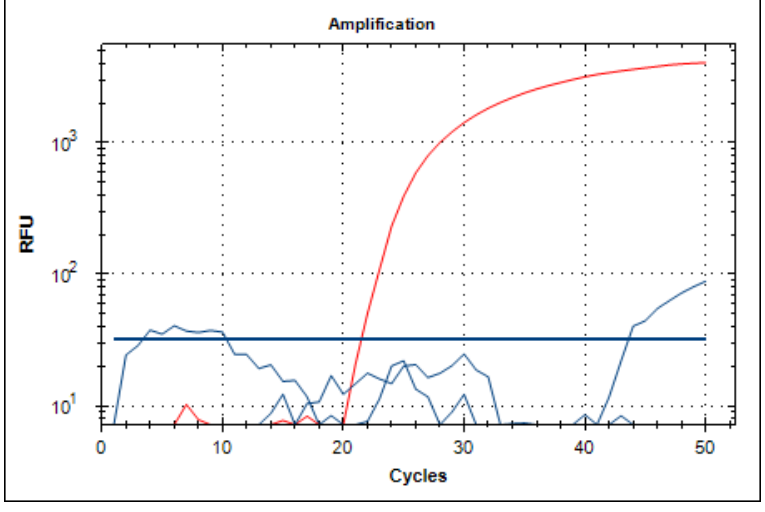
Time point 1

8 Positives



Time point 2

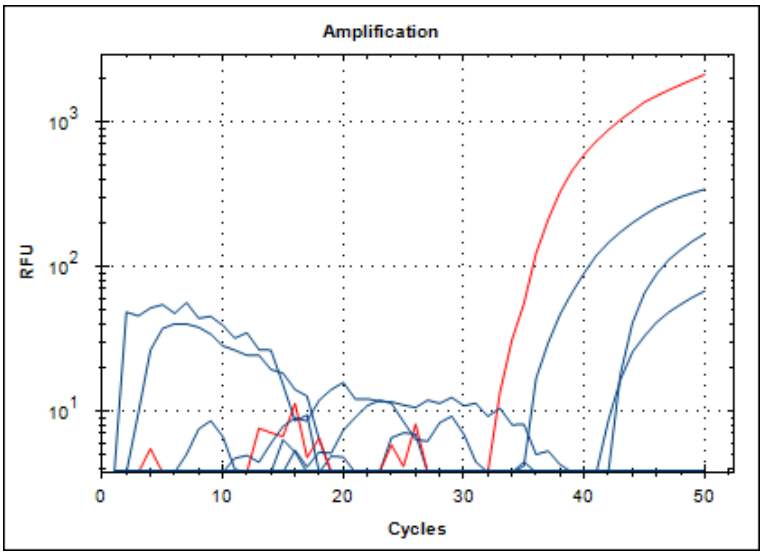
7 Positives



Time point 3

1 Positve

340116 Chr2_73620016



Time point 2

3 Positives